PRACE KÓŁ NAUKOWYCH Politechniki Rzeszowskiej w roku akademickim 2025/2026

Rzeszów 2026

Wydano za zgodą Rektora

R e c e n z e n c i

Ewelina NYCZ, Danuta PROSZAK, Piotr BOGUSZ, Mateusz SZARATA,
Piotr NAZARKO, Mateusz SALACH, Rafał OLIWA, Damian MAZUR,
Marcin GORAL, Karol SIWIEC, Mariusz NYCZ, Bogdan KWIATKOWSKI,
Mariusz DĘBSKI, Jan SADOLEWSKI, Marcin HUBACZ, Paweł DOBRZAŃSKI,
Tomasz KAPUŚCIŃSKI, Olga ŻUCZEK, Dawid KALANDYK, Tomasz KOSSOWSKI,
Mirosław SOŁTYSIAK, Krzysztof LEW, Dorota GŁOWACZ-CZERWONKA,
Kacper SKROBACZ, Patryk ORGANIŚCIAK, Arkadiusz STĘCHŁY, Mariusz RUSZEL,
Bartłomiej ŚNIEŻYŃSKI, Kalina DETKA, Sławomir PORADA, Maciej CIĘŻKOWSKI

R e d a k t o r n a c z e l n y
Wydawnictw Politechniki Rzeszowskiej
dr hab. inż. Lesław GNIEWEK, prof. PRz

R e d a k c j a n a u k o w a
dr inż. Bartosz TRYBUS
Politechnika Rzeszowska
Wydział Elektrotechniki i Informatyki
Katedra Informatyki i Automatyki

dr inż. Bartosz PAWŁOWICZ
Politechnika Rzeszowska
Wydział Elektrotechniki i Informatyki
Zakład Systemów Elektronicznych i Telekomunikacyjnych

Wydruk z matryc dostarczonych przez redaktorów naukowych.
W procesie wydawniczym pominięto etap opracowania językowego.

© Copyright by Oficyna Wydawnicza Politechniki Rzeszowskiej
Rzeszów 2026

p-ISBN 978-83-7934-870-1
e-ISBN 978-83-7934-872-5

Oficyna Wydawnicza Politechniki Rzeszowskiej
al. Powstańców Warszawy 12, 35-029 Rzeszów
https://oficyna.prz.edu.pl

Ark. wyd. 33,33. Ark. druk. 33,50.
Wydrukowano we wrześniu 2026 r.
Drukarnia Oficyny Wydawniczej,
al. Powstańców Warszawy 12, 35-029 Rzeszów
Zam. nr 46/26

SPIS TREŚCI

Bartosz Trybus, Bartosz Pawłowicz
Od redakcji **7**

Maja Jaszowska, Patryk Błoński, Sandra Kudła
Koło Naukowe Elektroniki i Technologii Informacyjnych
Analiza skuteczności platformy T-Pot w detekcji i profilowaniu zagrożeń sieciowych **8**

Norbert Mazur
Koło Naukowe Interakcji Człowiek - Komputer GEST
Implementacja modułu gry typu Tower Defense w środowisku platformy edukacyjnej RUT-AI z wykorzystaniem języka TypeScript **30**

Krystian Sołtys, Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST
Ewolucja inteligentnego urządzenia rehabilitacji ruchowej **45**

Paulina Wojdylak
Koło Naukowe Inżynierii Medycznej X-med
Przegląd systemów przeciwodleżynowych oraz koncepcja własnego rozwiązania **54**

Paweł Dźwięga
Koło Naukowe Inżynierii Medycznej X-med
Projekt konstrukcyjny mechatronicznej dłoni **67**

Daria Ciura, Ewelina Nycz
Studenckie Koło Naukowe Komunikacji Marketingowej „Brief”
Orientacja na klienta jako determinant przewagi konkurencyjnej współczesnych przedsiębiorstw **75**

Julia Ziąbek, Natalia Klocek, Maja Jaszowska, Dominik Dworak, Sebastian Cwynar
Koło Naukowe Elektroniki i Technologii Informacyjnych
Analiza bezpieczeństwa stron wygenerowanych za pomocą AI **87**

Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Krystian Sołtys, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST
Uniwersalny interfejs komunikacji między komputerem a wieloma projektami bazującymi na Raspberry Pi **117**

Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Krystian Sołtys, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST
Projekt i sterowanie robotem samobalansującym **130**

Dominik Dworak, Julia Ziąbek, Maja Jaszowska, Natalia Klocek, Sebastian Cwynar
Koło Naukowe Elektroniki i Technologii Informacyjnych
Identyfikacja układów scalonych bez pełnych oznaczeń **143**

Gabriela Bober, Dorota Głowacz-Czerwonka
Koło Naukowe Studentów Chemii „Esprit”
Wpływ grafitu ekspandowanego na właściwości fizyczne i ogniowe sztywnych pianek poliuretanowych zawierających kwiaty goździka **176**

Dorota Głowacz-Czerwonka, Aleksandra Brożyna
Koło Naukowe Studentów Chemii „Esprit”
Wpływ płatków róży i grafitu ekspandowanego na morfologię i odporność ogniową sztywnych pianek poliuretanowych **187**

Andżelika Gronko
Studenckie Koło Naukowe Inżynierii Środowiska
Ocena satysfakcji klientów przedsiębiorstwa wodociągowego z uwzględnieniem aspektów społecznych, technicznych i ekonomicznych **196**

Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Krystian Sołtys, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST
Robotyczne ramię sortujące **213**

Sandra Kudła
Koło Naukowe Elektroniki i Technologii Informacyjnych
Automatyzacja odczytu ksiąg metrykalnych urodzeń z wykorzystaniem *Google Vision API* **224**

Zuzanna Dörfler, Ewelina Nycz
Studenckie Koło Naukowe Komunikacji Marketingowej „Brief”
Moja marka, moja siła: znaczenie marki osobistej w rozwoju kariery zawodowej **243**

Zuzanna Dörfler, Ewelina Nycz
Studenckie Koło Naukowe Komunikacji Marketingowej „Brief”
Influencer marketing jako narzędzie komunikacji marketingowej i jego wpływ na zachowania konsumentów **258**

Aleksandra Brożyna, Dorota Głowacz-Czerwonka
Koło Naukowe Studentów Chemii „Esprit”
Wpływ płatków róży i grafitu ekspandowanego na właściwości fizyczne sztywnych pianek poliuretanowych **269**

Gabriela Bober, Dorota Głowacz-Czerwonka
Koło Naukowe Studentów Chemii „Esprit”
Wpływ dodatku suszonych kwiatów goździka na właściwości fizyczne i ogniowe sztywnych pianek poliuretanowych **279**

Natalia Klocek, Julia Ziąbek, Maja Jaszowska, Dominik Dworak, Sebastian Cwynar
Koło Naukowe Elektroniki i Technologii Informacyjnych
Wykorzystanie kwantyzacji wektorów uczących do analizy i klasyfikacji danych ewaluacyjnych asystentów nauczania **291**

Patryk Błoński, Maja Jaszowska, Sandra Kudła
Koło Naukowe Automatyków i Robotyków
Analiza działania i implementacja praktyczna transformatora Tesli **318**

Martyna Bielenda, Karolina Kisała, Anna Turek
Koło Naukowe Systemów Złożonych
Automatyzacja analizy stomatologicznych obrazów RTG przy użyciu algorytmów głębokiego uczenia **330**

Anna Turek, Karolina Kisała, Martyna Bielenda
Koło Naukowe Systemów Złożonych
Predykcja produkcji biogazu z uwzględnieniem czynników meteorologicznych w procesie fermentacji metanowej **339**

Karolina Kisała, Anna Turek, Martyna Bielenda
Koło Naukowe Systemow Złożonych
Adaptacja platformy LMS Moodle dla niezależnych ośrodków edukacyjnych z wykorzystaniem zewnętrznych systemów zarządzania **352**

Kajetan Żarski, Mirosław Sobaszek
Koło Podstaw Elektroniki
Model symulacyjny głośnika w obudowie zamkniętej **366**

Marcin Bator
Studenckie Koło Naukowe Informatyków „Kod”
Dostarczanie infrastruktury aplikacji z użyciem terraform na przykładzie chmury Azure i kontenerów Docker **378**

Maja Lis
Koło Naukowo-Badawcze Turystyki „Tuptuś”
Wpływ turystyki kosmicznej na środowisko **394**

Julia Widz
Koło Naukowo-Badawcze Turystyki „Tuptuś”
Metody analizy i oceny efektywności ekonomicznej działalności przedsiębiorstw turystycznych **402**

Karolina Czapka
Koło Naukowo-Badawcze Turystyki „Tuptuś”
Czynniki wpływające na wybór ubezpieczenia turystycznego przez osoby podróżujące **412**

Sebastian Cwynar, Maja Jaszowska, Natalia Klocek, Julia Ziąbek, Dominik Dworak
Studenckie Koło Naukowe Informatyków „Kod”
Metodyka projektowania systemów sterowania animacją szkieletową postaci 3D **426**

Maja Jaszowska, Natalia Klocek, Sebastian Cwynar, Julia Ziąbek, Dominik Dworak
Koło Naukowe Elektroniki i Technologii Informacyjnych
Bezpieczeństwo IoT w systemach krytycznej infrastruktury, sieć lokalna **445**

Eryk Kopciuch
Studenckie Koło Naukowe Informatyków „Kod”
Proceduralna generacja terenu z wykorzystaniem wielowarstwowego szumu Perlina w silniku Unity **461**

Maciej Karczmarz
Studenckie Koło Naukowe Informatyków „Kod”
Zintegrowany potok produkcyjny proceduralnego generowania środowisk trójwymiarowych z wykorzystaniem silnika Unreal Engine 5 oraz narzędzi Gaea i Houdini **467**

Przemysław Lib, Karolina Kielar, Michał Meszko, Kacper Szlachta, Bartosz Dusza, Filip Mordka, Kacper Barański
Koło Naukowe PRz Polycomp
Wykorzystanie recyklatów polimerowych w technologii druku 3D – przegląd właściwości i wyzwań technologicznych **477**

Aneta Miazga
Koło Naukowo-Badawcze Turystyki „Tuptuś”
Strategie optymalizacji wydatków w turystyce studenckiej – analiza postaw pokolenia Z **492**

Mikołaj Lech, Mirosław Sobaszek
Koło Naukowe Elektroników
Model impedancyjny głośnika dynamicznego w obudowie zamkniętej **510**

Małgorzata Żukowicz
Koło Naukowo-Badawcze Turystyki „Tuptuś”
Podatki turystyczne w Polsce i Europie – analiza porównawcza systemów **522**

Bartosz Trybus, Bartosz Pawłowicz

OD REDAKCJI

Niniejsza monografia stanowi efekt pracy studentów zrzeszonych w kołach naukowych działających w Politechnice Rzeszowskiej. Zawarte w niej rozdziały są wynikiem działalności badawczej, projektowej i konstrukcyjnej prowadzonej w wielu obszarach. Udział w pracach kół naukowych pozwala studentom rozwijać zainteresowania, zdobywać doświadczenie oraz podejmować pierwsze próby samodzielnego opracowania wyników. Dla wielu autorów publikacja rozdziału jest pierwszym doświadczeniem związanym z przygotowaniem tekstu naukowego.

Tematyka publikacji jest szeroka. Obejmuje m.in. informatykę, sztuczną inteligencję i cyberbezpieczeństwo, w tym analizę zagrożeń sieciowych, bezpieczeństwo systemów IoT, automatyczny odczyt dokumentów, analizę obrazów RTG oraz rozwiązania chmurowe. Istotną część stanowią także prace z zakresu automatyki, robotyki, elektroniki i systemów wbudowanych, dotyczące m.in. robotów, urządzeń rehabilitacyjnych, Raspberry Pi, modelowania głośników, transformatora Tesli, animacji 3D i proceduralnego generowania środowisk.

W monografii znalazły się również badania z zakresu chemii i inżynierii materiałowej, dotyczące pianek poliuretanowych i recyklatów polimerowych w druku 3D. Uzupełniają je prace poświęcone zagadnieniom środowiskowym i gospodarczym, a także rozdziały z zakresu zarządzania, marketingu i turystyki, obejmujące m.in. orientację na klienta, markę osobistą, influencer marketing, ubezpieczenia, podatki turystyczne i turystykę kosmiczną.

Wszystkie rozdziały zostały poddane recenzji. Recenzenci życzliwie podeszli do oceny prac, uwzględniając ich studencki charakter i fakt, że dla wielu autorów był to pierwszy kontakt z procesem recenzowania publikacji naukowej. Przekazane uwagi pozwoliły udoskonalić teksty pod względem merytorycznym i redakcyjnym. Wśród recenzentów znalazły się także osoby spoza Politechniki Rzeszowskiej.

Monografia jest świadectwem aktywności i różnorodności zainteresowań studentów skupionych w kołach naukowych. Działalność ta stanowi ważne uzupełnienie procesu kształcenia, umożliwiając łączenie wiedzy zdobywanej podczas studiów z pracą projektową, eksperymentalną i badawczą.

Maja Jaszowska, Patryk Błoński, Sandra Kudła
Koło Naukowe Elektroniki i Technologii Informacyjnych

dr inż. Bartosz Pawłowicz
opiekun Koła Naukowego

ANALIZA SKUTECZNOŚCI PLATFORMY T-POT W DETEKCJI I PROFILOWANIU ZAGROŻEŃ SIECIOWYCH

Streszczenie

Niniejszy rozdział prezentuje wdrożenie oraz weryfikację skuteczności platformy T-Pot, stanowiącej środowisko typu honeypot. System ten wykorzystuje technologię konteneryzacji Docker do uruchamiania usług symulujących luki w zabezpieczeniach systemów operacyjnych, baz danych oraz protokołów sieciowych. Celem badawczym była ocena platformy w zakresie gromadzenia danych o wektorach ataków w czasie rzeczywistym. Badania przeprowadzono w izolowanym środowisku sieci lokalnej, co zapewniło kontrolę nad ruchem i wyeliminowało ryzyko dla infrastruktury zewnętrznej. Działania ofensywne generowano z systemu Kali Linux, wykorzystując narzędzia audytowe: Nmap, Nikto, Hydra, Ncrack oraz Metasploit Framework. Symulacje wykazały, że sensory platformy celowo modyfikują reakcje na zapytania. Mechanizmy tarpit spowalniają skanowanie, wyczerpując zasoby napastnika. Emulatory usług masowo akceptują poświadczenia logowania, zachęcając zautomatyzowane skrypty do nawiązywania sesji i wykonywania komend. Kluczowym etapem była analiza danych zgromadzonych przez stos analityczny ELK. Statystyki wskazały na najwyższe natężenie ataków w obrębie usług bazodanowych (port 1433), aplikacji webowych (port 80) oraz usług zdalnego dostępu (port 22). Analiza przechwyconych danych uwierzytelniających potwierdziła, że kompromitacja systemów opiera się na atakach słownikowych, eksploatujących najczęściej używane hasła oraz konta o wysokich uprawnieniach. Uzyskane wyniki dowodzą, że integracja wielu sensorów dostarcza wiarygodnego i dobrego materiału badawczego. Środowisko T-Pot stanowi skuteczne rozwiązanie do detekcji, mapowania aktywności intruzów oraz profilowania zagrożeń sieciowych.

Słowa kluczowe: honeypot, platforma honeynet, wektory ataków, stos ELK, detekcja ataków.

1. Wstęp

Honeypoty to symulowane systemy lub segmenty sieci, które zostały zaprojektowane tak, aby wyglądały na podatne na ataki, wabiąc w ten sposób atakujących do kontrolowanego środowiska, w którym można ich monitorować i analizować [6]. Zastosowanie honeypotów kontrastuje z innymi technikami bezpieczeństwa tym, że nie mają one na celu bezpośredniego zapobiegania włamaniom [1]. Dodatkowo organizacje mogą używać honeypotów, które są systemami-przynętami zaprojektowanymi w celu odciągnięcia atakujących od rzeczywistych zasobów, aby gromadzić informacje o typach ataków uruchamianych przeciwko ich systemom. W ten sposób są w stanie wykryć ataki typu zero-day i zapewnić wgląd w działania oraz motywacje atakujących [6].

Platforma T-Pot jest systemem łączącym wiele zróżnicowanych usług honeypotowych za pomocą technologii kontenerów Docker. T-Pot tworzy realistyczne środowisko dla atakujących, oferując jednoczesną emulację szerokiej gamy protokołów i usług, takich jak SSH i Telnet (np. dzięki honeypotowi Cowrie), usługi internetowe HTTP i HTTPS (wykorzystując m.in. Snare, Glastopf czy Citrixhoneypot), a także ruch SMB, FTP czy bazy danych (głównie za pomocą honeypota Dionaea) [10]. Dzięki temu system zbiera szerokie spektrum danych dotyczących prób włamań. Wykorzystanie takiej platformy pozwala na kompleksową analizę ruchu złośliwego w czasie rzeczywistym. Niniejszy rozdział opisuje proces wdrożenia tego środowiska oraz weryfikację jego skuteczności w testach praktycznych.

Technologia taka jak honeypot jest dość nowym konceptem, przez co ciężko ją jednoznacznie zdefiniować. Chociaż koncepcja honeypotów była używana już w latach 90., jedna z pierwszych i najszerzej akceptowanych definicji w literaturze określa je po prostu jako zasoby bezpieczeństwa, których jedyną wartością jest bycie sondowanym, atakowanym lub kompromitowanym przez intruzów [7]. System ten nie ogranicza się do symulowania tylko jednego protokołu, ale oferuje bardzo szeroki zakres usług uruchamiając jednocześnie jak najwięcej punktów wejścia, by ściśle przypominać rzeczywiste, rozbudowane środowiska. Dzięki takiemu podejściu i zintegrowaniu wielu usług z narzędziami analitycznymi, badacze mogą analizować atakujących na podstawie ich zachowania, techniki oraz zróżnicowanych wektorach ataków [10].

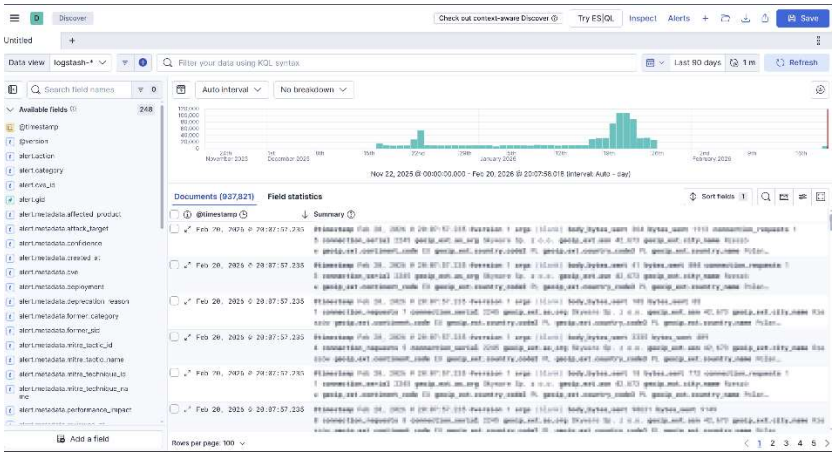
Podstawą technologiczną platformy T-Pot jest technologia Docker, która pozwala na konteneryzację i parawirtualizację poszczególnych usług [8]. Każdy honeypot działa jako oddzielna jednostka, co minimalizuje ryzyko całkowitego przejęcia systemu przez napastnika. Dzięki temu kontenery, choć są izolowanymi środowiskami, zużywają znacznie mniej zasobów sprzętowych w porównaniu do maszyn wirtualnych, zachowując przy tym podobną funkcjonalność. Dzięki zastosowaniu silnika Docker oraz Docker Compose, T-Pot może bezproblemowo i jednocześnie uruchamiać bardzo dużą liczbę honeypotów i narzędzi na tym samym interfejsie sieciowym, maksymalnie i efektywnie wykorzystując przy tym dostępne zasoby sprzętowe [8]. Rysunek 1. przedstawia homepage T-Pota.



Rys. 1. Strona startowa platformy T-Pot

Fig. 1. T-Pot platform home page

Platforma T-Pot wyróżnia się wśród samodzielnych honeypotów ze względu na to, jak łączy ze sobą różne usługi. O ile pojedyncze narzędzia trzeba ustawiać ręcznie i za każdym razem osobno, to T-Pot daje możliwość zarządzania wszystkim z jednego miejsca i automatyzuje konfigurację. Niestety wiąże się to kosztem większego zużycia zasobów systemowych. Automatyzacja procesów i utrzymanie jest tu priorytetem. Architektura T-Pot została zaprojektowana w modelu „zainstaluj i zapomnij" (low maintenance), co osiągnięto dzięki szerokiej automatyzacji zadań administracyjnych. Dzięki temu użytkownik może skupić się na analizie danych, a nie na serwisowaniu poszczególnych komponentów.



Rys. 2. Karta „Odkryj" na platformie T-Pot prezentuje wszystkie zebrane logi w danym przedziale czasowym

Fig. 2. The discover tab on the T-Pot platform presents all collected logs in a given time window

Zarządzanie platformą T-Pot odbywa się na dwa sposoby: z jednej strony jest to bezpieczny dostęp do terminala, z drugiej strony poprzez różne interfejsy webowe. Rozwiązanie to pozwala na pełną kontrolę nad środowiskiem z poziomu przeglądarki. Dla bezpieczeństwa zmieniono standardowe porty usług. Dzięki temu ruch administracyjny jest oddzielony od tego, który generują ataki na honeypoty. W ten sposób łatwiej jest kontrolować, co się dzieje i uniknąć niepotrzebnych zagrożeń. Takie rozgraniczenie kanałów komunikacji jest kluczowe w systemach przeznaczonych do kontaktu ze szkodliwym oprogramowaniem. Rysunek 2. przedstawia kartę odkryj z zebranymi logami.

Uruchomienie kilkudziesięciu kontenerów razem z całym stosem analitycznym Elastic (ELK) na jednej maszynie potrafi mocno obciążyć zasoby sprzętowe. Złożoność systemu wymaga odpowiedniego zaplecza technicznego, szczególnie w zakresie podsystemu pamięci masowej. Dysk SSD jest kluczowy, ponieważ generowana jest duża liczba operacji wejścia/wyjścia przez bazę Elasticsearch i logi honeypotów. Niedobór mocy obliczeniowej lub wolny zapis danych mogą prowadzić do opóźnień w indeksowaniu zdarzeń oraz niestabilności kontenerów.

2. Honeypoty w T-Pot

ADBHoney

ADBHoney to honeypot, symulujący protokół Android Debug Bridge. Jego głównym zadaniem jest monitorowanie prób nieautoryzowanego dostępu do urządzeń mobilnych oraz systemów IoT bazujących na systemie Android. Narzędzie to pozwala na wczesną identyfikację zagrożeń celujących w specyficzne luki w oprogramowaniu układowym [2].

- System nie tylko rejestruje wpisywane komendy, ale fizycznie pobiera i zapisuje pliki, takie jak malware czy skrypty dropperów, przesyłane przez atakujących.
- Nie uruchamia złośliwego kodu, co minimalizuje ryzyko infekcji hosta, lecz może zostać wykryte przez zaawansowane boty.
- Umożliwia wykrywanie nowych kampanii botnetów mobilnych i IoT, zanim zaatakują one realną infrastrukturę korporacyjną.

Beelzebub

Beelzebub wykorzystuje technologię dużych modeli językowych (LLM) do emulacji różnych usług sieciowych. Potrafi prowadzić bardziej wiarygodną interakcję z napastnikiem niż tradycyjne rozwiązania statyczne, co jest kluczowe w walce z systemami automatycznego rozpoznawania pułapek [5].

- Dzięki wykorzystaniu LLM, odpowiedzi serwera nie są statycznymi ciągami znaków, lecz są generowane dynamicznie.
- Dynamiczne odpowiedzi i zróżnicowany czas reakcji utrudniają zautomatyzowanym skanerom odróżnienie Beelzebuba od rzeczywistego systemu produkcyjnego.

CiscoASA

Honeypot CiscoASA służy do emulacji interfejsu urządzeń sieciowych Cisco ASA, koncentrując się na specyficznych podatnościach tych systemów. Pozwala on na zbieranie informacji o próbach eksploatacji błędów w usługach VPN oraz panelach administracyjnych [5].

- Rejestrowany ruch jest wysoce ukierunkowany (targeted), co odróżnia go od masowego skanowania usług takich jak SSH.
- System rejestruje przede wszystkim treść złośliwych pakietów XML, co pozwala na dogłębną analizę struktury exploita.
- Jako narzędzie niskiej interakcji, zwraca statyczne nagłówki, co może pozwolić zaawansowanemu agresorowi na zidentyfikowanie pułapki poprzez analizę stosu protokołów.
- Większość wykrywanych incydentów pochodzi od automatów skanujących sieć w poszukiwaniu starych, niezałatanych systemów i znanych luk z baz takich jak Exploit-DB.

CitrixHoneypot

CitrixHoneypot służy do emulacji podatnych usług dostępu zdalnego, takich jak Citrix ADC czy Gateway. Skupia się na przechwytywaniu prób wykorzystania krytycznych luk w infrastrukturze korporacyjnej [5].

- Narzędzie symuluje specyficzne nagłówki odpowiedzi HTTP serwera, aby zmylić proste skrypty skanujące, choć brak pełnego stosu SSL/TLS może zdradzić naturę pułapki.
- Alarm z tego systemu jest niemal zawsze sygnałem o poważnym zagrożeniu (np. rekonesans przed atakiem ransomware), co odróżnia go od masowego szumu sieciowego.

Conpot

Conpot to honeypot typu ICS (Industrial Control Systems), symulujący systemy automatyki przemysłowej i protokoły sterowników PLC. Pozwala on na monitorowanie prób infiltracji infrastruktury krytycznej [7].

- Emuluje standardy takie jak MODBUS TCP (TCP/502), który jest popularnym standardem przemysłowym. Obsługa funkcji master/slave. IPMI (UDP/623): Interfejs zarządzania sprzętem serwerowym (np. zasilania).
- Ze względu na niski poziom interakcji, służy głównie do wykrywania etapu zbierania informacji (Reconnaissance), nie pozwalając na obserwację fazy niszczenia procesów fizycznych.

Cowrie

Cowrie to jeden z najpopularniejszych honeypotów emulujący usługi SSH oraz Telnet. Służy do rejestrowania sesji interaktywnych oraz przechwytywania plików przesyłanych przez napastników [1].

- Rejestruje każdy krok intruza, w tym wpisywane komendy, próby logowania oraz pobierane skrypty powłoki [7].

- Udostępnia fałszywy system plików, co pozwala na obserwację zachowania napastnika po uzyskaniu rzekomego dostępu do powłoki [10].

DDoSpot

DDoSpot jest wyspecjalizowanym narzędziem do monitorowania i analizy ataków typu Distributed Denial of Service. Skupia się na identyfikacji botnetów wykorzystujących protokoły do wzmacniania ataków (np. NTP, DNS) [1].

- Wymaga ścisłej izolacji i limitowania ruchu (rate-limiting), aby sam nie stał się aktywnym elementem ataku DDoS [1].
- Pozwala na precyzyjne określenie, jakie protokoły są aktualnie nadużywane przez grupy przestępcze do zalewania celów ruchem [3].

Dicompot

Dicompot to honeypot symulujący standard DICOM, wykorzystywany w medycynie do przesyłania i przechowywania obrazów diagnostycznych. Narzędzie to imituje urządzenia takie jak tomografy czy systemy PACS [5].

- Pozwala na odróżnienie przypadkowego skanowania Internetu od ataków celowanych (Targeted Attacks) bezpośrednio w sektor medyczny.
- Umożliwia badanie metod, jakimi napastnicy próbują uzyskać dostęp do poufnych danych pacjentów lub manipulować aparaturą medyczną [9].

Dionaea

Dionaea to honeypot, którego głównym celem jest chwytanie złośliwego oprogramowania wykorzystującego luki w protokołach sieciowych. Narzędzie to emuluje szeroki wachlarz usług, aby skłonić napastnika do przesłania pliku binarnego [7].

- Każda przechwycona próbka malware może być automatycznie przesyłana do zewnętrznych serwisów w celu natychmiastowej analizy [7].
- System skutecznie emuluje protokoły takie jak SMB, FTP, TFTP oraz MSSQL, co pozwala na szeroki zakres detekcji [7].

ElasticPot

ElasticPot służy do symulowania instancji Elasticsearch, które są częstym celem ataków typu ransomware oraz nieautoryzowanej eksfiltracji danych [5].

- Honeypot odpowiada na zapytania REST API w sposób charakterystyczny dla prawdziwego serwera, co pozwala na logowanie prób kradzieży indeksów [7].
- Narzędzie identyfikuje próby wykorzystania znanych luk w starszych wersjach silnika Elastic, rejestrując złośliwe zapytania JSON [5].

Galah

Galah to nowoczesny honeypot HTTP o niskiej interakcji, zaprojektowany pod kątem szybkości i minimalnego zużycia zasobów [1].

- Pozwala na szybkie tworzenie fałszywych punktów końcowych API oraz stron logowania w celu zbierania danych uwierzytelniających [7].

- Dzięki lekkiej architekturze potrafi obsługiwać duży ruch bez wpływu na wydajność całego systemu T-Pot [5].

HellPot

HellPot to narzędzie typu tarpit (pułapka czasowa), zaprojektowane w celu zwalczania złośliwych botów poprzez ich spowalnianie i marnowanie ich zasobów [7].

- HellPot jest dla protokołu HTTP, natomiast Endlessh jest dla SSH; oba to tarpity, jednak HellPot generuje znacznie większy wolumen ruchu sieciowego.
- Dla człowieka jest łatwy do wykrycia przez nieskończone ładowanie strony, jednak dla prostych botów jest zabójczy; zaawansowane skanery z limitami czasowymi mogą go zignorować.

Heralding

Heralding to honeypot typu "credential catcher", którego zadaniem jest zbieranie danych uwierzytelniających przesyłanych przez napastników podczas prób logowania [5].

- Obsługuje dostęp zdalny (SSH, VNC), pocztę (SMTP, POP3, IMAP), usługi web (HTTP/S) oraz bazy danych (PostgreSQL) i FTP.
- Gromadzi ogromne ilości haseł w formie otwartego tekstu (cleartext), co wymaga rygorystycznej polityki retencji i zabezpieczenia logów przed wyciekiem.

Honeytrap

Honeytrap to dynamiczny system wykrywania ataków, który potrafi adaptować się do ruchu przychodzącego na porty, na których nie działają inne usługi [5].

- Posiada zdolność obsługi zarówno ruchu otwartego, jak i szyfrowanego (SSL/TLS), co jest kluczowe w analizie nowoczesnego złośliwego oprogramowania.
- Ze względu na generyczne odpowiedzi i brak pełnej emulacji stosu systemowego, jest stosunkowo łatwy do zidentyfikowania przez zaawansowane skanery.

h0neytr4p

h0neytr4p to nowoczesna, napisana w języku Go wersja klasycznego Honeytrap, zoptymalizowana pod kątem wydajności w środowiskach kontenerowych [4].

- Pełni rolę sieci zabezpieczającej; przejmuje każdy pakiet, który nie trafił do dedykowanego sensora, zapewniając 100% pokrycia prób połączeń.
- Pozwala na łatwe przekierowywanie ruchu do innych analizatorów lub logowanie surowych danych o połączeniach na nieużywanych portach [4].

Honeyaml

Honeyaml to wyspecjalizowane narzędzie do ochrony zasobów webowych przed automatycznym wyciąganiem danych [1].

- Pozwala identyfikować boty próbujące masowo pobierać dane z wystawionych publicznie API poprzez dostarczanie im sfałszowanych informacji.
- Umożliwia odróżnienie indeksujących robotów wyszukiwarek od agresywnych skanerów próbujących przejąć własność intelektualną serwisu [4].

IPPHoney

IPPHoney to honeypot emulujący protokół IPP (Internet Printing Protocol). W cyberbezpieczeństwie służy on do monitorowania zagrożeń wymierzonych w urządzenia biurowe i IoT, które często są pomijane w standardowych procedurach patchowania.

- Adresuje specyficzny segment bezpieczeństwa Internetu Rzeczy; drukarki sieciowe z przestarzałym firmwarem są atrakcyjnym celem dla atakujących szukających trwałości (persistence) w sieci.
- Pozwala na identyfikację skanerów poszukujących urządzeń peryferyjnych w celu ich dalszej eksploatacji lub kradzieży dokumentów z bufora wydruku [5].

Mailoney

Mailoney jest honeypotem niskiej interakcji symulującym serwer pocztowy SMTP. Jego głównym celem jest zbieranie danych o kampaniach spamowych oraz próbach wykorzystania serwerów do rozsyłania złośliwego oprogramowania [4].

- Analiza ruchu pozwala powiązać adresy IP z domenami znanymi z działalności botnetowej, co ułatwia atrybucję ataku do konkretnych grup przestępczych.
- Ze względu na uproszczoną implementację protokołu SMTP, zaawansowane narzędzia mogą łatwo odróżnić Mailoney od pełnoprawnych serwerów takich jak Postfix czy Exchange.

Medpot

Medpot emuluje protokół HL7 (Health Level Seven), który jest standardem wymiany danych w systemach medycznych. Jest to kluczowe narzędzie w ochronie infrastruktury sektora ochrony zdrowia [5].

- Szpitale są głównym celem ataków; aktywność na porcie 2575 w sieci zewnętrznej jest silnym indykatorem wrogiego rekonesansu poprzedzającego atak szyfrujący.
- Pozwala na monitorowanie prób dostępu do systemów informacji medycznej (HIS) bez ryzyka narażenia realnych danych pacjentów [2].

RedisHoneyPot

RedisHoneyPot symuluje bazę danych NoSQL Redis, która w przypadku błędnej konfiguracji często pozostaje otwarta na dostęp z zewnątrz bez uwierzytelnienia [2].

- Zwraca statyczne odpowiedzi i symuluje dane za pomocą prostej hashmapy, co pozwala na logowanie prób kradzieży danych lub infekcji serwera.
- Brak dynamicznej reakcji na obciążenie oraz ograniczony zestaw komend sprawiają, że jest łatwy do odróżnienia od produkcyjnej instancji przez zaawansowane skanery.

SentryPeer

SentryPeer to honeypot dedykowany dla protokołu SIP (Session Initiation Protocol), służący do wykrywania ataków na systemy telefonii IP (VoIP) oraz oszustw telekomunikacyjnych [2].

- Sensor rejestruje specyficzne metody, takie jak INVITE (próba połączenia), REGISTER (próba przejęcia konta) oraz OPTIONS (rekonesans).
- Pomaga w identyfikacji serwerów typu fraud oraz masowych prób łamania haseł do kont telefonii internetowej [2].

Wordpot

Wordpot to narzędzie emulujące instalację systemu CMS WordPress. Pozwala ono na zbieranie informacji o najczęstszych atakach wymierzonych w najpopularniejszą platformę blogową na świecie [5].

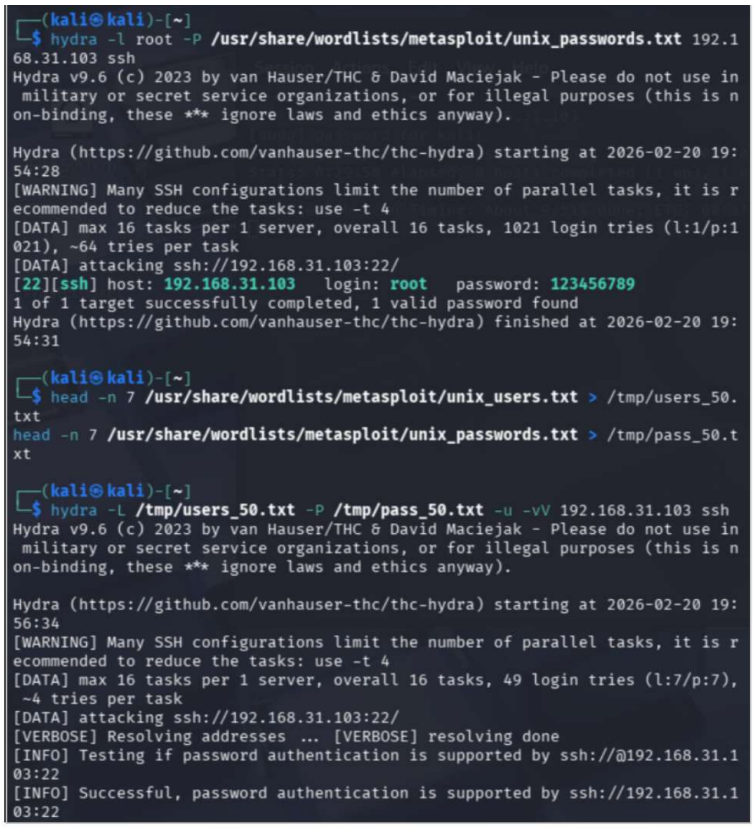
- Skutecznie identyfikuje próby enumeracji użytkowników, skanowanie podatnych wtyczek (Plugin Exploitation) oraz ataki typu brute-force na panel administracyjny.
- Ze względu na statyczną naturę plików, może być rozpoznany jako fałszywa instancja poprzez analizę nagłówków HTTP lub brak dynamicznego generowania treści PHP.

3. Wyniki przeprowadzonych testów

Weryfikację skuteczności platformy T-Pot przeprowadzono z wykorzystaniem systemu Kali Linux, wyposażonego w zestaw specjalistycznych narzędzi do audytu bezpieczeństwa. Wszystkie testy zostały zrealizowane w odizolowanym środowisku sieci lokalnej (LAN). Zapewniło to pełną kontrolę nad generowanym ruchem i wyeliminowało ryzyko przypadkowego oddziaływania na infrastrukturę zewnętrzną. Przyjęty model badawczy pozwolił na bezpieczne symulowanie realnych wektorów ataków. Dzięki takiemu podejściu możliwe było określenie szybkości reakcji poszczególnych honeypotów oraz poprawności logowania zdarzeń przez stos analityczny ELK.

3.1. Wykorzystane narzędzia

Pierwszym etapem testów był rekonesans platformy przy użyciu skanera Nmap. Zastosowano polecenie sudo nmap -A -p- <IP_T_POTA>, mające na celu identyfikację wszystkich otwartych portów, detekcję wersji usług oraz próbę rozpoznania systemu operacyjnego. Proces ten trwał 30 minut, po czym został przerwany ze względu na bardzo wysoki przewidywany czas ukończenia, wynoszący 7 godzin. Decyzja o przerwaniu skanowania wynikała z charakterystyki honeypotów typu tarpit, takich jak HellPot, które celowo spowalniają sesje TCP w celu wyczerpania zasobów skanera. Mimo niedokończonego procesu, zebrane w tym czasie logi pozwoliły na wstępną ocenę zdolności platformy do rejestrowania masowych zapytań o porty.



Rys. 3. Zdjęcie terminala Kali Linux przedstawiające kroki przygotowywania plików dla programu Hydra

Fig. 3. A picture of the Kali Linux terminal showing the steps involved in preparing files for the Hydra program

Wstępna próba użycia narzędzia Hydra wykazała, że domyślna konfiguracja programu przerywa działanie natychmiast po znalezieniu pierwszej poprawnej pary danych uwierzytelniających. Ponieważ celem badania nie było samo uzyskanie dostępu, lecz weryfikacja zdolności T-Pota do logowania wielu prób jednocześnie, konieczna była zmiana podejścia. Postanowiono zmodyfikować parametry ataku tak, aby proces nie ulegał przerwaniu przed sprawdzeniem wszystkich przygotowanych rekordów.

W celu zapewnienia pełnej kontroli nad przebiegiem testu, przygotowano dwa oddzielne pliki tekstowe zawierające po 50 rekordów: jeden z loginami, drugi z hasłami. Do przeprowadzenia ataku wykorzystano odpowiednią flagę wymuszającą kontynuację pracy. Przykładem zastosowanej składni było: hydra -L login_list.txt -P pass_list.txt <IP_T_POTA> ssh. Takie rozwiązanie pozwoliło na wygenerowanie kontrolowanego, seryjnego ruchu skierowanego na port 22, symulowanego przez honeypot Cowrie. Rysunek 3. przedstawia terminal w Kali Linux z komendami.

Ostatnim etapem tej fazy testów była analiza zachowania systemu pod wpływem powtarzalnych zapytań. Dzięki wymuszeniu pełnej sekwencji 50 prób, możliwe było sprawdzenie, czy interfejs analityczny Kibana poprawnie grupuje ataki pochodzące z jednego źródła oraz czy wszystkie pary danych są rejestrowane w bazie Elasticsearch.



Rys. 4. Zdjęcie z terminala Kali Linux pokazujące wprowadzone loginy i hasła

Fig. 4. A photo from the Kali Linux terminal showing what the input logins and passwords were

Podczas realizacji ataków odnotowano specyficzne zachowanie narzędzia Hydra, które wynikało z natury zaimplementowanego honeypota. Program wyświetlał powiadomienia wskazujące, że każdy wykorzystany login oraz każde hasło z przygotowanej listy są poprawne i pozwalają na autoryzację. Jest to efekt celowego działania honeypota Cowrie, który został skonfigurowany w trybie akceptowania wszystkich prób logowania. Taka strategia ma na celu zachęcenie napastnika do kontynuowania interakcji i przejścia do fazy wprowadzania komend systemowych. Z punktu widzenia badacza pozwoliło to na zweryfikowanie, czy

każda z 50 prób została prawidłowo zarejestrowana jako udany dostęp w logach platformy T-Pot. Rysunek 4. przedstawia terminal Kali Linux z loginami i hasłami.

Kolejny etap testów koncentrował się na analizie bezpieczeństwa warstwy aplikacyjnej przy użyciu skanera Nikto. Wykorzystano polecenie nikto -h http://192.168.31.x, aby zidentyfikować potencjalne luki w konfiguracji serwera HTTP emulowanego przez platformę T-Pot. Narzędzie to, w odróżnieniu od ogólnego skanowania portów, skupia się na wykrywaniu niebezpiecznych plików, przestarzałego oprogramowania oraz błędów w nagłówkach odpowiedzi serwera. Wykonany test dostarczył obszernego raportu, który potwierdził zdolność honeypotów webowych (takich jak np. Tanner czy Galah) do symulowania systemów zawierających liczne, krytyczne podatności, co ma na celu zatrzymanie uwagi napastnika na dłuższy czas.

Na załączonym zrzucie ekranu widoczna jest lista krytycznych podatności typu Cross-Site Scripting (XSS) oraz liczne błędy w konfiguracji bezpieczeństwa. System zidentyfikował brak nagłówków ochronnych, takich jak X-Frame-Options (ochrona przed Clickjackingiem) oraz X-Content-Type-Options, co w realnym środowisku ułatwiłoby przeprowadzenie ataku na użytkowników końcowych. Raport wskazuje również na obecność podatnych komponentów, takich jak Mambo PHP Portal, które są znane z luk umożliwiających wstrzykiwanie skryptów (CVE-2003-1204). Wykrycie tak dużej liczby "pozytywnych" wyników przez Nikto świadczy o wysokiej jakości emulacji platformy T-Pot; honeypot celowo podaje fałszywe informacje o zainstalowanych usługach i ich wersjach, aby sprowokować atakującego do podjęcia próby eksploatacji nieistniejących w rzeczywistości błędów. Rysunek 5. przedstawia raport po użyciu nikto.

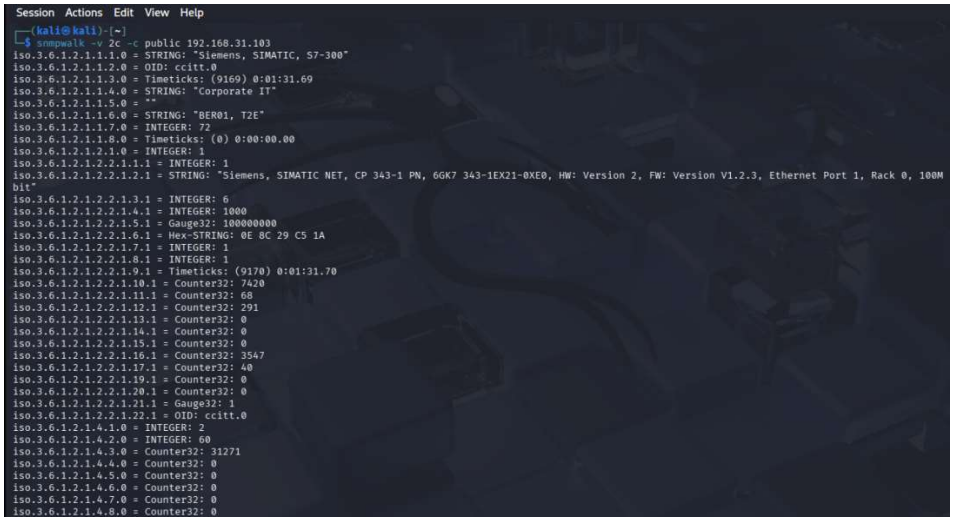


Rys. 5. Raport wygenerowany podczas korzystania z oprogramowania nikto

Fig. 5. Report generated while using nikto software

Kolejny etap testów objął weryfikację usług zarządzania sieciowego poprzez wykonanie polecenia snmpwalk -v 2c -c public 192.168.31.103. Test ten miał na celu sprawdzenie, jak platforma T-Pot reaguje na próby odpytywania o obiekty bazy MIB (Management Information Base) przy użyciu domyślnego ciągu uwierzytelniającego "public". Operacja ta jest standardowym elementem rekonesansu, pozwalającym napastnikowi na uzyskanie szczegółowych informacji o sprzęcie, konfiguracji sieciowej oraz zainstalowanym oprogramowaniu bez konieczności łamania zaawansowanych zabezpieczeń.

Analiza wyników przedstawionych na zrzucie ekranu wskazuje na bardzo wysoką jakość emulacji komponentów przemysłowych przez honeypot Conpot. Narzędzie zwróciło identyfikatory OID wskazujące na obecność sterownika Siemens SIMATIC S7-300 oraz modułu komunikacyjnego CP 343-1 PN. Widoczne dane, takie jak wersja oprogramowania układowego (FW: Version V1.2.3) oraz lokalizacja zasobu (Corporate IT, BER01, T2E), są spreparowanymi informacjami, które mają utwierdzić napastnika w przekonaniu, że trafił na rzeczywistą infrastrukturę sterowania procesami technologicznymi. Sukces tego testu potwierdza, że T-Pot skutecznie maskuje swoją naturę, wystawiając na widok publiczny parametry typowe dla systemów automatyki przemysłowej, co pozwala na monitorowanie specyficznych zagrożeń typu ICS. Rysunek 6. przedstawia raport po użyciu snmpwalk.



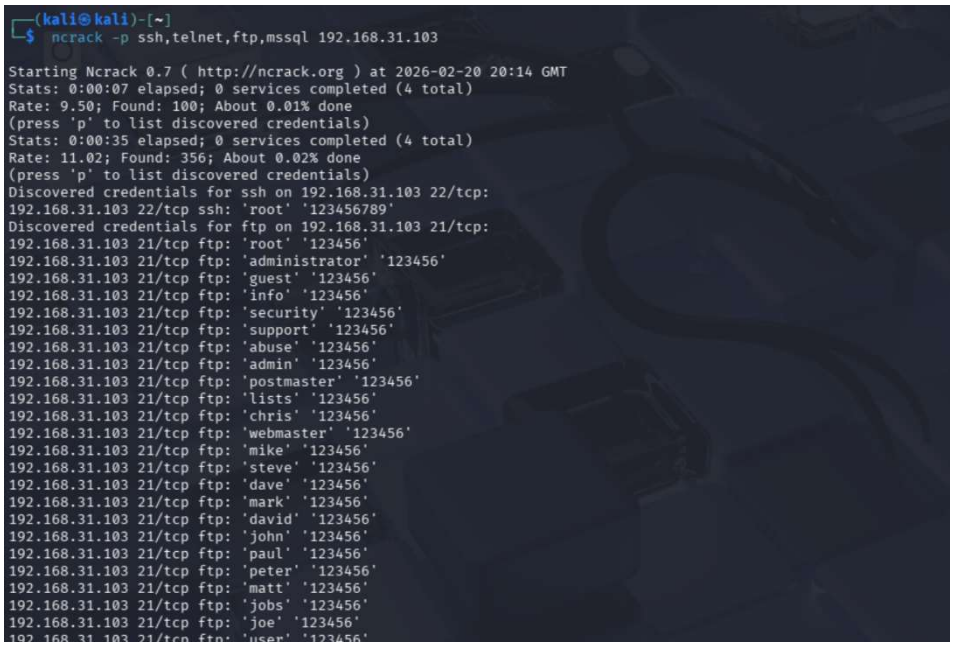
Rys. 6. Raport wygenerowany podczas korzystania z oprogramowania snmpwalk

Fig. 6. Report generated while using snmpwalk software

Następnym etapem testów penetracyjnych było użycie narzędzia Ncrack, zorientowanego na szybkie i wieloprotokołowe łamanie haseł w sieciach o dużej skali. Wykonano polecenie ncrack -p ssh,telnet,ftp,mssql 192.168.31.103, aby zweryfikować responsywność platformy T-Pot na jednoczesne próby logowania

do różnych usług systemowych. W przeciwieństwie do Hydry, Ncrack optymalizuje proces skanowania pod kątem wielu celów i portów jednocześnie, co pozwoliło na szybkie zidentyfikowanie domyślnych poświadczeń akceptowanych przez emulowane serwery. Test ten potwierdził, że system zbiera logi z wielu sensorów (takich jak Cowrie dla SSH czy Dionaea dla FTP i MSSQL) w sposób równoległy i spójny.

Wyniki widoczne na zrzucie ekranu potwierdzają specyficzną strategię obronną zaimplementowaną w T-Pot. Narzędzie Ncrack w bardzo krótkim czasie (zaledwie 35 sekund) odnalazło setki "poprawnych" danych uwierzytelniających dla usług SSH oraz FTP. Przykładowo, dla protokołu FTP na porcie 21 system zaakceptował szeroki wachlarz loginów, takich jak administrator, guest, support czy webmaster, z identycznym hasłem 123456. Tak masowe i natychmiastowe potwierdzanie poprawności haseł przez honeypoty ma na celu zachęcenie automatów i botnetów do nawiązania pełnej sesji, co pozwala platformie na przechwycenie dalszych poleceń i pobieranych skryptów złośliwego oprogramowania. Dla badacza jest to dowód na to, że T-Pot skutecznie "zatrzymuje" agresora, tworząc iluzję łatwego przejęcia systemowych kont dostępowych. Rysunek 7. przedstawia raport wygenerowany po użyciu ncrack.



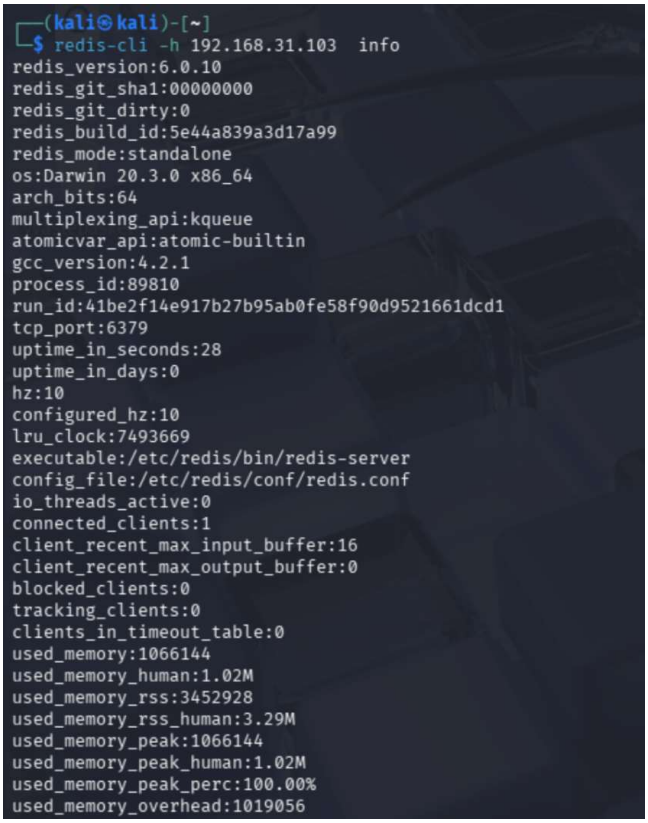
Rys. 7. Raport wygenerowany podczas korzystania z oprogramowania ncrack

Fig. 7. Report generated while using ncrack software

Kolejnym elementem weryfikacji usług bazodanowych była próba uzyskania nieautoryzowanego dostępu do instancji Redis przy użyciu narzędzia redis-cli. Test

polegał na próbie nawiązania połączenia bez podania hasła oraz wykonaniu komendy info, która w rzeczywistych warunkach zwraca szczegółowe dane o stanie serwera, wersji oprogramowania i wykorzystaniu zasobów. Z punktu widzenia napastnika, brak wymuszenia autoryzacji w usłudze Redis jest krytycznym błędem konfiguracyjnym, który pozwala na pełną eksploatację bazy, w tym kradzież danych lub ich usunięcie. Wykonanie zapytania redis-cli -h <IP_T_POTA> info miało na celu sprawdzenie, jak honeypot RedisHoneyPot emuluje tę konkretną podatność.

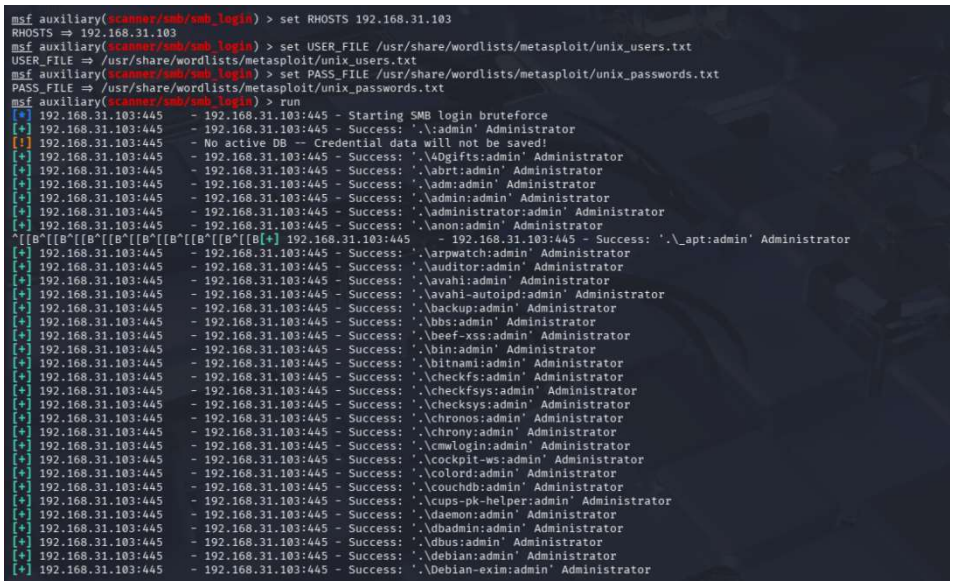
Przeprowadzony test wykazał, że honeypot poprawnie symuluje otwartą i podatną instancję bazy danych. Zamiast odrzucenia połączenia, system zaakceptował komendę i zwrócił spreparowane informacje systemowe, co pozwoliło na zarejestrowanie adresu IP agresora oraz typu wykonanego zapytania. W badaniach potwierdzono, że RedisHoneyPot wykorzystuje proste mapowanie odpowiedzi, które mimo swojej statycznej natury, jest wystarczające do zmylenia zautomatyzowanych skanerów poszukujących niezabezpieczonych baz danych. Rysunek 8. przedstawia raport z użycia redis.



Rys. 8. Raport wygenerowany podczas korzystania z oprogramowania redis

Fig. 8. Report generated while using redis software

Ostatnim etapem badań była weryfikacja interakcji T-Pota z zaawansowanymi narzędziami do automatyzacji ataków, takimi jak Metasploit Framework (MSF). Wykorzystano moduły skanujące oraz gotowe exploity, aby sprawdzić, jak platforma reaguje na próby przejęcia kontroli nad systemem. W przeciwieństwie do prostych skanerów portów, Metasploit wysyła specyficzne pakiety mające na celu wywołanie błędu w usłudze, co pozwala na uruchomienie własnego kodu (payloadu). Testy koncentrowały się na popularnych podatnościach w protokołach SMB oraz HTTP, emulowanych przez sensory Dionaea oraz Tanner, aby ocenić głębię symulowanej interakcji. Rysunek 9. przedstawia raport po użyciu oprogramowania Metasploit Framework.



Rys. 9. Raport wygenerowany podczas korzystania z oprogramowania Metasploit Framework

Fig. 9. Report generated while using Metasploit Framework

Analiza zachowania platformy wykazała, że T-Pot skutecznie identyfikuje aktywność frameworka Metasploit dzięki zintegrowanemu systemowi wykrywania intruzów (IDS) Suricata. Każda próba użycia znanego exploita generowała odpowiedni alert w panelu Kibana, precyzyjnie wskazując na rodzaj zastosowanego ataku (np. EternalBlue czy próby Directory Traversal). Mimo że honeypoty o niskiej interakcji nie pozwalają na pełne wykonanie payloadu, ich zdolność do symulowania podatnych odpowiedzi serwera jest wystarczająca, aby zmylić narzędzie i zarejestrować unikalne sygnatury ataku. Pozwala to na gromadzenie cennych informacji o trendach w wykorzystywaniu konkretnych luk bezpieczeństwa przez współczesnych agresorów.

3.2. Analiza zebranych danych w T-Pot

Kluczowym etapem badania była analiza danych zgromadzonych przez stos analityczny ELK (Elasticsearch, Logstash, Kibana), który stanowi integralną część platformy T-Pot. Wszystkie próby interakcji podjęte przy użyciu narzędzi z systemu Kali Linux zostały zidentyfikowane, skategoryzowane i poddane procesowi wizualizacji w czasie rzeczywistym. Poniższa sekcja prezentuje interpretację wyników przedstawionych na zrzutach ekranu z panelu administracyjnego, co pozwala na ocenę skuteczności sensorów w detekcji poszczególnych wektorów ataków. Analiza ta skupia się na sposobie prezentacji surowych logów oraz ich transformacji w czytelne wykresy statystyczne i mapy aktywności.

Rysunek 10. prezentuje tabelę oraz powiązane z nią liczniki stanowią syntetyczne podsumowanie całej sesji testowej, agregując dane z wielu honeypotów w jednym widoku. Kluczowym elementem tego zestawienia jest całkowita liczba zarejestrowanych incydentów (Total Events). Dane te pozwalają na błyskawiczną ocenę skali ataku oraz identyfikację najbardziej aktywnych sensorów w danym przedziale czasowym. Tabela porządkuje ataki według ilości, co pozwala badaczowi stwierdzić, że najbardziej obciążonymi elementami systemu były usługi SSH oraz HTTP.

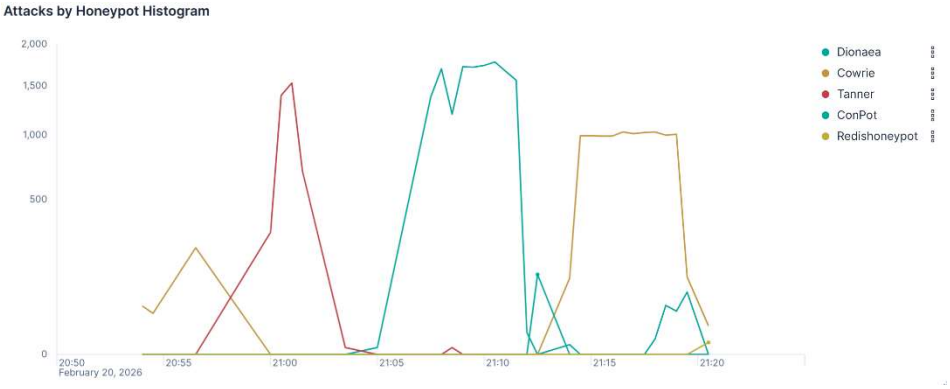


Rys. 10. Statystyki zbiorcze zdarzeń oraz wskaźniki aktywności sensorów

Fig. 10. Aggregate event statistics and sensor activity indicators

Wykres na rysunku 11. przedstawia intensywność ruchu sieciowego zarejestrowanego przez poszczególne sensory platformy T-Pot w przedziale czasowym od 20:50 do 21:20. Widoczna jest sekwencyjna aktywacja różnych honeypotów, co odpowiada kolejnym fazom testów penetracyjnych przeprowadzanych z systemu Kali Linux. Najwyższą aktywność odnotowano dla sensora Dionaea (ataki na protokoły sieciowe) oraz Tanner (emulacja luk w aplikacjach webowych), gdzie liczba zdarzeń przekroczyła poziom 1500 w szczytowych momentach.

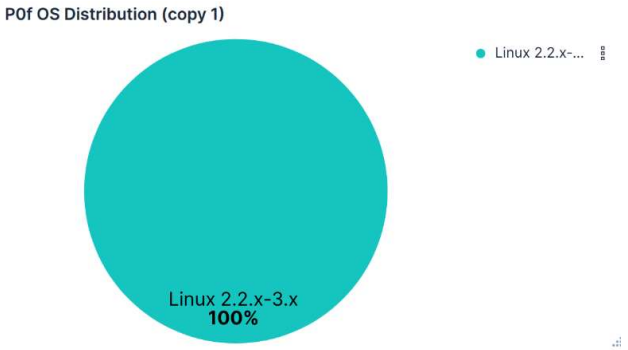
Stabilny wzrost oraz utrzymujący się wysoki poziom logów dla honeypota Cowrie potwierdza przeprowadzenie skutecznej symulacji ataku brute-force. Zróżnicowana dynamika wykresów potwierdza selektywne działanie wykorzystanych narzędzi skanujących i exploitów na konkretne usługi systemowe. Pozostałe sensory, takie jak ConPot czy Redishoneypot, wykazują śladową aktywność, co wskazuje na specyficzny dobór wektorów ataku ograniczonych do najpopularniejszych usług sieciowych i webowych.



Rys. 11. Histogram natężenia ataków zarejestrowanych przez poszczególne sensory systemu T-Pot w czasie rzeczywistym

Fig. 11. Histogram of attack intensity recorded by individual T-Pot system sensors in real-time

Wykres kołowy przedstawia wyniki pasywnej identyfikacji systemów operacyjnych (fingerprinting) wykonanej przez narzędzie p0f wewnątrz platformy T-Pot. Dane wskazują na pełną jednorodność źródła ruchu sieciowego, gdzie 100% zidentyfikowanych sygnatur przypisano rodzinie systemów Linux (jądro 2.2.x-3.x). Wynik ten jest bezpośrednim następstwem konfiguracji środowiska testowego, w którym wszystkie działania ofensywne były generowane wyłącznie z jednej stacji roboczej opartej na systemie Kali Linux. Brak dywersyfikacji w tym obszarze potwierdza, że mechanizmy detekcji T-Pot poprawnie sklasyfikowały stos TCP/IP maszyny atakującej, co jest kluczowe dla późniejszej analizy profilu intruza.



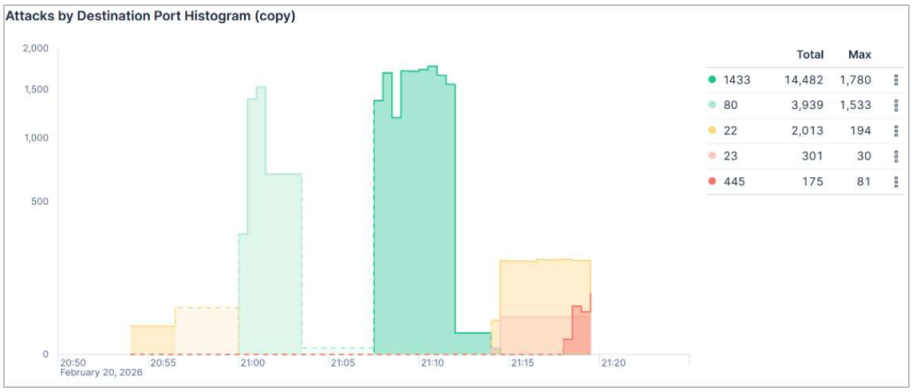
Rys. 12. Rozkład systemów operacyjnych zidentyfikowanych przez sensor p0f na podstawie analizy pakietów

Fig. 12. Distribution of operating systems identified by the p0f sensor based on packet analysis

Wykres na rysunku 12. przedstawia rozkład natężenia ataków w odniesieniu do konkretnych portów docelowych usług emulowanych przez T-Pot. Najwyższą

aktywność, przekraczającą 14 000 zdarzeń (Total), odnotowano na porcie 1433 (MSSQL), co skorelowane jest z intensywną pracą sensora Dionaea w fazie testów eksploitacji baz danych. Znaczący wolumen ruchu zarejestrowano również na porcie 80 (HTTP), gdzie liczba incydentów w szczytowym punkcie wyniosła 1533, co potwierdza symulację ataków na warstwę aplikacji webowej.

Rysunek 13. przedstawia histogram natężenia ruchu sieciowego. Ruch na portach 22 (SSH) oraz 23 (Telnet) wykazuje mniejszą dynamikę, lecz stałą obecność w końcowej fazie badanego interwału czasowego. Wartości te, wynoszące odpowiednio 2013 oraz 301 zdarzeń, wskazują na przeprowadzenie prób uwierzytelnienia oraz skanowania usług zdalnego dostępu. Minimalna aktywność na porcie 445 (SMB) sugeruje ograniczony zakres testów dotyczących protokołów wymiany plików, co pozwoliło na precyzyjne odizolowanie ruchu generowanego przez konkretne narzędzia z systemu Kali Linux.



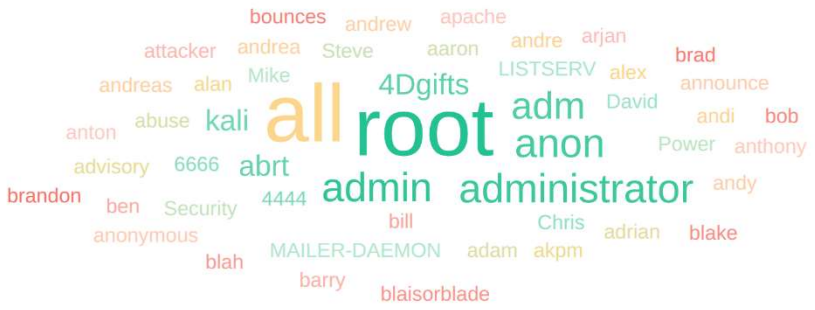
Rys. 13. Histogram natężenia ruchu sieciowego w podziale na porty docelowe z uwzględnieniem statystyk sumarycznych i wartości maksymalnych

Fig. 13. Histogram of network traffic intensity by destination ports, including total statistics and peak values

Prezentowane chmury tagów (ang. word clouds) na rysunku 14. obrazują statystyczną częstotliwość danych przechwyconych przez sensory typu honeypot w ramach platformy T-Pot. Pierwsza wizualizacja dokumentuje najczęściej wykorzystywane nazwy użytkowników podczas prób nieautoryzowanego dostępu. Dominacja terminów takich jak root, admin oraz administrator wskazuje na ukierunkowanie ataków na konta o najwyższych uprawnieniach systemowych. Obecność nazwisk i imion sugeruje natomiast wykorzystanie słowników predefiniowanych do ataków typu brute force.

Druga wizualizacja na rysunku 15. przedstawia zestawienie najpopularniejszych fraz wprowadzanych w polach haseł. Analiza wykazuje wysoką powtarzalność ciągów numerycznych (np. 123456) oraz prostych haseł słownikowych (np. password, iloveyou). Dane te potwierdzają, że automatyczne skrypty atakujące

(boty) w pierwszej kolejności eksploatują najsłabsze ogniwa autoryzacji, opierając się na globalnych listach najczęściej używanych i domyślnych poświadczeń.



Rys. 14. Ranking najczęściej wykorzystywanych nazw użytkowników i haseł podczas prób nieautoryzowanego dostępu

Fig. 14. Ranking of the most frequently used usernames and passwords during unauthorized access attempts



Rys. 15. Zestawienie najczęściej wykorzystywanych haseł przechwyconych przez sensory honeypot

Fig. 15. Distribution of the most frequently utilized passwords captured by honeypot sensors

Podsumowując, przeprowadzone testy weryfikują wysoką skuteczność platformy T-Pot w zakresie detekcji i profilowania zagrożeń sieciowych. Zastosowany stos analityczny ELK precyzyjnie zarejestrował sekwencje działań ofensywnych

generowanych ze środowiska Kali Linux. Analiza zgromadzonych logów wskazała na największe natężenie ruchu w obrębie usług operujących na portach HTTP, SSH oraz MSSQL. Wskazuje to na wyraźne ukierunkowanie zautomatyzowanych ataków na powszechnie wykorzystywane protokoły komunikacyjne i bazy danych.

Ponadto, przechwycone dane uwierzytelniające bezspornie dowodzą polegania na słownikowych metodach brute-force w początkowych fazach kompromitacji systemów. Pełna identyfikacja źródła ruchu sieciowego oraz precyzyjne mapowanie aktywności potwierdzają zdolności analityczne systemu. Integracja różnorodnych sensorów honeypot dostarcza w konsekwencji wiarygodnego materiału badawczego do monitorowania wektorów ataków w czasie rzeczywistym.

4. Podsumowanie

System T-Pot jest oparty na technologii konteneryzacji Docker, integruje szereg zróżnicowanych usług typu honeypot oraz stos analityczny ELK. Badania skuteczności przeprowadzono w izolowanym środowisku sieci lokalnej. Wykorzystano system Kali Linux do symulacji realnych wektorów ataków. Zastosowana metodologia objęła użycie narzędzi audytowych, takich jak Nmap, Hydra, Nikto, Ncrack oraz Metasploit Framework. Wykazano, że platforma skutecznie wchodzi w interakcję ze skanerami. Mechanizmy typu tarpit spowalniają proces skanowania portów w celu wyczerpania zasobów napastnika. Inne sensory, jak Cowrie, celowo akceptują wprowadzane poświadczenia logowania. Strategia ta zachęca automatyczne skrypty do nawiązywania sesji, co umożliwia szczegółową rejestrację komend systemowych. Platforma poprawnie symuluje również krytyczne luki w aplikacjach webowych, bazach danych oraz systemach automatyki przemysłowej.

Kluczowym elementem oceny systemu była analiza logów zgromadzonych w środowisku Kibana. Zarejestrowane statystyki wykazały najwyższe natężenie wrogiego ruchu sieciowego w obrębie emulowanych usług bazodanowych, na porcie 1433. Znaczące obciążenie odnotowano również dla aplikacji webowych na porcie 80 oraz dla usług zdalnego dostępu na porcie 22. Analiza przechwyconych danych uwierzytelniających potwierdziła, że początkowe fazy kompromitacji systemów opierają się na słownikowych atakach typu brute-force. Zautomatyzowane skrypty masowo wykorzystują domyślne nazwy kont administratorów oraz proste sekwencje numeryczne w polach haseł. Zgromadzone wyniki dowodzą ostatecznie, że integracja wielu sensorów w systemie T-Pot pozwala na precyzyjne mapowanie aktywności intruzów oraz profilowanie współczesnych zagrożeń w czasie rzeczywistym.

Literatura

[1] Devi Priya V. S., Sibi Chakkaravarthy S.: Containerized cloud-based honey-pot deception for tracking attackers. Scientific Reports, 13(1):1437, 2023.

[2] Dokumentacja Beelzebub: https://github.com/mariocandela/beelzebub (dostęp 12.12. 2025).
[3] Fan W., Du Z., Fernández D., Villagrá V. A.: Enabling an anatomic view to inves-tigate honeypot systems: A survey. IEEE Systems Journal, 12(4):3906–3919, 2017.
[4] Hilt S., Maggi F., Perine C., Remorin L., Rösler M., Vosseler R.: Caught in the act: Running a realistic factory honeypot to capture real threats. Trend Micro Research, 2020.
[5] Machmeier S.: Honeypot implementation in a cloud environment. arXiv pre-print arXiv:2301.00710, 2023.
[6] Mesbah M., Elsayed M. S., Jurcut A. D., Azer M.: Analysis of ICS and SCADA systems attacks using honeypots. Future Internet, 15(7), 241. 2023.
[7] Nawrocki M., Wählisch M., Schmidt T. C., Keil C., Schönfelder J.: A survey on honeypot software and data analysis. arXiv pre-print arXiv:1608.06249, 2016.
[8] Sivamohan S., Sridhar S. S., Krishnaveni S.: Efficient multi-platform honey- pot for capturing real-time cyber attacks. In Intelligent Data Communication Technol-ogies and Internet of Things: Proceedings of ICICI 2021, pages 291–308. Springer, 2022.
[9] Washofsky A. D.: Deploying and analyzing containerized honeypots in the cloud with t-pot. PhD thesis, Monterey, CA; Naval Postgraduate School, 2021.
[10] Zymberi I.: Honeypots: A means of sensitizing awareness of cybersecu- rity con-cerns. 2021.

ANALYSIS OF THE T-POT PLATFORM'S EFFECTIVENESS IN DETECTING AND PROFILING NETWORK THREATS

Summary

This article presents the deployment and effectiveness verification of the T-Pot platform, which serves as a honeypot environment. This system utilizes Docker containerization technology to run services that simulate vulnerabilities in operating systems, databases, and network protocols. The research objective was to evaluate the platform in terms of collecting data on attack vectors in real-time. The research was conducted in an isolated local area network environment, which ensured control over the traffic and eliminated the risk to the external infrastructure. Offensive actions were generated from a Kali Linux system, utilizing auditing tools: Nmap, Nikto, Hydra, Ncrack, and Metasploit Framework. Simulations showed that the platform's sensors intentionally modify responses to queries. Tarpit mechanisms slow down scanning, exhausting the attacker's resources. Service emulators massively accept login credentials, encouraging automated scripts to establish sessions and execute commands. A key stage was the analysis of data collected by the ELK analytics stack. Statistics indicated the highest intensity of attacks within database services (port 1433), web applications (port 80), and remote access services (port 22). The analysis of captured authentication data confirmed that system compromise relies on dictionary attacks, exploiting the most commonly used passwords and high-privilege accounts. The obtained results prove that the integration of multiple sensors provides reliable and solid research material. The T-Pot environment constitutes an effective solution for detection, mapping intruder activity, and profiling network threats.

Keywords: honeypot, T-Pot platform, attack vectors, ELK stack, threat detection.

Norbert Mazur
Koło Naukowe Interakcji Człowiek - Komputer GEST

dr hab. inż. Tomasz Kapuściński, prof. PRz, dr inż. Dawid Kalandyk
opiekunowie Koła Naukowego

IMPLEMENTACJA MODUŁU GRY TYPU TOWER DEFENSE W ŚRODOWISKU PLATFORMY EDUKACYJNEJ RUT-AI Z WYKORZYSTANIEM JĘZYKA TYPESCRIPT

Streszczenie

W rozdziale zaprezentowano proces projektowania, implementacji oraz parametryzacji dwuwymiarowej gry z gatunku Tower Defense. Aplikacja została utworzona jako nowy element edukacyjnej platformy webowej RUT-AI Games, której głównym celem jest dostarczenie środowiska do testowania i ewaluacji algorytmów sztucznej inteligencji. Projekt zrealizowano z wykorzystaniem silnie typowanego języka TypeScript, co pozwoliło na przygotowanie przejrzystej, modularnej architektury oraz ułatwiło integrację z frameworkiem Angular, na którym opiera się cała platforma. Za warstwę wizualną, w tym renderowanie obiektów i animacji w przeglądarce odpowiada interfejs programistyczny HTML5 Canvas w kontekście dwuwymiarowym. W pracy omówiono kluczowe aspekty inżynierii oprogramowania gier wideo. Głównym elementem mechaniki poruszania się jednostek wroga było zastosowanie algorytmu przeszukiwania wszerz (BFS – Breadth-First Search), który zaadaptowano do dynamicznego wyznaczania najkrótszych ścieżek po dyskretnej siatce przestrzennej. Dodatkowo przeprowadzono analizę modelu ekonomicznego gry, obejmującą parametryzację kosztów budowy wież obronnych, ocenę ich efektywności bojowej (wskaźnik obrażeń na sekundę) oraz mechanizmy skalowania poziomu trudności w kolejnych falach przeciwników. Zrealizowany projekt stanowi w pełni funkcjonalną i zbalansowaną aplikację. Dzięki przemyślanej architekturze i zastosowaniu powszechnie uznanych wzorców projektowych, gra stanowi doskonały fundament do dalszych prac badawczo-rozwojowych, w tym przede wszystkim do wdrażania zaawansowanych systemów uczenia maszynowego w ramach ekosystemu RUT-AI Games.

Słowa kluczowe: Tower Defense, TypeScript, algorytm BFS, ekonomia rozgrywki, architektura oprogramowania webowego.

1. Wprowadzenie

Gry od wielu lat stanowią nie tylko gałąź przemysłu rozrywkowego, ale także doskonałe środowisko do badań nad inżynierią oprogramowania oraz algorytmami sztucznej inteligencji (AI). Szczególnie interesującym gatunkiem w kontekście testowania optymalizacji przestrzennej oraz procesów decyzyjnych jest Tower Defense. Rozgrywka w produkcjach tego typu polega na odpieraniu kolejnych fal przeciwników poprzez strategiczne rozmieszczenie wież obronnych na określonej mapie. Złożoność problemów pojawiających się w takich grach: dynamiczne wyznaczanie najkrótszej ścieżki, zarządzanie ograniczonymi zasobami czy optymalizacja renderowania wielu obiektów jednocześnie - czyni je idealnym polem do praktycznej implementacji złożonych struktur danych i algorytmów.

W odpowiedzi na potrzebę tworzenia nowoczesnych środowisk badawczo-edukacyjnych, rozwijana jest platforma webowa RUT-AI Games [7]. Jest to zintegrowany ekosystem, którego głównym celem jest dostarczenie infrastruktury do uruchamiania, testowania i wizualizacji działania różnorodnych mechanik oraz algorytmów sztucznej inteligencji.

2. Cel pracy

Głównym celem pracy było zaprojektowanie, implementacja oraz parametryzacja w pełni grywalnej aplikacji z gatunku Tower Defense. Została ona stworzona jako rozszerzenie funkcjonalności istniejącego ekosystemu RUT-AI Games, funkcjonując jako nowy, niezależny komponent bezpośrednio w strukturze repozytorium platformy. Zastosowanie takiego podejścia pozwoliło na zachowanie niezależności logiki gry przy jednoczesnym działaniu w tym samym środowisku uruchomieniowym.

3. Modelowanie przestrzeni i algorytm wyznaczania ścieżki

Kluczowym elementem mechaniki gier z gatunku Tower Defense jest system poruszania się jednostek wroga od punktu odrodzenia do bazy gracza. Wymaga to odpowiedniego zamodelowania przestrzeni rozgrywki oraz implementacji wydajnego algorytmu przeszukiwania grafu.

3.1. Dyskretna reprezentacja mapy

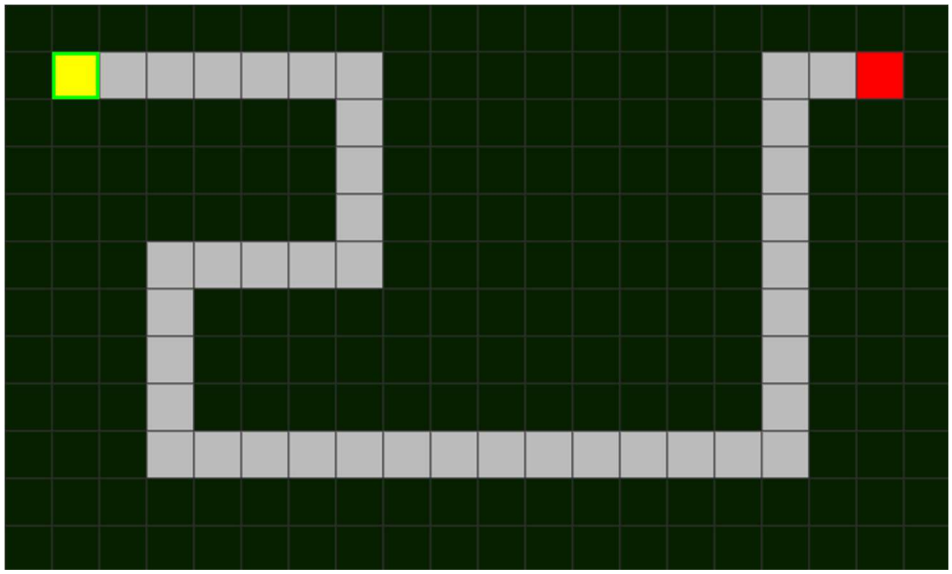
Świat gry został zaprojektowany z wykorzystaniem struktury dyskretnej siatki dwuwymiarowej. W pamięci aplikacji mapa przechowywana jest pod postacią dwuwymiarowej tablicy liczb całkowitych (*number[][]*), gdzie każda wartość numeryczna determinuje typ i właściwości danego pola.

Zastosowano następujące kodowanie przestrzeni (rys. 1.):

- 0 - pole puste, na którym gracz może wznosić swoje wieże obronne,

- 1 - wyznaczona ścieżka, po której poruszają się jednostki naziemne,
- 2 - punkt początkowy, w którym generowane są fale przeciwników,
- 3 - punkt końcowy, którego gracz musi bronić

Taka reprezentacja pamięciowa redukuje problem znajdowania ścieżki do klasycznego problemu przeszukiwania grafu, w którym każdy kafelek ścieżki jest węzłem, a dozwolone ruchy (góra, dół, lewo, prawo) stanowią krawędzie o jednakowej wadze przypisanej do przejścia.



Rys. 1. Widok podstawowej struktury mapy statycznej z punktami START i BAZA
Fig. 1. View of the basic static map structure with START and BASE points

3.2. Implementacja algorytmu BFS (Breadth-First Search)

Ponieważ w zrealizowanym środowisku koszt przejścia pomiędzy sąsiadującymi polami ścieżki jest stały, do dynamicznego znajdowania optymalnej trasy dla przeciwników wykorzystano algorytm przeszukiwania wszerz – BFS [2, 8]. W odróżnieniu od heurystycznych algorytmów takich jak A*, BFS jest idealnie dopasowany do niemających wag siatek prostopadłych, gwarantując odnalezienie najkrótszej możliwej ścieżki przy optymalnej złożoności obliczeniowej:

$$O(V+E) \tag{1}$$

gdzie: V - liczba węzłów,
E - liczba krawędzi.

Algorytm (Listing 1.) rozpoczyna działanie od zlokalizowania współrzędnych punktu startowego oraz końcowego. Główny proces przeszukiwania opiera się na strukturze kolejki typu FIFO (First-In, First-Out). Należy jednak wziąć pod uwagę, że złożoność całej operacji jest ściśle zależna od złożoności usuwania pierwszego elementu z kolejki. W przypadku natywnej metody w języku JavaScript (*queue.shift()*), operacja ta może osiągać złożoność O(n) ze względu na konieczność przesunięcia całej tablicy w pamięci. Zjawisko to przy bardzo rozbudowanych grafach może wpłynąć na ogólną wydajność algorytmu. W celu śledzenia postępu oraz późniejszego odtworzenia najkrótszej drogi, wykorzystano strukturę słownikową *Map<string, object>* dostępną w języku TypeScript [9]. Struktura ta (*cameFrom*) mapuje unikalny klucz znakowy, stanowiący współrzędne węzła (np. „y,x"), na obiekt reprezentujący węzeł, z którego do niego dotarło. Zastosowanie mapy zamiast standardowej tablicy zapewnia złożoność czasową O(1) dla operacji sprawdzania czy dany węzeł został już odwiedzony. Po osiągnięciu punktu docelowego algorytm przechodzi do fazy rekonstrukcji ścieżki. Wykorzystując zgromadzone dane w strukturze *cameFrom*, algorytm wykonuje iterację wstecz (tzw. backtracking) od punktu *endPos* do *startPos*. Otrzymana tablica współrzędnych jest następnie odwracana za pomocą metody *reverse()*, tworząc gotowy sekwencyjny bufor instrukcji ruchu dla obiektów klasy *Enemy*. Tak zaimplementowany moduł jest odporny na zmiany ukształtowania terenu i wylicza trasę w sposób dynamiczny, co stanowi solidny fundament dla przyszłych prac.

```
while (queue.length > 0) {
    const current = queue.shift();
    if (!current) {
        console.error("BFS Error: Queue shift returned undefined unexpectedly.");
        break;
    }
    // warunek końcowy
    if (current.x === endPos.x && current.y === endPos.y)
    {
         finalEndPos = current;
         break;
    }
    const neighbors = [ // wyznaczenie 4 kierunków
        { x: current.x, y: current.y - 1 }, { x: current.x, y: current.y + 1 },
        { x: current.x - 1, y: current.y }, { x: current.x + 1, y: current.y },
    ];
    for (const next of neighbors) {
        const nextKey = `${next.y},${next.x}`;
        if (next.x < 0 || next.x >= mapWidth || next.y < 0 || next.y >= mapHeight)
        {
            continue; // weryfikacja granic mapy
         }
         const tileValue = map[next.y][next.x];
         // weryfikacja czy pole jest ścieżką i nie zostało odwiedzone
```

```
        if ((tileValue === 1 || tileValue === 3) && !cameFrom.has(nextKey))
        {
            queue.push(next);
            cameFrom.set(nextKey, current);
        }
    }
}
```

Listing 1. Fragment implementacji głównej pętli algorytmu BFS w języku TypeScript
Listing 1. Fragment of the main BFS algorithm loop implementation in TypeScript

3.3. Specyfika ruchu jednostek latających

Ważnym rozszerzeniem zaimplementowanej logiki poruszania się jest obsługa jednostek powietrznych (np. Helikopter, Odrzutowiec). W przeciwieństwie do jednostek naziemnych, które są ściśle ograniczone przez wyznaczoną w tablicy ścieżkę, jednostki latające posiadają parametr *isFlying* ustawiony na *true*, więc nie muszą podążać wyznaczoną ścieżką. Z punktu widzenia algorytmicznego, dla tych jednostek mapa staje się grafem pełnym. W praktyce oznacza to, że jednostki te mogą poruszać się „na skróty" – bezpośrednio w linii prostej od punktu startu do bazy, ignorując ścieżkę wyznaczoną dla jednostek naziemnych.

4. Architektura systemu, zarządzanie stanem i warstwa wizualna

Aby umożliwić bezproblemową integrację gry z edukacyjną platformą RUT-AI Games, konieczne było zaprojektowanie modularnej architektury opartej na jednoznacznym oddzieleniu logiki i stanu gry od warstwy prezentacji wykorzystując do tego sprawdzone wzorce projektowe [3].

4.1. Zarządzanie stanem i interfejs dla algorytmów AI (I/O)

Sercem aplikacji jest klasa *TowerDefenseState*, pełniąca rolę kontenera na wszystkie dynamiczne parametry środowiska. Przechowuje ona m.in. stan zasobów (*gold*), stan zdrowia bazy (*baseHealth*), mapę (*map*) oraz tablice aktywnych obiektów (*towers*, *enemies*, *bullets*). Kluczowym aspektem integracji z frameworkiem uczelnianym było zdefiniowanie ścisłego interfejsu wejścia/wyjścia (I/O). W klasie głównej *TowerDefense*, dziedziczącej po bazowej klasie *Game* platformy RUT-AI, zaimplementowano pole *outputSpec*. Formatuje ono aktualny stan obiektu *TowerDefenseState* do ustandaryzowanego łańcucha znaków, zrozumiałego dla zewnętrznych agentów sztucznej inteligencji.

Równie istotne jest mapowanie wejść (działań użytkownika/agenta). Dyskretna przestrzeń akcji została zdefiniowana za pomocą konfiguracji *players*:

- Sterowanie nawigacyjne (*move*): przypisane bezpośrednio do zdarzeń klawiatury (klawisze kierunkowe oraz W, S, A, D) i zmapowane na dyskretne

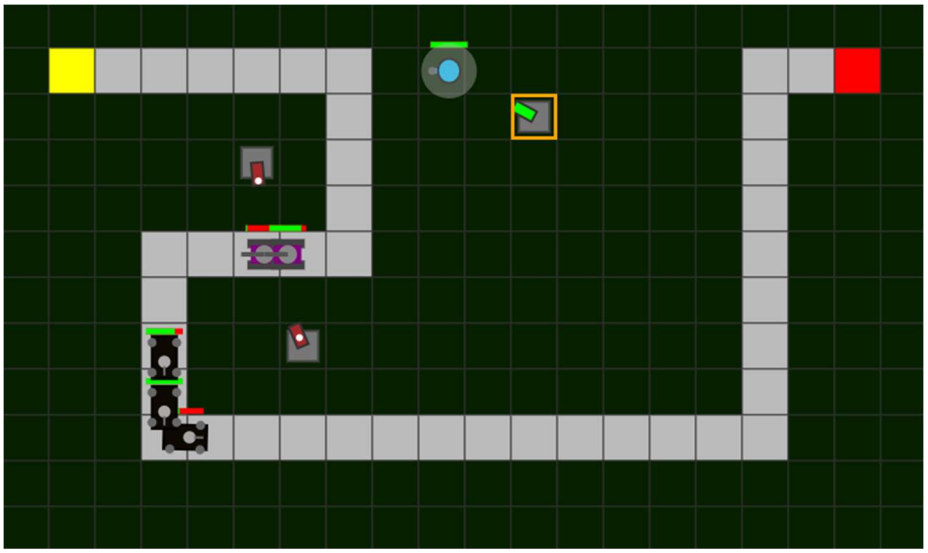
wartości numeryczne {1,2,3,4} - odpowiednio ruch w górę, dół, lewo, prawo. System ten modyfikuje pozycję wskaźnika wyboru na siatce (*cursorX* oraz *cursorY*), co stanowi idealnie zdefiniowaną przestrzeń akcji dla przyszłych agentów AI.

- Interakcje i decyzje (*action*): przypisane do klawiszy: Spacja, Enter, Q oraz E, zmapowane na wartości {1,2,3,4} odpowiednio: budowa/ulepszenie wieży, start kolejnej fali, zmiana typu wieży, sprzedaż wieży.

Taka parametryzacja wejść pozwala na łatwe zastąpienie sterowania manualnego (klawiatury) przez sieć neuronową lub algorytm ewolucyjny, który będzie podejmował decyzję przez wysyłanie prostych sygnałów liczbowych.

4.2. Proceduralne renderowanie warstwy wizualnej

W przeciwieństwie do tradycyjnego podejścia opartego na gotowych zasobach graficznych (ang. *sprites*), w projekcie zdecydowano się na autorskie, proceduralne renderowanie obiektów w czasie rzeczywistym (rys. 2.). Wykorzystanie HTML5 Canvas API [5] pozwala na płynne renderowanie stanu gry na podstawie danych z obiektu *TowerDefenseState* bez obciążania systemu zewnętrznymi silnikami graficznymi [6].



Rys. 2. Zrzut ekranu z uruchomionej aplikacji Tower Defense prezentujący efekty renderowania proceduralnego obiektów oraz interfejs rozgrywki przy użyciu interfejsu Canvas API

Fig. 2. Screenshot of the running Tower Defense application presenting the effects of procedural object rendering and the gameplay interface using the Canvas API

Przeprowadzone testy wydajnościowe wykazały, że zastosowane podejście pozwala na utrzymanie stabilnych 60 klatek na sekundę (FPS) w najpopularniejszych przeglądarkach internetowych (m.in. Google Chrome oraz Mozilla Firefox),

nawet przy jednoczesnym renderowaniu ponad 100 dynamicznych obiektów na ekranie. Wszystkie jednostki występujące w tej grze (przeciwnicy, wieże, pociski) są generowane dynamicznie za pomocą zestawu instrukcji rysowania kształtów. Przykładem wykorzystania tego mechanizmu jest funkcja odpowiedzialna za renderowanie kursora wyboru (Listing 2.), która reaguje na stan ekonomiczny gracza, zmieniając kolor obramowania na zielony (*lime*) lub pomarańczowy (*orange*) w zależności od tego, czy gracz posiada wystarczającą ilość funduszy (*gold*) na budowę wybranej wieży na danym kafelku.

```
let canAfford = false;
// sprawdzenie obecności wieży
const towerOnTile = state.towers.find(t => t.x === cursorX && t.y === cursorY);
if (towerOnTile) {
   const towerData = TowerTypes[towerOnTile.type.toUpperCase() as keyof typeof
TowerTypes];
   if (towerData.upgradeCost) {
      // weryfikacja funduszy na ulepszenie
      canAfford = gold >= towerData.upgradeCost;
   }
} else {
   const selectedData = getSelectedTowerData();
   if (selectedData.cost > 0) {
      // weryfikacja funduszy na budowę
      canAfford = gold >= selectedData.cost;
   }
}

// wizualna informacja zwrotna
context.strokeStyle = canAfford ? 'lime' : 'orange';
context.lineWidth = 3;
context.strokeRect(x + 1.5, y + 1.5, tileSize - 3, tileSize - 3);
```

Listing 2. Fragment implementacji wizualnej informacji zwrotnej w oparciu o stan gry
Listing 2. Fragment of the visual feedback implementation based on the game state

Przykładowo renderowanie przeciwnika typu *TANK* (Listing 3.) polega na złożeniu obiektu z kilku prymitywów np. prostokąt imitujący kadłub i gąsienice. Dzięki takiemu podejściu umożliwiono łatwą implementację efektów wizualnych np. przezroczyste kręgi zasięgu wież (*rgba*) widoczne w fazie przygotowawczej do fali.

```
case 'TANK': {
   const tankWidth = currentRadius * 2;
   const tankLength = currentRadius * 2.2;
   context.fillStyle = '#444';
   context.fillRect(-tankLength / 2, -tankWidth / 2, tankLength, tankWidth * 0.3);
```

```
    context.fillRect(-tankLength / 2, tankWidth / 2 - tankWidth * 0.3, tankLength, tank-
Width * 0.3);

    context.fillStyle = enemy.color;
    context.fillRect(-tankLength / 2 * 0.8, -tankWidth / 2 * 0.6, tankLength * 0.8, tank-
Width * 0.6);
    context.fillStyle = '#888';
    context.beginPath();
    context.arc(0, 0, currentRadius * 0.6, 0, Math.PI * 2);
    context.fill();
    context.fillStyle = '#555';
    context.fillRect(0, -currentRadius * 0.15, currentRadius * 1.5, currentRadius * 0.3);
    break;
}
```

Listing 3. Fragment kodu odpowiedzialny za proceduralne rysowanie jednostki typu TANK

Listing 3. Code fragment responsible for procedural drawing of a TANK unit

4.3. Przestrzenna detekcja celów i logika ataku

Integralną częścią silnika gry jest tzw. pętla walki, odpowiadająca za interakcje przestrzenne między wieżami a jednostkami wroga. Ponieważ gra działa w czasie rzeczywistym, system musi w każdej klatce (ang. *frame*) oceniać stan otoczenia dla każdej aktywnej struktury obronnej. Proces ten jest realizowany w dwóch etapach: poprzez znalezienie odpowiedniego celu oraz obliczenie trajektorii ataku.

Proces obierania celu opiera się na klasycznej, matematycznej kalkulacji dystansu euklidesowego pomiędzy środkiem wieży a pozycją przeciwnika [4]:

$$d = \sqrt{(x_e - x_t)^2 + (y_e - y_t)^2} \tag{2}$$

gdzie: x_e - pozycja na osi X wroga,
y_e - pozycja na osi Y wroga,
x_t - pozycja na osi X wieży,
y_t - pozycja na osi Y wieży.

Algorytm (Listing 4.) iteruje po tablicy aktywnych wrogów. Jeśli wyliczony dystans jest mniejszy lub równy zasięgowi wieży (*range*), następuje weryfikacja logiki domenowej – system sprawdza za pomocą flag konfiguracyjnych, czy dana wieża jest uprawniona do ataku na konkretny typ jednostki wroga (np. weryfikacja parametrów *isFlying* oraz *canHitAir*).

```
private findTarget(tower: ITower): IEnemy | undefined {
    // wyznaczenie środka geometrycznego wieży
    const towerCenterX = (tower.x + 0.5) * this.game.state.tileSize;
    const towerCenterY = (tower.y + 0.5) * this.game.state.tileSize;
```

```
    const towerData = TowerTypes[tower.type.toUpperCase() as keyof typeof
TowerTypes];
    for (const enemy of this.game.state.enemies) {
        // obliczenie dystansu euklidesowego do celu
        const distance = Math.sqrt(Math.pow(enemy.x - towerCenterX, 2) +
Math.pow(enemy.y - towerCenterY, 2));
        // filtracja jednostek wroga poza zasięgiem strzału
        if (distance > tower.range) {
            continue
        }
        // weryfikacja uprawnień do ataku
        if (enemy.isFlying && !towerData.canHitAir) {
            continue;
        }
        if (!enemy.isFlying && !towerData.canHitGround) {
            continue;
        }
        return enemy; // cel znaleziony
    }
    return undefined; // brak celów
}
```

Listing 4. Fragment kodu odpowiedzialny za detekcję celu

Listing 4. Code fragment responsible for target detection

Po pomyślnej identyfikacji celu, system przechodzi do kalkulacji wizualnych i inicjacji ataku. Aby wieża mogła płynnie śledzić poruszający się obiekt, jej kąt rotacji obliczany jest w każdej klatce z wykorzystaniem odwrotnej funkcji trygonometrycznej arcus tangens (*Math.atan2*). Zwraca ona kąt w radianach pomiędzy osią X a punktem docelowym, co pozwala na realistyczne renderowanie obrotu lufy wieży przed oddaniem strzału. Wystrzelenie pocisku (wygenerowanie obiektu klasy *Bullet*) odbywa się asynchronicznie względem klatek animacji – jest warunkowana przez wewnętrzny licznik *cooldown*, który po osiągnięciu limitu wyznaczonego przez parametr *fireRate*, ulega wyzerowaniu.

5. Model ekonomiczny i parametryzacja rozgrywki

Kluczowym aspektem gier typu Tower Defense, warunkującym zaangażowanie gracza i sensowność wyzwania, jest odpowiedni balans ekonomiczny (tzw. *game balance*) [1]. W zrealizowanym projekcie mechanika finansowa i trudność zostały sparametryzowane w oparciu o analizę efektywności kosztowej wież oraz specjalizację jednostek przeciwnika.

5.1. Analiza efektywności bojowej i kosztowej wież

Aby zbalansować rozgrywkę, parametry wież zostały zróżnicowane pod kątem obrażeń zadawanych na sekundę (ang. *Damage Per Second - DPS*) w stosunku do ich kosztu. Zakładając, że główna pętla gry aktualizowana jest z częstotliwością 60 klatek na sekundę, wartość *DPS* wyliczana jest na podstawie bazowych obrażeń (*O*) i parametru *fireRate* (*C*), określającego przerwę między strzałami w klatkach obrazu:

$$DPS = O \times \frac{60}{C} \tag{3}$$

gdzie: O - bazowe obrażenia,
C - częstotliwość strzelania (*fireRate*).

Wydajność ekonomiczną wieży (*W*) można zdefiniować jako stosunek *DPS* do kosztu budowy (*K*):

$$W = \frac{DPS}{K} \tag{4}$$

gdzie: DPS - Obrażenia zadawane na sekundę,
K - Koszt wieży.

Zestawienie podstawowych struktur obronnych zaimplementowanych w projekcie prezentuje tabela 1.

Tabela 1. Zestawienie podstawowych struktur obronnych

Table 1. Summary of basic defensive structures

Typ Wieży (Kod)	Koszt budowy	Obrażenia	Fire Rate	Obliczony DPS	Wydajność (W)
BASIC	100	48	30	96	0.96
CANNON	200	72	90	48	0.24
AA_GUN	150	105	30	210	1.40

Analiza danych pokazuje, że wieża przeciwlotnicza (*AA_GUN*) posiada doskonały współczynnik wydajności 1,40, jednak zostało to okupione silną restrykcją w strzelaniu – wieża nie może atakować jednostek naziemnych (*canHitGround: false*). Z kolei wieża *CANNON* posiada najniższą efektywność, jednak wynika to z faktu zadawania przez nią obrażeń obszarowych, które nie zostały tutaj uwzględnione.

5.2. Parametryzacja ulepszeń i mechanika obrażeń obszarowych

Model ekonomiczny gry został wzbogacony o mechanizm ulepszania struktur (ang. *upgrades*). Zjawisko to sparametryzowano w oparciu o prawo malejącej stopy zwrotu. Przykładowo, ulepszenie wieży *BASIC* do poziomu drugiego kosztuje dwukrotnie więcej niż bazowa budowa nowej wieży (koszt: 200), podnosząc obrażenia o 50% (z 48 na 72). Taki model zmusza gracza do analizy: inwestycja w wieżę w dobrym miejscu (np. na zakręcie mapy), czy budowa nowej wieży w mniej dogodnej pozycji. Warto zaznaczyć, że inwestycja w strukturę obronną wpływa także na jej promień zasięgu (zwiększenie parametru *range*).

Dodatkowym elementem bilansującym jest parametr *splashRadius* w wieżach typu *CANNON*. W ich przypadku wprowadzono pojęcie efektywnego *DPS* (DPS_{eff}), który rośnie proporcjonalnie do liczby jednostek wroga znajdujących się w polu rażenia:

$$DPS_{eff} = DPS \times n \tag{5}$$

gdzie: DPS – obrażenia na sekundę,
n – liczba jednostek wroga w polu rażenia.

Mechanika ta optymalizuje obronę przeciwko falom składającym się z dużej liczby słabych, zgrupowanych przeciwników.

5.3. Pętla zasobów i skalowanie poziomu trudności

Ekonomia gry opiera się na eliminacji przeciwników, która generuje walutę pozwalającą na ekspansję infrastruktury obronnej. Wartości nagród zostały rygorystycznie skorelowane z punktami życia (*HP*). Zwykła jednostka (*TANK*, 350 HP) generuje nagrodę w wysokości 10 jednostek waluty, co stanowi podstawę wczesnej ekonomii. Skalowanie trudności odbywa się poprzez tablicę definicji fal (*WaveDefinitions*), która w miarę postępu czasu wprowadza na mapę jednostki o wyższych parametrach, takie jak *BOSS_TANK* (4000 HP). Eliminacja bossa nagradzana jest kwotą 200 jednostek waluty, ale jego przepuszczenie do punktu docelowego skutkuje surową karą, odejmując 10 punktów życia bazy (baza na starcie mapy posiada 20 punktów życia). Standardowe jednostki zabierają jedynie 1 punkt życia bazy. Wymusza to na graczu ciągłą reinwestycję zasobów.

5.4. Kategoryzacja i parametryzacja jednostek przeciwnika

Aby zapobiec monotonii i wymusić na graczu ciągłą adaptację strategii, zaimplementowano pięć unikalnych klas przeciwników. Ich parametry zostały dobrane w taki sposób, aby testować różne słabe punkty systemów obronnych (tabela 2.).

Tabela 2. Zestawienie przeciwników

Table 2. Summary of enemy units

Typ jednostki (Kod)	Punkty życia (HP)	Prędkość bazowa	Nagroda (Gold)	Typ poruszania
TANK	350	0.7	10	Naziemny
VEHICLE	220	1.5	5	Naziemny
BOSS_TANK	4000	0.5	200	Naziemny
HELICOPTER	175	1.2	15	Powietrzny
JET	120	2.5	15	Powietrzny

Każda z jednostek pełni ściśle określoną funkcję:

- Jednostki podstawowe (*TANK*): Stanowią fundament ekonomiczny gry. Charakteryzują się optymalnym stosunkiem punktów życia do prędkości.
- Jednostki szybkie (*VEHICLE*): Testują zdolność układu obronnego do szybkiej reakcji. Mają blisko 40% mniej HP niż standardowy czołg, ale poruszają się dwukrotnie szybciej. Oferują również niską nagrodę (5 jednostek waluty).
- Cele lotnicze (*JET*, *HELICOPTER*): Całkowicie ignorują układ labiryntu (co opisano w rozdziale 3.3). Wymuszają budowę struktur typu *AA_GUN*.
- Jednostka specjalna (*BOSS_TANK*): Jej pojawienie się wymusza kumulację obrażeń z wielu źródeł jednocześnie. Nagroda (200 jednostek waluty) stanowi potężny zastrzyk gotówki pozwalający na rozwój bazy.

5.5. Iteracyjny proces balansowania i testy z użytkownikami

Ważnym etapem kształtowania modelu ekonomicznego oraz parametryzacji poziomu trudności był iteracyjny proces tworzenia rozgrywki (ang. *playtesting*). Ze względu na to, że czyste wyliczenia matematyczne nie zawsze w pełni oddają dynamikę gry w czasie rzeczywistym, wczesne wersje modułu zostały udostępnione wybranej grupie testowej. Grupę tę stanowili inni współtwórcy ekosystemu RUT-AI Games, co zapewniło wysoki poziom kompetencji technicznych testerów. W zorganizowanych badaniach wzięło udział 5 testerów, z których każdy rozegrał po 10 pełnych sesji. Zebrana od nich informacja zwrotna (ang. *feedback*) posłużyła do empirycznej weryfikacji początkowych założeń. Na podstawie obserwacji zachowań graczy i ich decyzji przeprowadzono 3 główne iteracje testowe, podczas których dokonano edycji parametrów. Pozwoliło to między innymi na:

- Wyeliminowanie tzw. strategii dominujących, czyli sytuacji, w której budowa wyłącznie jednego typu wieży (w tym przypadku *BASIC*) gwarantowała wygraną. Osiągnięto to m.in. poprzez zmniejszenie parametru bazowych obrażeń dla wieży *BASIC* z 60 na 48 oraz zwiększenie obrażeń zadawanych przez wieżyczki *AA_GUN* z 90 na 105.

- Odpowiednie dostrojenie nagród za pokonanie wrogów, co polegało np. na zwiększeniu nagrody za zniszczenie podstawowych czołgów z 5 do 10 jednostek waluty.

Zastosowanie takiego podejścia udowodniło, że ostateczny kształt parametrów jest nie tylko poprawny matematycznie, ale zapewnia przede wszystkim angażujące oraz zbalansowane doświadczenie użytkownika.

6. Kierunki dalszego rozwoju aplikacji

Zaimplementowana gra Tower Defense stanowi elastyczny fundament pod dalszą rozbudowę. Architektura systemu jest otwarta na modyfikacje, co pozwala na wyznaczenie głównych planów rozwoju w kolejnych iteracjach projektu.

6.1. Integracja z algorytmami uczenia ze wzmocnieniem

Biorąc pod uwagę, że gra stanowi moduł edukacyjnej platformy RUT-AI Games, jej naturalnym rozwinięciem jest implementacja inteligentnych agentów potrafiących samodzielnie grać i optymalizować strategie obronne. Przygotowany format wyjścia (*outputSpec*) oraz dyskretna przestrzeń akcji (*players*) tworzą gotowe środowisko dla algorytmów uczenia ze wzmocnieniem.

6.2. Proceduralne generowanie poziomów

Obecnie mapy w grze przechowywane są jako statyczne, predefiniowane w pliku konfiguracyjnym (*towerdefense.maps.ts*). Kolejnym krokiem mogłaby być implementacja systemu proceduralnego generowania zawartości. Przykładowo wykorzystując algorytmy takie jak błądzenie losowe lub automaty komórkowe, system mógłby generować nieskończoną liczbę unikalnych map przy każdym uruchomieniu gry. Warunkiem brzegowym dla takiego generatora byłaby walidacja (np. z użyciem istniejącego algorytmu BFS) sprawdzająca, czy wygenerowana losowo trasa posiada przejście od punktu startowego do bazy. Wprowadzenie takiego generowania zwiększyłoby grywalność aplikacji i dostarczyło zróżnicowanych danych testowych dla uczących się algorytmów sztucznej inteligencji.

6.3. Rozszerzona ewaluacja z udziałem niezależnej grupy użytkowników

Z uwagi na to, że dotychczasowe testy gry przeprowadzono w wykwalifikowanym gronie współtwórców platformy RUT-AI Games, obecne wyniki mogą nie w pełni reprezentować zachowania typowego odbiorcy końcowego. Współtwórcy projektu znają mechaniki platformy od strony technicznej, przez co ich styl gry różni się od przeciętnego gracza. Zaplanowanym kierunkiem dalszych prac jest udostępnienie modułu szerszej, niezależnej grupie badawczej. Obserwacja rozgrywki osób z zewnątrz pozwoli na jeszcze dokładniejsze dostosowanie poziomu trudności i ekonomii gry do potrzeb zwykłych użytkowników.

7. Podsumowanie

Głównym celem niniejszej pracy było zaprojektowanie, implementacja oraz zbalansowanie w pełni grywalnej aplikacji z gatunku Tower Defense, stanowiącej nowy moduł platformy edukacyjnej RUT-AI Games. Cel ten został w pełni zrealizowany, a ostateczny produkt pokazuje, że nowoczesne technologie webowe pozwalają na tworzenie wydajnych, złożonych środowisk symulacyjnych bez konieczności polegania na zewnętrznych, ciężkich silnikach graficznych.

W trakcie realizacji projektu z powodzeniem zaimplementowano szereg zaawansowanych mechanizmów inżynierii oprogramowania. Zastosowanie języka TypeScript pozwoliło na stworzenie bezpiecznej, modularnej architektury, przygotowując tym samym idealne środowisko interfejsów (I/O) dla przyszłych agentów sztucznej inteligencji. Silnik renderujący, oparty na proceduralnym rysowaniu kształtów za pomocą API Canvas, zapewnił wysoką wydajność i płynność działania aplikacji. Jak potwierdziły testy, utrzymanie 60 FPS przy znacznym obciążeniu obiektami na różnych przeglądarkach udowadnia, że gra nie wymaga wsparcia zewnętrznych silników. Rozwiązano także kluczowe problemy algorytmiczne dla tego gatunku gier. Dynamiczna implementacja przeszukiwania wszerz (BFS) na dyskretnej siatce dwuwymiarowej zagwarantowała optymalne wyznaczanie najkrótszych ścieżek dla jednostek naziemnych, zachowując jednocześnie elastyczność w obsłudze jednostek latających. Dodatkowo zrównoważenie modelu ekonomicznego obejmujące efektywność bojową wież, mechanikę obrażeń obszarowych oraz system fal przeciwników – wszystko wsparte praktycznymi testami - zaowocowało stworzeniem zbalansowanego środowiska.

Zrealizowany projekt jest wzbogaceniem ekosystemu platformy RUT-AI Games o solidne ramy do testowania algorytmów. Stanowi on doskonały punkt wyjścia do dalszych prac naukowych, w szczególności w zakresie implementacji sztucznej inteligencji oraz proceduralnego generowania map.

Literatura oraz źródła internetowe

[1] Adams E.: Fundamentals of Game Design, New Riders, Berkeley 2013.

[2] Cormen T. H., Leiserson C. E., Rivest R. L., Stein C.: Wprowadzenie do algorytmów, Wydawnictwo Naukowe PWN, Warszawa 2012.

[3] Nystrom R.: Game Programming Patterns, Genever Benning, 2014.

[4] https://developer.mozilla.org/en-US/docs/Games/Techniques/2D_collision_detection [dostęp: 17.02.2026].

[5] https://developer.mozilla.org/en-US/docs/Web/API/Canvas_API [dostęp: 17.02.2026].

[6] https://developer.mozilla.org/en-US/docs/Web/API/Canvas_API/Tutorial/Drawing_shapes [dostęp: 17.02.2026].

[7] https://rutai.kia.prz.edu.pl/ [dostęp: 17.02.2026].

[8] https://www.redblobgames.com/pathfinding/a-star/introduction.html#breadth-first-search [dostęp: 17.02.2026].

[9] https://www.typescriptlang.org/docs/ [dostęp: 17.02.2026].

DESIGN AND IMPLEMENTATION OF A TOWER DEFENSE GAME MODULE IN THE ENVIRONMENT OF THE RUT-AI EDUCATIONAL PLATFORM USING TYPESCRIPT

Summary

This article presents the process of designing, implementing, and parameterizing a two-dimensional Tower Defense game. The application was created as a new component of the RUT-AI Games educational web platform, whose main goal is to provide an environment for testing and evaluating artificial intelligence algorithms. The project was carried out using the strongly typed TypeScript language, which allowed for the preparation of a clear, modular architecture and facilitated integration with the Angular framework, on which the entire platform is based. The visual layer, including the rendering of objects and animations in the browser, is handled by the HTML5 Canvas application programming interface in a two-dimensional context. The paper discusses key aspects of video game software engineering. The main element of the enemy units' movement mechanics was the application of the Breadth-First Search (BFS) algorithm, which was adapted for the dynamic determination of the shortest paths on a discrete spatial grid. Additionally, an analysis of the game's economic model was conducted, including the parameterization of defensive tower construction costs, evaluation of their combat effectiveness (damage per second metric), and mechanisms for scaling the difficulty level in successive waves of opponents. The completed project is a fully functional and balanced application. Thanks to a well-thought-out architecture and the application of widely recognized design patterns, the game constitutes an excellent foundation for further research and development work, primarily for the implementation of advanced machine learning systems within the RUT-AI Games ecosystem.

Keywords: Tower Defense, TypeScript, BFS algorithm, game economy, web software architecture.

Krystian Sołtys, Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST

dr hab. inż. Tomasz Kapuściński, prof. PRz, dr inż. Dawid Kalandyk
opiekunowie Koła Naukowego

EWOLUCJA INTELIGENTNEGO URZĄDZENIA REHABILITACJI RUCHOWEJ

Streszczenie

Rozdział przedstawia proces ewolucji i optymalizacji inteligentnego urządzenia do rehabilitacji ruchowej, pełniącego funkcję autonomicznej wyrzutni piłek sterowanej gestami dłoni. Głównym celem pracy było rozwiązanie problemów zidentyfikowanych w pierwszym prototypie, takich jak trudności w serwisowaniu wynikające z monolitycznej budowy, przegrzewanie się komponentów, chaos w okablowaniu oraz niska wydajność systemu wizyjnego. W ramach metod badawczych zastosowano iteracyjne projektowanie CAD, technologię przyrostową (druk 3D) do budowy elementów mechanicznych oraz analizę wydajnościową algorytmów sztucznej inteligencji. W warstwie mechanicznej opracowano nową, w pełni modułową podstawę, podzieloną na trzy niezależne sekcje: zasilającą z wewnętrznym przyłączem na napięcie 230V, obrotową ze zintegrowanym systemem prowadzenia przewodów oraz sterująco-obliczeniową bazującą na mikrokomputerze Raspberry Pi. Istotnym wynikiem w warstwie oprogramowania jest pomyślna implementacja technologii przetwarzania brzegowego. Przeniesienie analizy obrazu bezpośrednio na układ obliczeniowy kamery OAK-D pozwoliło na odciążenie głównego procesora. Pomiary porównawcze wykazały ponad dwukrotny wzrost wydajności systemu wizyjnego, z 13,4 klatek na sekundę w pierwszej wersji do 28,1 klatek na sekundę w układzie zmodernizowanym. Główne wnioski z przeprowadzonych prac wskazują, że wprowadzone modyfikacje architektoniczne i programistyczne wyeliminowały problemy termiczne oraz znacząco zwiększyły responsywność maszyny. Zaproponowane rozwiązanie stanowi niezawodne i bezpieczne narzędzie asystujące, które może skutecznie wspierać zautomatyzowane treningi motoryczne, odciążając personel medyczny w codziennej pracy z pacjentami.

Słowa kluczowe: rehabilitacja ruchowa, przetwarzanie brzegowe, robotyka modułowa, ewolucja projektu, druk 3D.

1. Wstęp

Pierwsza wersja omawianego urządzenia stanowiła prototyp inteligentnej wyrzutni piłeczek, zaprojektowanej z myślą o wsparciu procesu rehabilitacji oraz utrzymywania możliwości ruchowych ludzkiego organizmu. Głównym założeniem tamtego projektu było zbudowanie wyrzutni potrafiącej autonomicznie śledzić cel i reagować na gesty dłoni wykonywane przez pacjenta, co miało stanowić inżynierską odpowiedź na rosnące potrzeby placówek medycznych [1].

Konstrukcja tej pierwotnej iteracji opierała się w głównej mierze na elementach mechanicznych wytworzonych w technologii druku 3D. Za zmysł wzroku robota odpowiadała kamera OAK-D, natomiast mózgiem całego układu, sterującym silnikami krokowymi oraz turbiną odpowiedzialną za wyrzut piłeczek do ping-ponga, była jednostka obliczeniowa Raspberry Pi. Wykonanie pierwszego prototypu pozwoliło na zweryfikowanie założeń teoretycznych oraz rozszerzenie kompetencji autorów w zakresie elektroniki, robotyki i praktycznej implementacji sztucznej inteligencji.

Jednakże proces budowy i rygorystycznego testowania tego prototypu pozwolił na zidentyfikowanie szeregu ważnych problemów, których wyeliminowanie okazało się niezbędne, aby odpowiednio ulepszyć robota i przystosować go do realnej pracy z użytkownikiem. Wśród głównych mankamentów pierwszej wersji znalazły się między innymi trudności z odprowadzaniem ciepła prowadzące do nagrzewania się silnika krokowego, skomplikowany demontaż elementów oraz ogólna potrzeba zwiększenia wydajności obliczeniowej.

Mając to na uwadze, niniejszy rozdział przedstawia ewolucję i proces optymalizacji urządzenia. Skupiono się w nim na omówieniu nowej, w pełni modułowej konstrukcji mechanicznej, przeprojektowaniu mechanizmów miotających i zasilających, a także na wdrożeniu technologii przetwarzania brzegowego (Edge AI) [6] w systemie wizyjnym. Wprowadzone innowacje mają na celu maksymalizację wydajności, bezpieczeństwa oraz ergonomii, przybliżając prototyp do formy w pełni funkcjonalnego i niezawodnego asystenta rehabilitacji ruchowej.

2. Założenia projektowe i koncepcja nowej generacji

Po zbudowaniu i przetestowaniu pierwszego prototypu urządzenia, zidentyfikowano szereg problemów oraz kwestii wymagających stanowczej poprawy. Sam proces testowania i wprowadzania drobnych poprawek pozwolił na dokładną ewaluację błędów konstrukcyjnych, które stały się punktem wyjścia do opracowania całkowicie nowej, zoptymalizowanej generacji wyrzutni.

2.1. Identyfikacja problemów i ograniczeń pierwszego prototypu

Analiza działania pierwszej wersji urządzenia ujawniła ograniczenia zarówno w warstwie mechanicznej, jak i sprzętowo-programowej. Konieczność wielokrot-

nego rozkładania i składania robota w celu wykonania nawet drobnych korekt okazała się wysoce problematyczna. Monolityczna budowa podstawy sprawiała, że serwisowanie układu było nieefektywne. Wewnątrz konstrukcji brakowało również wystarczającej ilości miejsca na ułożenie elementów elektronicznych oraz odpowiednich przelotów na okablowanie. Ponadto lufa okazała się zbyt krótka, a jej szerokość niedopasowana do wielkości piłeczki. Powodowało to zbyt małą prędkość wylotową, co w konsekwencji nie pozwalało na precyzyjne trafianie w zadany cel.

System zasilania oparty na zewnętrznych zasilaczach (osobnych dla mikrokomputera Raspberry Pi oraz dla silników krokowych) generował dużą ilość luźnych przewodów. Brak odpowiedniego zarządzania okablowaniem nie tylko negatywnie wpływał na estetykę, ale stwarzał realne zagrożenie splątania kabli podczas obrotu urządzenia w osi poziomej.

Scentralizowane przetwarzanie obrazu, w którym mikrokomputer Raspberry Pi odpowiadał jednocześnie za sterowanie silnikami, logikę układu oraz analizę klatek wideo z kamery DepthAI [4], prowadziło do znacznych opóźnień. Zbyt mała wydajność przeliczania klatek na sekundę (FPS) negatywnie wpływała na responsywność maszyny na gesty użytkownika.

2.2. Wdrożenie przetwarzania brzegowego (Edge AI)

W odpowiedzi na zidentyfikowane problemy, postanowiono zmienić architekturę urządzenia. W nowej koncepcji zmieniono sposób mocowania elementów oraz obudowy na taki, który zdecydowanie ułatwia proces ich demontażu. Głównym założeniem stał się podział podstawy na trzy niezależne, logiczne sekcje. Wprowadzono wydzielony moduł zasilania, nowy łącznik obrotowy wyposażony w system bezpiecznego prowadzenia przewodów oraz osobną część mieszczącą sterowniki i jednostkę obliczeniową. Dzięki takiemu rozwiązaniu zapewniono znacznie więcej miejsca na wszystkie elementy oraz większe przeloty na okablowanie.

Aby całkowicie uniezależnić robota od obecności zewnętrznych źródeł zasilania, przeprojektowano układ elektroniczny. W drugiej wersji prototypu zasilacze zostały na stałe zamknięte i zabezpieczone wewnątrz dolnej obudowy. Na zewnątrz wyprowadzono jedynie gniazda zasilające przystosowane do standardowego napięcia sieciowego 230V, eliminując tym samym problem wystających przewodów zasilających i podnosząc ergonomię całego systemu.

2.3. Założenia optymalizacyjne nowej wersji wyrzutni

Ostatnim istotnym elementem nowej koncepcji jest optymalizacja wydajności inteligentnego systemu sterowania. W pierwszej iteracji scentralizowane przetwarzanie obrazu oraz jednoczesne sterowanie mechaniką realizowane w całości przez jednostkę centralną Raspberry Pi prowadziło do zauważalnych opóźnień.

Aby usprawnić ten proces, autorzy zdecydowali się na wykorzystanie specjalnego oprogramowania dostarczanego z kamerą DepthAI OAK-D. Wprowadzono

technologię przetwarzania brzegowego (Edge AI), uruchamiając rozpoznawanie gestów dłoni w tzw. trybie krawędziowym (Edge Mode). Oznacza to przeniesienie głównego ciężaru analizy obrazu bezpośrednio na układ obliczeniowy wbudowany w samą kamerę. Odciążenie magistrali komunikacyjnej i procesora Raspberry Pi pozwala na uzyskanie znacznie wyższej płynności działania, co przekłada się na szybszą reakcję robota na ruchy rehabilitowanego użytkownika.

3. Modułowa architektura konstrukcji mechanicznej

W odpowiedzi na zidentyfikowane problemy związane z koniecznością wielokrotnego rozkładania i składania robota w celu wykonania drobnych korekt, proces projektowania nowej generacji urządzenia skupił się na całkowitej reorganizacji jego budowy. Postanowiono zmienić sposób mocowania elementów oraz obudowy na taki, który zdecydowanie ułatwia proces ich demontażu. Zrezygnowano z monolitycznej bryły na rzecz w pełni modułowej architektury, zachowując przy tym założenie, że większość elementów zostanie zrealizowana za pomocą technologii druku 3D.

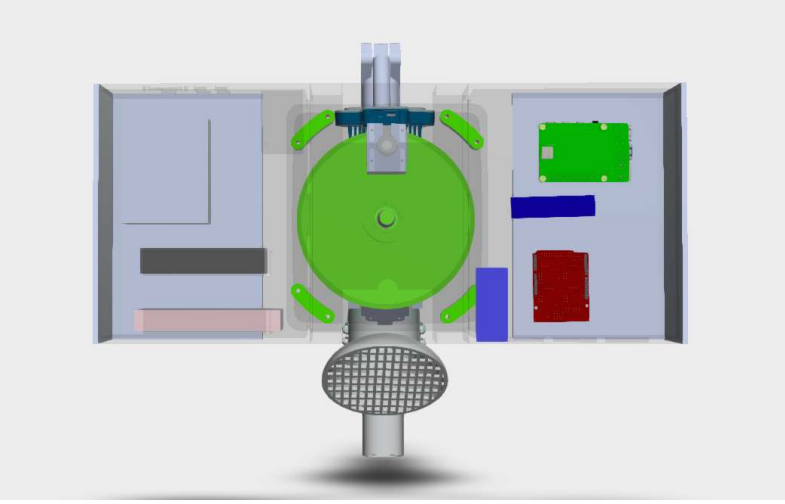
3.1. Trójdzielna konstrukcja podstawy urządzenia

Największą innowacją w warstwie mechanicznej jest nowa podstawa, przedstawiona na rys. 1., która została rozdzielona na trzy odrębne części. Konstrukcja ta umożliwia niezależne rozłączanie poszczególnych modułów, co znacząco redukuje czas potrzebny na ewentualne prace serwisowe oraz ułatwia testowanie poszczególnych podzespołów. Zapewniono dzięki temu więcej miejsca na wszystkie elementy oraz wyznaczono większe przeloty na okablowanie.

Po lewej stronie urządzenia wydzielono dedykowany moduł zasilania. Jego implementacja jest bezpośrednią odpowiedzią na chęć uniezależnienia robota od obecności zewnętrznych źródeł zasilania w postaci nieporęcznych zasilaczy wtyczkowych. Znajduje się w niej zasilacz produkujący napięcie stałe 12V, akumulator litowo-polimerowy oraz sterownik turbiny odpowiadającej za wyrzut piłeczki.

Po prawej stronie podstawy zlokalizowano moduł sterujący i obliczeniowy. Znalazło się tam miejsce dla sterowników obsługujących silniki krokowe, a także dla centralnej jednostki obliczeniowej, którą w dalszym ciągu pozostaje układ Raspberry Pi.

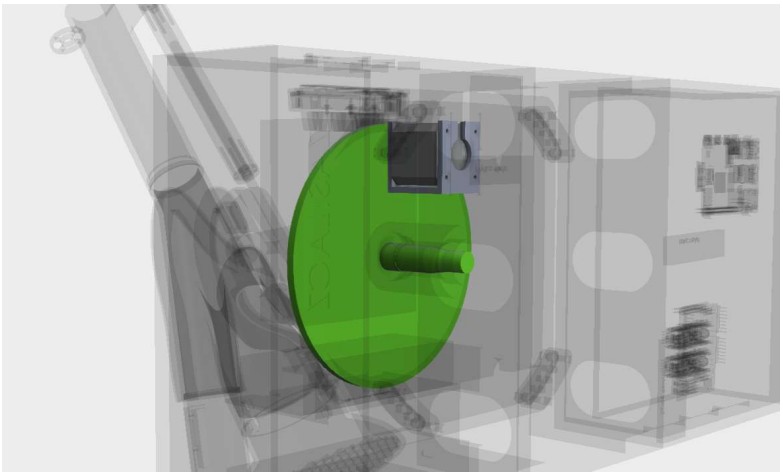
Środkowa część konstrukcji pełni funkcję głównego węzła mechanicznego i komunikacyjnego. Znajduje się w niej łącznik na przewody, zintegrowany z obracaną podstawą wyrzutni. Moduł ten odpowiada za bezpieczne przeprowadzenie wiązek elektrycznych pomiędzy strefą zasilania (lewą) a strefą sterowania (prawą) i elementami ruchomymi górnej partii robota.



Rys. 1. Trójdzielna konstrukcja podstawy urządzenia
Fig. 1. Three-part base design

3.2. Zmodernizowany mechanizm obrotowy

Nowy projekt obracającej się podstawy wyrzutni (rys. 2.), zlokalizowany w centralnym module, został poddany rygorystycznej optymalizacji. Mechanizm obrotu musi charakteryzować się wysoką precyzją, zważywszy na konieczność szybkiego kierowania piłki w stronę dłoni użytkownika wykonującego gest. Nowa konstrukcja węzła obrotowego eliminuje naprężenia kabli, które pojawiały się podczas pracy pierwszego prototypu.



Rys. 2. Mechanizm obrotowy
Fig. 2. Rotating mechanism

3.3. Przeprojektowany układ miotający i nowy magazynek

Zmiany objęły również nadbudowę robota, czyli sekcję odpowiedzialną bezpośrednio za wyrzut piłeczek do ping-ponga, w której za nadanie prędkości wciąż odpowiada wysokowydajna turbina. Lufa została wydłużona i zwężona tak by umożliwić lepszą prędkość wylotową (rys. 3.).



Rys. 3. Przeprojektowany układ miotający z magazynkiem
Fig. 3. Redesigned firing mechanism with magazine

Dodatkowo, dla zwiększenia autonomii pracy robota, opracowano nowy magazynek na piłeczki. Charakteryzuje się on powiększoną pojemnością oraz zmienionym kątem zsypu, co przy współpracy ze zoptymalizowanym mechanizmem przeładowania gwarantuje ciągłe i bezkolizyjne podawanie kolejnych piłek wprost do komory wyrzutowej turbiny.

4. Optymalizacja oprogramowania i systemu wizyjnego

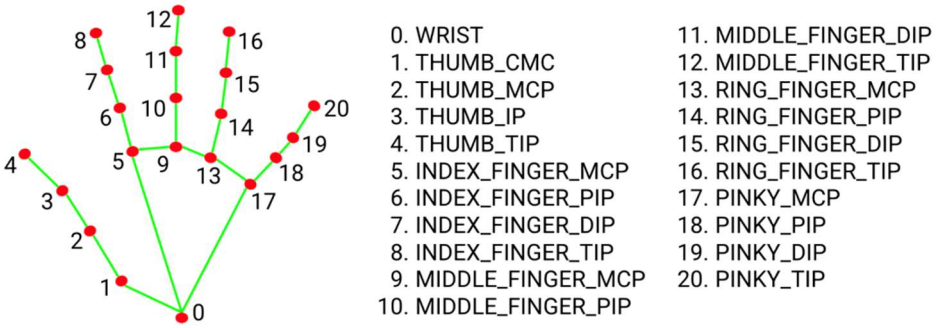
W pierwszej wersji prototypu za wizję robota odpowiadała nowoczesna kamera DepthAI OAK-D. Zdecydowano wówczas, że mózgiem całego układu będzie płytka Raspberry Pi w wersji 4 model B, która pozwalała na instalację systemu Linux wraz z niezbędnym oprogramowaniem i posiadała szereg przydatnych portów sterujących. Całość kodu sterującego pracą robota została wykonana w języku Python [5], a do wykrywania różnych gestów dłoni wykorzystano otwarte oprogramowanie Google MediaPipe. Niestety, architektura oparta na centralizacji zadań, w której mikrokomputer odpowiadał jednocześnie za sterowanie silnikami,

obsługę logiki maszyny stanów oraz wymagającą analitykę obrazu wideo, okazała się niewystarczająca. Prowadziło to do małej wydajności systemu wizyjnego, co objawiało się niską liczbą analizowanych klatek na sekundę oraz irytującymi opóźnieniami w reakcji na ruchy użytkownika.

4.1. Architektura przetwarzania brzegowego (Edge AI)

Zdecydowano się na pełniejsze wykorzystanie potencjału kamery DepthAI OAK-D, która jest fabrycznie wyposażona w specjalne oprogramowanie ułatwiające wykorzystanie obrazu jako źródła danych dla algorytmów przetwarzania obrazu oraz metod sztucznej inteligencji. Zamiast polegać na obliczeniach głównego procesora w Raspberry Pi, wprowadzono nowe oprogramowanie uruchamiające rozpoznawanie dłoni bezpośrednio na kamerze, wykorzystując tak zwany tryb krawędziowy (Edge Mode).

Implementacja przetwarzania brzegowego oznacza, że dedykowana sieć neuronowa odpowiadająca za skanowanie obrazu w poszukiwaniu pojedynczej dłoni ułożonej w gest pięści jest ewaluowana przez wewnętrzny procesor wizyjny (VPU) [3] samej kamery. Wykorzystano w tym celu model MediaPipe Hands [2]. W nowym modelu kamera działa jako inteligentny, autonomiczny sensor. Zamiast przesyłać „ciężki" strumień wideo o wysokiej rozdzielczości do jednostki centralnej, urządzenie wizyjne samodzielnie lokalizuje dłoń i analizuje jej ułożenie. Do głównego układu sterującego Raspberry Pi trafiają jedynie „lekkie", dane o 21 punktach charakterystycznych dłoni (rys. 4.), które następnie pozwalają metodą geometryczną rozpoznać wykonywany gest i na jego podstawie odpowiednio dostosowywane jest sterowanie wyrzutnią.



Rys. 4. Punkty charakterystyczne dłoni w bibliotece MediaPipe

Źródło: https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker (dostęp: 15.02.2026).

Fig. 4. Landmarks of the hand in the MediaPipe library

Source: https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker (accessed: 15.02.2026).

Efektem wprowadzonych prac optymalizacyjnych i zastosowania trybu krawędziowego jest wzrost wydajności całego systemu wizyjnego. Pomiary porównawcze wykazały, że w pierwszej iteracji urządzenia, gdzie główny ciężar analizy obrazu spoczywał na jednostce centralnej, algorytmy rozpoznawania gestów osiągały wydajność na poziomie zaledwie 13,4 klatek na sekundę. Przeniesienie procesów wnioskowania bezpośrednio na procesor kamery w ramach architektury Edge AI pozwoliło na ponad dwukrotny skok tej wartości. Zmodernizowane oprogramowanie, przetestowane z powodzeniem w środowisku laboratoryjnym, wykazuje stabilną pracę z prędkością 28,1 klatek na sekundę. Taka poprawa parametrów gwarantuje wysoce płynną reakcję układu sterowania na ruchy pacjenta, eliminując opóźnienia mogące dekoncentrować użytkownika.

5. Podsumowanie i perspektywy rozwoju

Proces projektowania i udoskonalania inteligentnej wyrzutni piłek stanowi przykład tego, jak ewolucyjne podejście do inżynierii pozwala na sukcesywne eliminowanie błędów i podnoszenie jakości całego systemu. Sam proces testowania i poprawiania pierwszego prototypu umożliwił zidentyfikowanie szeregu problemów, których rozwiązanie stało się fundamentem nowej koncepcji. Należy podkreślić, że zmodernizowana wersja urządzenia w ujęciu mechanicznym znajduje się obecnie w fazie zaawansowanego projektowania i produkcji elementów, nie będąc jeszcze w pełni zmontowaną konstrukcją. Mimo to nowa architektura oprogramowania została już zaimplementowana i zweryfikowana w środowisku testowym. Przejście od monolitycznej bryły do modułowej podstawy, opracowanie nowej prowadnicy oraz bezpieczne zintegrowanie układu zasilania wewnątrz obudowy stanowią naturalny krok w ewolucji urządzenia, zdecydowanie zwiększając jego bezawaryjność oraz ułatwiając przyszłe prace serwisowe.

W najbliższym czasie głównym celem zespołu badawczego jest fizyczne urzeczywistnienie i zmontowanie wszystkich wyprodukowanych modułów drugiej generacji robota. Po integracji nowej mechaniki z wysoce wydajnym oprogramowaniem brzegowym, przeprowadzone zostaną kompleksowe testy obciążeniowe w warunkach rzeczywistych. W dalszych fazach projektu autorzy planują rozwój funkcjonalności urządzenia o kolejne tryby pracy oraz implementacja interfejsu komunikacyjnego pozwalającego na zdalne przesyłanie i sterowanie plikami symulacyjnymi.

Literatura

[1] Kalandyk D., Ciećko K., Sołtys K.: Propozycja inteligentnego urządzenia rehabilitacji ruchowej, IV Konferencja Kół Naukowych w ramach Politechnicznej sieci Via Carpatia im. Prezydenta Lecha Kaczyńskiego, 2024, Lublin.

[2] „Hand landmarks detection guide”
https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker
(dostęp: 15.02.2026 r.).
[3] „Myriad X Product Brief”
www.intel.com/content/dam/www/public/us/en/documents/productbriefs/
myriad-x-product-brief.pdf (dostęp: 15.02.2026 r.).
[4] „OAK-D Datasheet”
https://github.com/luxonis/oak-hardware/blob/master/BW1098OAK_USB3C/Datasheet/
OAK-D_Datasheet.pdf (dostęp: 15.02.2026 r.).
[5] „Python 3 Documentation” https://docs.python.org/3 (dostęp: 15.02.2026 r.).
[6] „Understanding Edge AI: Artificial Intelligence Meets IoT”
https://article.murata.com/en-global/article/what-is-edge-ai (dostęp: 15.02.2026 r.).

EVOLUTION OF THE INTELLIGENT MOTOR REHABILITATION DEVICE

Summary

This article presents the evolution and optimization of an intelligent physical rehabilitation device, functioning as an autonomous, gesture-controlled ball launcher. The main objective was to resolve problems identified in the first prototype, including servicing difficulties caused by a monolithic design, overheating components, wiring chaos, and low vision system performance. Applied research methods involved iterative CAD design, 3D printing for mechanical components, and performance analysis of AI algorithms. Mechanically, a fully modular base was developed, divided into three sections: a power supply with an internal 230V connection, a rotating section with integrated wire management, and a computing section based on a Raspberry Pi. The original military-like launching element was replaced by a visually friendly rehabilitation guide and a higher-capacity magazine, enhancing patient comfort. A key software result is the successful implementation of Edge AI technology. Transferring image analysis directly to the DepthAI Oak-D camera drastically offloaded the main processor. Comparative measurements demonstrated a performance increase from 13.4 to 28.1 frames per second. The main conclusions indicate that these architectural and software modifications eliminated thermal issues and significantly improved responsiveness. The proposed solution constitutes a reliable assistive tool that can effectively support automated motor training, relieving medical staff in their daily work.

Keywords: motor rehabilitation, edge computing, modular robotics, project evolution, 3D printing.

Paulina Wojdylak
Koło Naukowe Inżynierii Medycznej X-med

PRZEGLĄD SYSTEMÓW PRZECIWODLEŻYNOWYCH ORAZ KONCEPCJA WŁASNEGO ROZWIĄZANIA

Streszczenie

Odleżyny są jednym z najpoważniejszych powikłań u pacjentów długotrwale unieruchomionych, szczególnie gdy odżywianie i pielęgnacja są nieodpowiednie. W zależności od stopnia zaawansowania zmiany odleżynowe mogą przybierać różnorodne formy. W przypadkach zaawansowanych często dochodzi do martwicy i zakażeń. Profilaktyka przeciwodleżynowa znacznie obniża ryzyko ich powstawania oraz zmniejsza ich częstość występowania. Gojenie ran odleżynowych może trwać od kilku do nawet kilkunastu dni. Dobór właściwych metod pielęgnowania oraz miejscowego leczenia uzależniony jest od specyfiki rany. Produkty przeciwodleżynowe to nieodzowne akcesoria, które mogą ochronić osoby stale bądź czasowo unieruchomione, przed odleżynami. Należą do nich zarówno kosmetyki ochronne, jak i także krążki, poduszki czy materace przeciwodleżynowe. Celem pracy jest przedstawienie problematyki odleżyn, omówienie dostępnych rozwiązań profilaktycznych i leczniczych oraz zaprezentowanie koncepcji materaca niskobudżetowego z systemem alarmowym monitorującym długotrwały ucisk ciała pacjenta.

Słowa kluczowe: odleżyny, leczenie ran, pielęgnacja, sprzęt medyczny, materace przeciwodleżynowe.

1. Wprowadzenie

Odleżyny (łac. *decubitus*) powstają u chorych, którzy z jakiegoś powodu nie mogą samodzielnie zmieniać pozycji ciała. Są jednym z najpoważniejszych powikłań w opiece długoterminowej i wymagają szczególnej uwagi zarówno w warunkach domowych, jak i szpitalnych [1, 3]. Ich występowanie wiąże się z długotrwałym uciskiem tkanek w miejscach kostnych, ograniczoną perfuzją, tarciem i działaniem sił ścinających, co prowadzi do niedokrwienia i uszkodzenia skóry oraz tkanek podskórnych [5, 10].

Pacjenci długotrwale unieruchomieni, niemogący samodzielnie zmieniać pozycji, skazani są na pomoc osób z otoczenia. Gdy przebywają w domu, podlegają zazwyczaj opiece indywidualnej członków rodziny lub wynajętych opiekunów. W takim przypadku opieka nad pacjentem jest priorytetem dla opiekunów. Dzięki

temu mogą regularnie zmieniać ich pozycje ciała i zachowywać odpowiednie warunki zapobiegające powstawaniu odleżyn.

Pacjenci, których rodzina z różnych powodów nie może się zająć w domu są w innej sytuacji. Wówczas chory trafia do szpitala. Opieka indywidualna stwarza bardzo dobre warunki dla działań profilaktycznych, ale jest niemożliwa do zapewnienia w szpitalu. Wyposażenie oddziałów w sprzęt przeciwodleżynowy, często pozostawia wiele do życzenia, a na dodatek dyżurujące pielęgniarki obarczone są licznymi obowiązkami. Nawet jeżeli rodzina pacjenta zaopatrzy go w specjalistyczny materac, to personel w natłoku zajęć, często po prostu nie pamięta o regularnej zmianie pozycji chorego.

Powaga problemu wskazuje na konieczność opracowania sposobu skutecznego działania w kierunku zapobieganiu powstawania odleżyn u osób unieruchomionych. Szczególnie dotyczy to oddziałów szpitalnych, gdzie przebywa większa liczba pacjentów, a personel pielęgniarski jest niewystraczający. W odpowiedzi na te trudności powstał pomysł na zaprojektowanie materaca przeciwodleżynowego, który przypominałby pielęgniarkom o potrzebie zmiany pozycji pacjenta. Taki sprzęt powinien być w miarę tani, aby szpitale stać było na jego zakup w odpowiedniej ilości [13].

Regularna ocena ryzyka odleżyn, zastosowanie odpowiednich powierzchni przeciwodleżynowych oraz monitorowanie stanu skóry pacjentów stanowią podstawę skutecznej profilaktyki, zarówno w domu, jak i w szpitalu [9, 13].

2. Charakterystyka ran odleżynowych

2.1. Patomechanizm powstawania odleżyn

Najczęściej odleżyny powstają w tych miejscach ciała człowieka, w których kości usytuowane są blisko powierzchni skóry, a tkanki miękkie nie zapewniają wystarczającej amortyzacji. Są to przede wszystkim okolice łokci i nadgarstków, kości krzyżowej oraz kości guzicznej, bioder, krętarzy większych, łopatek, kręgosłupa, a także kostek i pięt. Jednak w zależności od pozycji w jakiej długotrwale ułożone jest ciało pacjenta odleżyny mogą rozwijać się również w miejscach nietypowych, takich jak. potylica lub okolice małżowiny usznej [3]. Na rysunku 1. przedstawiono schematyczną lokalizację najczęściej występujących odleżyn u pacjentów leżących na plecach.

Podstawowymi czynnikami prowadzącymi do powstawania odleżyn są ciśnienie proste, tarcie oraz boczne siły ścinające. Każda odleżyna powstaje zasadniczo jako efekt zamknięcia światła naczyń krwionośnych przez ciśnienie działające z zewnątrz oraz śródbłonkowych uszkodzeń w mikrokrążeniu spowodowanych siłami ścinającymi. Są to dwa procesy, które praktycznie zachodzą równocześnie, powodując cykl patofizjologicznych zmian skutkujących uszkodzeniem skóry i powstaniem rany odleżynowej [10].



Rys. 1. Lokalizacja miejsc powstawania odleżyn u pacjentów leżących na plecach (opracowanie własne na podstawie [3])

Fig. 1. Location of pressure ulcer formation in patients lying on their backs (adapted from [3])

Istotne znaczenie ma w tym procesie wielkość oraz czas trwania ciśnienia oddziałującego na skórę (długotrwałe działanie niskiego ciśnienia, jak i krótkotrwałe działanie wysokiego ciśnienia), a także jej wrażliwość i wrażliwość struktur niżej położonych, które poddawane są jego działaniu. Tarcie, które często dołącza się do tego procesu bezpośrednio uszkadza skórę, powoduje to również powtarzający się ucisk przekraczający średnie ciśnienie w naczyniach włosowatych, powodujące zniszczenie śródbłonka oraz zakrzepy płytkowe [10].

Kolejnym istotnym czynnikiem powodującym powstawanie odleżyn są siły ścinające, które oddziałują na skórę w płaszczyźnie równoległej i powodują jej przesuwanie względem struktur głębiej położonych odwrotnie jak ciśnienie działające prostopadle do jej powierzchni. Ponieważ te dwie siły występują razem, dopływ krwi do rozległych obszarów skóry zostaje zamknięty [5].

Szereg zewnętrznych i wewnętrznych czynników predysponuje do zmniejszenia indywidualnej tolerancji skóry na działające siły, co sprzyja powstawaniu odleżyn. Zestawienie tych czynników przedstawiono w tabeli 1. Jak podkreśla Boyko i wsp. [3], do najistotniejszych czynników ryzyka należą wszystkie stany prowadzące do unieruchomienia pacjenta, ograniczenia lub zniesienia czucia oraz niedożywienia, które w istotny sposób zwiększają podatność tkanek na uszkodzenia niedokrwienne.

Tabela 1. Czynniki predysponujące do powstania odleżyn

Table 1. Factors predisposing to the development of pressure ulcers

Czynniki wewnętrzne (specyficzne dla pacjenta)
→ ogólny stan organizmu, → stan odżywienia, → typ budowy (np. otyłość, wyniszczenie organizmu), → ruchomość ciała (np. niedowład), → nietrzymanie moczu i/lub stolca, → czynniki neurologiczne (brak czucia, bólu, ucisku), → czynniki naczyniowe (miażdżyca, niewydolność krążenia, POChP, cukrzyca),
Czynniki zewnętrzne (otoczenie pacjenta)
→ ciśnienie (wielkość i czas utrzymywania się), → tarcie i siły ściągające (głównie u chorych z wrażliwą skórą), → stan skóry (uszkodzenia przez nietrzymanie moczu i/lub kału).

Źródło: opracowanie własne.

Czynniki wewnętrzne są specyficzne dla indywidualnego chorego i upośledzają zdolność skóry do odpowiedzi na działające siły, wpływając na czucie, przepływy tkankowe i metabolizm skóry. Ogólny stan organizmu oraz odżywienie mają kluczowe znaczenie dla prawidłowego przebiegu procesów metabolicznych i regeneracyjnych. W przypadku złego odżywienia dochodzi do zahamowania procesów naprawczych tkanek, co powoduje zmniejszenie ich odporności na uszkodzenia.

U wyniszczonego pacjenta występuje zanik podskórnej tkanki tłuszczowej, co zwiększa ryzyko powstawania odleżyn, zwłaszcza w miejscach kostnych wyniosłości i ścięgien. Również osoby otyłe narażone są na występowanie odleżyn. Pomimo lepszej ochrony przed uciskiem, większa masa ciała sprzyja nasileniu sił ścinających oraz utrudnia prowadzenie skutecznej profilaktyki [9].

Pacjenci z ograniczoną ruchomością cierpią na brak spontanicznie pojawiającej się redystrybucji ciśnienia w miejscach ucisku, co naraża ich na powstawanie i rozwój odleżyn. Często współwystępuje u nich nietrzymanie moczu lub stolca, prowadzące do maceracji skóry. Powoduje to jej chemiczne i mechaniczne uszkodzenie oraz sprzyja rozwojowi infekcji, szczególnie w obecności tarcia i sił ścinających [9]. Szczególnie wysokie ryzyko rozwoju odleżyn dotyczy pacjentów z uszkodzeniem rdzenia kręgowego, u których współwystępują unieruchomienie oraz zaburzenia czucia, co potwierdzają obserwacje kliniczne opisane przez Boyko i wsp. [3].

Problemy neurologiczne i naczyniowe są powodem zaburzeń w skórze oraz mięśniach, wpływając na ich funkcjonowanie, a także perfuzję tkankową. Czynniki te w sposób istotny zwiększają podatność pacjentów na rozwój odleżyn.

2.2. Rodzaje i stopnie odleżyn

W zależności od stopnia zaawansowania zmiany, odleżyny mogą przybierać różnorodne formy. Początkowo mogą to być jedynie zaczerwienienie skóry oraz lekki obrzęk, przekształcające się czasem w pęcherze oraz nadżerki. Mogą one prowadzić do głębokich ubytków tkanki, które sięgają aż do mięśni, a w skrajnych przypadkach aż do kości. W stanach zaawansowanych często dochodzi do martwicy i zakażeń.

Dla precyzyjnej diagnozy, monitorowania postępu choroby oraz doboru terapii, Europejski Panel Doradczy ds. Odleżyn (EPUAP) oraz Narodowy Panel Doradczy ds. Odleżyn w USA (NPIAP) pracowały klasyfikacje określające stopień zaawansowania zmian [6]. W praktyce klinicznej stosuje się m.in. klasyfikacje EPUAP/NPIAP (tab. 2.), skalę Torrance'a z pięciostopniowym podziałem zmian skórnych (tab. 3.) oraz kolorową skalę ułatwiającą szybkie określenie stopnia zaawansowania i dominującego rodzaju tkanki w ranie (tab. 4.).

Tabela 2. Klasyfikacja odleżyn EPUAP/NPIAP

Table 2. EPUAP/NPIAP Classification of Pressure Ulcers

Stopnie zmian	Charakterystyka zmiany
I stopień	Niebledнący rumień – zaczerwienienie, które nie blednie, mogące sprawiać ból. Skóra może być cieplejsza lub chłodniejsza, twardsza lub bardziej miękka od sąsiadującej tkanki.
II stopień	Zmniejszona grubość – skóra właściwa staje się cieńsza, zaczerwienienie zmienia się w płytkie owrzodzenie z różowoczerwonym dnem lub pęcherze z płynem surowiczym/surowiczo-krwistym.
III stopień	Pełna utrata grubości skóry – odsłonięcie tkanki podskórnej, możliwa obecność martwicy wilgotnej, podminowanie i tunelowanie rany.
IV stopień	Pełna utrata grubości tkanki – odsłonięcie kości, mięśni lub ścięgien, ryzyko zapalenia szpiku/kości. Często podminowanie i tunelowanie, obecność martwicy lub strupów.

Źródło: opracowanie własne na podstawie [6, 8].

Tabela 3. Skala Torrance'a, która uwzględnia pięciostopniowy podział zmian skórnych

Table 3. Torrance Scale – Five-Stage Classification

Stopnie zmian	Charakterystyka zmiany
I stopień	Zaczerwienienie nad wyniosłością kostną, które może blednąć pod uciskiem. Skóra nie jest naruszona, ale może być cieplejsza/chłodniejsza, obrzęknięta, bolesna.
II stopień	Zaczerwienienie niebledнące (uszkodzenie mikrokrążenia), uszkodzenie naskórka i skóry właściwej. Mogą wystąpić pęcherze, nadżerki, owrzodzenia, często z bólem.
III stopień	Głębszy ubytek, brzegi obrzęknięte, dno rany z czerwoną ziarniną lub żółtymi masami tkanek. Czasem martwica.

Tabela 3 (cd.). Skala Torrance'a, która uwzględnia pięciostopniowy podział zmian skórnych

Table 3 (cont.). Torrance Scale – Five-Stage Classification

Stopnie zmian	Charakterystyka zmiany
IV stopień	Brzeg rany wyraźnie odgraniczony, dno z czarną martwicą, możliwe zakażenie i przetoki.
V stopień	Powstają jamy łączące się ze sobą, rana wypełniona czarno-brązowymi masami, martwica obejmuje mięśnie, czasem kości i stawy.

Źródło: opracowanie własne na podstawie [7].

Tabela 4. Kolorowa skala różnicująca rany odleżynowe

Table 4. Color-Based Classification of Pressure Ulcers

Kolor zmian	Charakterystyka zmiany
Czarny (martwica)	Rana pokryta suchą, twardą lub miękką martwicą, odleżyny są w charakterystycznym czarnym lub brązowym kolorze, mogą występować uchyłki i kieszenie.
Żółty (infekcja)	Nagromadzenie mas martwych komórek, głównie tłuszczowej tkanki podskórnej, duża ilość wysięku, źródło pożywki dla drobnoustrojów.
Czerwony (ziarnina)	Rozpoczyna się proces ziarninowania, tworzą się nowe naczynia włosowate, ziarnina jest bardzo delikatna, łatwo ulega zranieniu i krwawi. Czasami kolor przechodzi w ciemnoczerwony, brunatny lub czarny, co może świadczyć o zakażeniu.
Różowy (naskórkowanie)	Ostatnia faza procesu gojenia ran, dno rany wypełnione jest ziarniną, a na nią od brzegów rany do środka nachodzi nowo wytworzony naskórek.

Źródło: opracowanie własne na podstawie [11].

Powyższy system nie wskazuje na głębokość uszkodzenia tkanek, jednakże ułatwia określenie, jaki rodzaj tkanki dominuje obecnie w łożysku rany. Należy odróżnić te skale wizualnej oceny rany od skal oceny ryzyka wystąpienia odleżyn. Do najpopularniejszych skal tego rodzaju należą [8]:

1. Skala Norton - ocena od 1 do 4 pięć czynników: stan fizyczny, stan świadomości, aktywność, możliwość samodzielnego zmieniania pozycji, sprawność zwieraczy odbytu i cewki moczowej. Wynik poniżej 14 punktów wskazuje na ryzyko wystąpienia odleżyn.
2. Skala Waterflow - ocena wielu czynników ryzyka podzielonych na 10 kategorii: budowa ciała (0-3 pkt), stan skóry (0-3 pkt), płeć i wiek (1-5 pkt), trzymanie moczu i stolca (0-3 pkt), mobilność (0-5 pkt), apetyt (0-5 pkt), odżywienie komórkowe (1-8 pkt), choroby towarzyszące (4-6 pkt), urazy i przebyte zabiegi (5 pkt), przyjmowane leki (4 pkt za każdy). Wynik równy lub wyższy niż 10 pkt to stan alarmowy, równy lub wyższy niż 15 pkt stanowi zagrożenie, zaś wynik równy lub wyższy niż 20 pkt wskazuje na bardzo wysokie ryzyko odleżyn.
3. Skala Bradena - ocena 6 czynników ryzyka (percepcja, wilgotność, mobilność, aktywność, stan odżywienia, tarcie i siły ściskania) w 4-stopniowej skali.

Zagrożenie wystąpienia odleżyn wzrasta, gdy wynik pacjenta wynosi poniżej 16 punktów.

Ocena zaawansowania odleżyn jest kluczowa dla wdrożenia odpowiedniego leczenia [1]. Skale wizualne pozwalają określić stan rany, natomiast skale ryzyka umożliwiają ocenę stanu pacjenta, identyfikację czynników ryzyka i wdrożenie działań profilaktycznych.

2.3. Leczenie ran odleżynowych

Profilaktyka przeciwodleżynowa znacznie obniża ryzyko i zmniejsza częstość występowania ran odleżynowych, których gojenie może trwać od kilku, do nawet kilkunastu dni [2, 11]. Jest to jednak zawsze trudny oraz kosztowny proces, wymagający zaangażowania całego zespołu specjalistów (leczenie, pielęgnacja, rehabilitacja) [12]. Dobór właściwych metod pielęgnowania oraz miejscowego leczenia uzależniony jest od specyfiki rany [7].

W przypadku, gdy profilaktyka jest niewystarczająca, konieczne jest jak najszybsze działanie, ponieważ rany odleżynowe są wyjątkowo trudne do wyleczenia. Na rynku istnieje obecnie wiele specjalistycznych opatrunków, powstałych w celu leczenia odleżyn niezależnie od stopnia ich zaawansowania.

3. Współczesne systemy przeciwodleżynowe

3.1. Charakterystyka systemów przeciwodleżynowych

Produkty przeciwodleżynowe to nieodzowne akcesoria, które mogą ochronić osoby stale bądź czasowo unieruchomione, przed odleżynami. Należą do nich zarówno kosmetyki ochronne, jak i także krążki, poduszki czy materace przeciwodleżynowe [4].

Materace przeciwodleżynowe są nieocenionym elementem profilaktyki i leczenia odleżyn. Dzięki nowoczesnym technologiom pozwalają zachować integralność skóry, poprawiają komfort pacjentów i znacząco zmniejszają ryzyko wystąpienia poważnych powikłań. Obecnie ich stosowanie jest standardem w opiece długoterminowej, zarówno w warunkach szpitalnych, jak i domowych [13].

Materace przeciwodleżynowe to specjalistyczne wyroby medyczne, których zadaniem jest zmniejszanie nacisku punktowego na ciało pacjenta oraz poprawa cyrkulacji powietrza i krwi w miejscach kontaktu ze skórą. Działają one poprzez naprzemienne pompowanie i spuszczanie powietrza z komór (w materacach pneumatycznych), równomierne rozłożenie ciężaru ciała na powierzchni (w materacach piankowych viscoelastic lub żelowych), zapewnienie mikroprzepływu powietrza, który zmniejsza potliwość i wilgoć skóry. Rodzaje materacy przeciwodleżynowych obejmują materace [6, 13]:

1. Statyczne (piankowe i żelowe) - wykonane z pianki poliuretanowej lub viscoelastic (tzw. „pianka z pamięcią”), działają poprzez równomierne rozłożenie

masy ciała i zmniejszenie punktowego nacisku. Są zalecane osobom z niskim ryzykiem powstania odleżyn.

2. Dynamiczne (pneumatyczne) - składają się z wielokomorowych poduszek lub materacy napompowanych powietrzem, w których ciśnienie jest cyklicznie zmieniane dzięki kompresorowi. Zmiana punktów ucisku co kilka minut poprawia ukrwienie tkanek, co wspomaga profilaktykę i leczenie odleżyn u osób z wysokim ryzykiem lub już istniejącymi zmianami skórnymi.
3. Specjalistyczne – obejmują zaawansowane rozwiązania:
   - materace z funkcją mikrowentylacji lub chłodzenia,
   - systemy hybrydowe, łączące zalety materacy piankowych i dynamicznych.

Tabela 5. Wybrane modele materacy przeciwodleżynowych

Table 5. Selected Pressure Ulcer Mattress Models

Materac	Typ	Materiał	Kluczowe cechy
SYSTAM 	Statyczny (piankowy)	Pianka poliuretanowa, viscoelastic	Gofrowana powierzchnia, wentylacja, różna gęstość bloków, wspomaga profilaktykę i leczenie I–II stopnia odleżyn, kompatybilny z łóżkami rehabilitacyjnymi
AirMATE Timago 	Dynamiczny (pneumatyczny)	Wielokomorowy PVC	Naprzemienne pompowanie, zmiana punktów nacisku, wspomaga krążenie i komfort pacjenta
TGR-Y MR 001-5 	Dynamiczny (pneumatyczny)	PVC antyalergiczne	Cykl zmiany ciśnienia, wyjmowane komory, łatwa dezynfekcja, poprawia ukrwienie tkanek
Gofer 	Statyczny (piankowy)	Pianka PU T25	Sprężysta gąbka, pokrowiec paroprzepuszczalny, pełne podparcie kręgosłupa, nie zaburza funkcji oddechowej i krążeniowej
ASX DIGITAL Novacare 	Dynamiczny (pneumatyczny)	Nylon z powłoką poliuretanową	18 komór, cyfrowa kontrola ciśnienia, tryb statyczny i opieki, sensor ciśnienia, do leczenia odleżyn I–IV stopnia

Źródło: opracowanie własne.

Cierpienie, ból, ryzyko infekcji, a także wysokie koszty związane z zakupem środków opatrunkowych, materaców przeciwodleżynowych i rehabilitacją diametralnie wpływają na jakość życia pacjentów z ranami odleżynowymi [2].

3.2. Dostępność systemów przeciwodleżynowych w polskich szpitalach

W polskich szpitalach pacjenci, którzy są unieruchomieni przez dłuższy czas i rokują z jakichś powodów powstanie odleżyn, mają zazwyczaj dostęp do materacy przeciwodleżynowych lub innych powierzchni odciążających. Szpitale stosują materace przeciwodleżynowe (zarówno statyczne jak i dynamiczne/zmiennociśnieniowe) pozyskując je na przykład w ramach programów profilaktycznych lub kupując na wyposażenie.

W polskich standardach akredytacyjnych dla szpitali profilaktyka i zapobieganie odleżynom są ujęte jako część opieki pacjenta, choć konkretne wymagania dotyczące sprzętu mogą nie być zapisane wprost w szczegółach. W zaleceniach medycznych przy profilaktyce odleżyn wymienia się stosowanie poduszek i materacy przeciwodleżynowych jako jedną z metod zmniejszania ryzyka powstawania tych zmian skórnych. Jednakże polska szpitalna rzeczywistość nie zawsze jest tak optymistyczna. Dostępność materacy może się różnić pomiędzy placówkami zwłaszcza w mniejszych miejscowościach, lub jeżeli szpital jest słabiej wyposażony. NFZ nie zawsze zapewniaj pełne pokrycie kosztów materacy przeciwodleżynowych dla wszystkich pacjentów.

Ponadto należy zauważyć, że materace szpitalne, jak i pozostały sprzęt, są produktami wielokrotnego użytku przez długi okres czasu. Braki finansowe polskiej służby zdrowia nie pozwalają na zakup tego sprzętu indywidualnie dla każdego pacjenta, jest on po prostu przekazywany kolejnym tak długo, dopóki się nie zużyje. Dlatego rodziny chorych często na własną rękę organizują zakup materaca i dostarczają go do szpitala. Mając skierowanie lub ordynację od lekarza mogą oni uzyskać refundację od Narodowego Funduszu Zdrowia lub ubiegać się o wsparcie finansowe lub dofinansowanie do sprzętu rehabilitacyjnego w ramach pomocy społecznej lub Funduszu Osób Niepełnosprawnych. Pomoc taką organizują też hospicja i opieka domowa.

4. Projekt koncepcyjny niskobudżetowego materaca przeciwodleżynowego

4.1. Uzasadnienie potrzeby opracowania rozwiązania

Problem finansowania zakupu materacy przeciwodleżynowych dla jednostek medycznych, w których przebywają chorzy długoterminowo unieruchomieni, nie jest jedynym wyzwaniem związanym z profilaktyką odleżyn. Innym, równie istotnym czynnikiem jest niewystarczająca liczba personelu pielęgniarskiego. Pielęgniarki dyżurujące, przy natłoku obowiązków, nie zawsze są w stanie zapewnić

regularną zmianę pozycji wszystkim pacjentom zagrożonym wystąpieniem ran odleżynowch.

W związku z tym zasadne wydaje się opracowanie rozwiązania wspomagającego personel medyczny poprzez monitorowanie czasu przebywania pacjenta w jednej pozycji. Dobrym rozwiązaniem byłby materac przeciwodleżynowy wyposażony w system powiadamiania o zbyt długim przebywaniu pacjenta w jednej pozycji. Mógłby on stanowić dodatkowe narzędzie wspierające działania profilaktyczne.

4.2. Założenia projektowe

Projekt zakłada stworzenie materaca, który:

- wykrywa nacisk ciała pacjenta,
- mierzy czas przebywania w jednej pozycji,
- sygnalizuje alarmem zbyt długie pozostawanie pacjenta w niezmienionej pozycji.

System nie będzie bezpośrednio ingerował w ułożenie pacjenta. Jego zadaniem będzie wyłączeni informowanie opiekuna o potencjalnym zagrożeniu wynikającym z długotrwałego ucisku tkanek.

4.3. Proponowana budowa i zasada działania

Projektowany materac przeciwodleżynowy wyposażony byłby w układ czujników nacisku oraz prosty system elektroniczny, którego zadaniem byłoby monitorowanie czasu przebywania pacjenta w jednej pozycji leżącej. Działanie systemu rozpoczynałoby się w momencie ułożenia pacjenta na materacu i włączenia zasilania urządzenia.

W przedstawionej koncepcji klasyczny materac piankowy stanowi bazę konstrukcyjną, wewnątrz której planuje się umieszczenie trzech lub czterech czujników nacisku. Czujniki te byłyby rozmieszczone w newralgicznych strefach ciała, takich jak okolice pleców, kości krzyżowej oraz ud lub pięt. Poszczególne elementy projektowanego materaca oraz ich funkcje przedstawiono w tabeli 6.

Sygnały analogowe z czujników byłyby cyklicznie odczytywane przez mikrokontroler (np. Arduino lub ESP32), którego przetwarzałby dane oraz monitorowałby czas przebywania pacjenta w jednej pozycji. Na podstawie odczytów z czujników mikrokontroler określałby aktualny rozkład nacisku, traktowany jako pozycja odniesienia pacjenta. Następnie uruchamiany byłby licznik czasu, który mierzy okres przebywania pacjenta w tej samej pozycji. W trakcie pracy systemu kolejne pomiary byłyby porównywane z wartościami zapisanymi wcześniej.

W przypadku wykrycia istotnej zmiany rozkładu nacisku, świadczącej o zmianie pozycji pacjenta, licznik czasu zostałby wyzerowany. W przypadku braku zmian w rozkładzie nacisku przez czas dłuższy niż ustalony limit (np. 120 minut), mikrokontroler generowałby sygnał alarmowy.

Tabela 6. Poszczególne elementy projektowanego materaca

Table 6. Individual Components of the Designed Mattress and Their Functions

Element	Opis	Funkcja
Czujnik nacisku	Reaguje na zmianę obciążenia mechanicznego poprzez zmianę rezystancji.	Wykrywa obecność pacjenta oraz monitoruje nacisk wywierany przez ciało w poszczególnych strefach materaca.
Mikrokontroler	Programowalny mikrokontroler (np. Arduino Uno lub kompatybilny), wyposażony w wejścia analogowe i cyfrowe.	Odpowiada za odczyt danych z czujników, analizę rozkładu nacisku, pomiar czasu oraz sterowanie elementami sygnalizacyjnymi.
Dioda LED	Dioda elektroluminescencyjna o niskim poborze mocy.	Wizualna sygnalizacja stanu pracy systemu (normalna praca lub stan alarmowy).
Brzęczyk	Element elektroakustyczny generujący sygnał dźwiękowy po podaniu napięcia.	Akustycznie powiadamia opiekuna o zbyt długim przebywaniu pacjenta w jednej pozycji.
Zasilacz	Zasilacz sieciowy lub akumulator dostarczający napięcie robocze układu.	Zasila mikrokontroler oraz pozostałe elementy systemu.
Materac piankowy	Standardowy materac piankowy przystosowany do integracji z czujnikami.	Baza konstrukcyjna systemu, zapewniająca komfort pacjenta.

Źródło: opracowanie własne.

Sygnał alarmowy mógłby być realizowany poprzez uruchomienie diody sygnalizacyjnej LED (zielona – prawidłowa praca systemu, czerwona – alarm) oraz brzęczyka akustycznego, informujących opiekuna o konieczności zmiany pozycji pacjenta. Po ręcznym wyłączeniu alarmu lub po wykryciu zmiany pozycji pacjenta system wracałby do normalnego trybu pracy. Zasilanie planowane jest poprzez powerbank lub zasilacz o napięciu 5–12 V.

Zaprojektowany system nie ingerowałby bezpośrednio w pozycję pacjenta, a jego zadaniem byłoby wyłącznie informowanie o potencjalnym zagrożeniu powstania odleżyn, wynikającym z długotrwałego ucisku tkanek.

4.4. Szacunkowa analiza kosztów

Kwestią zasadniczą jest koszt zakupu powyższego materaca. Koszt realizacji projektu materaca przeciwodleżynowego z funkcją sygnalizacji oszacowano na podstawie cen detalicznych dostępnych na rynku krajowym. Analiza kosztów obejmuje wyłącznie wykonanie prototypu i nie uwzględnia kosztów związanych z certyfikacją medyczną, produkcją seryjną ani pracami badawczo-rozwojowymi. Koszty poszczególnych elementów materaca zostały zestawione w tabeli 7.

Tabela 7. Przybliżony koszt poszczególnych elementów materaca

Table 7. Approximate Cost of Individual Components of the Mattress

Element	Ilość	Cena jednostkowa [zł]	Koszt [zł]
Czujnik nacisku FSR	4	25–35	100-140
Mikrokontroler (Arduino Uno/ klon)	1	40-70	40-70
Dioda LED + rezystor	2	1-2	2-4
Brzęczyk	1	5-10	5-10
Zasilacz 5 V / powerbank	1	30-50	30-15
Przewody, płytka stykowa	1	15-25	15-25
Materac piankowy	1	150-300	150-300
Obudowa elektroniki	1	20-40	20-40

Źródło: opracowanie własne.

Całkowity koszt wykonania prototypu waha się w granicach od około 360 zł do 640 zł, w zależności od jakości zastosowanych komponentów oraz źródła ich zakupu. Średni koszt realizacji projektu można przyjąć na poziomie około 500 zł, co czyni zaproponowane rozwiązanie ekonomicznie uzasadnionym w kontekście projektu.

Dla porównania, komercyjnie dostępne materace przeciwodleżynowe wyposażone w zaawansowane systemy monitorowania lub aktywnej zmiany nacisku osiągają ceny rzędu kilku tysięcy złotych. W związku z tym zaproponowany projekt, mimo swojej uproszczonej konstrukcji, pozwala na znaczące obniżenie kosztów przy jednoczesnym zachowaniu podstawowej funkcjonalności związanej z profilaktyką odleżyn. Ponadto koszt materaca można obniżyć poprzez użycie 3 czujników zamiast 4 (-30 zł), zastosowanie Arduino Nano zamiast Uno (-20 zł) oraz użycie używanego materaca piankowego (-100 zł).

Literatura

[1] Bąk O.: Odleżyny – rodzaje i stopnie odleżyn oraz ich klasyfikacja, [online] https://www.doz.pl/czytelnia/a18258-Odlezyny__rodzaje_i_stopnie_odlezyn_oraz_ich_klasyfikacja#4 (dostęp: 27 grudnia 2025 r.).

[2] Bazaliński D., Wołkowicz M., Więch P., Kózka M.: Leczenie piorunującej rany odleżynowej z wykorzystaniem rekomendacji dotyczących profilaktyki przeciwodleżynowej – opis przypadku, Leczenie Ran 2017; 14/2, s. 65-69.

[3] Boyko T. V., Longaker M. T., Yang G. P.: Review of the Current Management of Pressure Ulcers, Advances in Wound Care, 7(2), 2018, s. 57–67, https://doi.org/10.1089/wound.2016.0697.

[4] Cox J: Pressure injury prevention: clinical practice guidelines. In: Kottner J., Gefen A. (eds). Pressure Ulcer Prevention and Treatment. Springer, 2018.

[5] Czeczelewska E.: Etiologia i charakterystyka ran odleżynowych, [online] https://nursing.com.pl/artykul/etiologia-i-charakterystyka-ran-odlezynowych-61a65c7e4170277e5ab42645, (data publikacji: 30 listopada 2021 r.).

[6] European Pressure Ulcer Advisory Panel (EPUAP), National Pressure Injury Advisory Panel (NPIAP), Pan Pacific Pressure Injury Alliance (PPPIA): Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline, 2019.

[7] Gluz-Golik U.: Skala Torrance'a – co to jest i do czego służy?, [online] https://karierabohatera.pl/artykuly/skala-torrance#odlezyny, (data publikacji: 25 maja 2022 r.).

[8] Klasyfikacja odleżyn – poznaj stopnie zaawansowania i fazy gojenia, [online] https://winncare.pl/klasyfikacja-odlezyn-poznaj-stopnie-zaawansowania-i-fazy-gojenia/ (data publikacji: 9 września 2020 r.).

[9] Klukowski M.: Odleżyny – zapobieganie i leczenie, Ogólnopolski Przegląd Medyczny, 2009; (9/10), s. 36-38.

[10] Rewerska B., Sianos G., Kowalik J., Kopański Z., Skura-Madziała A., Ptak W.: Patofizjologia, profilaktyka i leczenie odleżyn, Journal of Public Health, Nursing and Medical Rescue, nr 2/2011, s. 14-19.

[11] Szewczyk M. T., Cwajda-Białasik J., Mościcka P., Cierzniakowska K., Bazaliński D., Jawień A., Spannbauer A., Polak A., Sopata M., Kozłowska E., Popow A., Jabłońska R., Przybek-Mita J., Sierżantowicz R., Kuberka I., Ślusarz R., Mrozikiewicz-Rakowska B., Kózka M., Jakubowska U., Bakowska M., Augusewicz Z., Samson I.: Leczenie odleżyn – zalecenia Polskiego Towarzystwa Leczenia Ran. Część II, Leczenie Ran 2020; 17(4), s. 151-184.

[12] Szewczyk M. T., Jawień A. (red.): Leczenie ran przewlekłych, Wydawnictwo Lekarskie PZWL, Warszawa 2019.

[13] Thomas DR: Prevention and treatment of pressure ulcers: what works? J Am Med Dir Assoc. 2006;7(1):46–59.

REVIEW OF PRESSURE ULCER PREVENTION SYSTEMS AND CONCEPT OF OWN SOLUTION

Summary

Pressure ulcers are among the most serious complications in long-term immobilized patients, especially when nutrition and care are inadequate. Depending on the severity, pressure ulcers may vary in appearance, from mild erythema and swelling to advanced cases involving tissue necrosis and infection. Preventive measures significantly reduce the risk and frequency of pressure ulcer development. Healing of pressure ulcers can last from several to over a dozen days, and the choice of proper care and local treatment depends on the wound characteristics. Pressure-relieving products are essential tools to protect patients from developing pressure ulcers and include protective creams, cushions, and specialized mattresses. This paper aims to present the issue of pressure ulcers, review currently available preventive and therapeutic solutions, and propose a low-cost mattress concept with an alarm system for monitoring prolonged pressure on the patient's body.

Keywords: pressure ulcers, wound care, nursing, medical equipment, pressure-relieving mattresses.

Paweł Dźwięga
Koło Naukowe Inżynierii Medycznej X-med

PROJEKT KONSTRUKCYJNY MECHATRONICZNEJ DŁONI

Streszczenie

Celem projektu było zaprojektowanie oraz wykonanie projektu konstrukcyjnego mechatronicznej dłoni przy zastosowaniu technologii druku 3D, oraz ocena możliwości odwzorowania podstawowych funkcji ruchowych ludzkiej dłoni. Założono uzyskanie konstrukcji umożliwiającej zginanie palców oraz zachowanie budowy segmentowej odpowiadającej anatomii. W części badawczej zastosowano metodę projektowo-eksperymentalną. Opracowano model trójwymiarowy w środowisku CAD w programie Autodesk Fusion, gdzie zaprojektowano przeguby i otwory montażowe. Następnie wykonano fizyczny prototyp metodą przyrostową wykorzystując drukarkę Creality CR-6 SE z użyciem filamentu PLA. Po wydruku przeprowadzono montaż mechaniczny, łącząc paliczki za pomocą osi wykonanych za pomocą filamentu oraz stosując żyłki pełniące funkcję ścięgien jak i gumek umożliwiających powrót do pozycji początkowej. Działanie modelu przetestowano poprzez sprawdzenie zakresu i płynności ruchu palców. W wyniku realizacji projektu uzyskano sprawnie działający model mechatronicznej dłoni, który umożliwia niezależne zginanie poszczególnych palców. Konstrukcja zachowała stabilność mechaniczną oraz poprawne dopasowanie elementów bez konieczności dodatkowej obróbki. Wnioski wskazują, że technologia druku 3D stanowi efektywne i niskokosztowe narzędzie do tworzenia prototypów urządzeń mechatronicznych. Opracowany model stanowi solidną podstawę do dalszych prac rozwojowych obejmujących integrację napędu elektrycznego oraz zastosowanie bardziej wytrzymałych materiałów konstrukcyjnych.

Słowa kluczowe: mechatronika, druk 3D, sztuczna dłoń, CAD, modelowanie 3D, proteza.

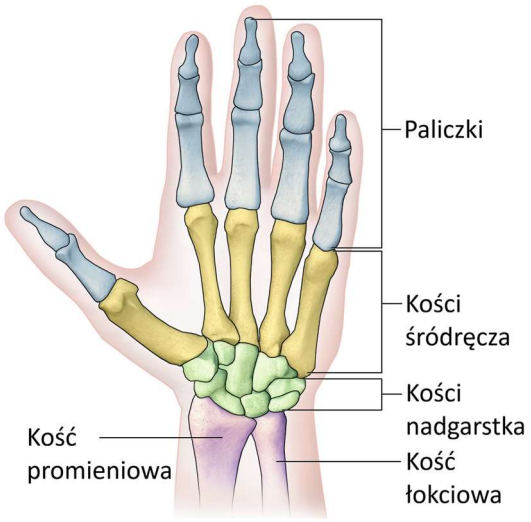
1. Wstęp

Od wieków ludzie poszukują metod zastępowania brakujących kończyn przy pomocy protez, co umożliwia częściowe lub całkowite przywrócenie sprawności manualnej oraz poprawę jakości życia. Początkowo były to konstrukcje mechaniczne, które były proste, natomiast dynamiczny rozwój różnych technologii

w ostatnich dekadach znacząco zwiększył możliwości funkcjonalne protez. Szczególnie ważną rolę odegrał postęp w dziedzinie mechatroniki, która to łączy w sobie różnorodne elementy mechaniki, elektroniki czy informatyki, co umożliwia tworzenie zaawansowanych oraz funkcjonalnych urządzeń [6].

Równolegle z rozwojem mechatroniki do powszechnego użytku weszła technologia przyrostowa (ang. additive manufacturing, AM), znana także jako druk 3D, która polega na warstwowym wytwarzaniu obiektów na podstawie modeli cyfrowych. Umożliwia ona szybkie i stosunkowo tanie wytwarzanie złożonych elementów konstrukcyjnych, często niemożliwych do wykonania tradycyjnymi metodami, przy jednoczesnym minimalizowaniu odpadów materiałowych [10]. Integracja tych dwóch dziedzin stworzyła nowe możliwości projektowania nowoczesnych protez kończyn, które są coraz bliższe do perfekcyjnego odwzorowania naturalnej budowy i funkcjonalności ludzkiej ręki.

Ludzka ręka stanowi jeden z najbardziej złożonych narządów ruchu, charakteryzujący się skomplikowaną budową anatomiczną oraz wysokim stopniem funkcjonalności. Jej struktura obejmuje układ kostny złożony z kości śródręcza oraz paliczków bliższych, środkowych i dalszych, połączonych ruchomymi stawami umożliwiającymi precyzyjne ruchy zginania i prostowania. Istotną rolę odgrywa również układ ścięgien i mięśni, który pozwala na przenoszenie sił oraz koordynację ruchów palców. Odwzorowanie naturalnej budowy i funkcjonalności ludzkiej ręki w konstrukcjach mechatronicznych wymaga uwzględnienia tej segmentacji oraz zasad działania przegubów, co stanowi jedno z kluczowych wyzwań w projektowaniu nowoczesnych protez kończyn górnych (rys. 1.) [1].



Rys. 1. Kości dłoni i nadgarstka [3]

Fig. 1. Hand and wrist bones [3]

Obecne protezy myoelektryczne oferują różne liczby stopni swobody, mechanizmy przegubów i możliwości aktuacji, co pozwala na odtwarzanie naturalnych ruchów dłoni. Projektowane modele 3D uwzględniają podział na palce i śródręcze oraz mechanizmy sprzężenia stawów, co zwiększa funkcjonalność przy ograniczonej liczbie napędów [2]. Mechatroniczna dłoń jest przykładem rozwiązania, które niedawno było traktowane jako wizja przyszłości, a obecnie znajduje się w zasięgu ręki.

Celem niniejszego projektu było zaprojektowanie oraz wykonanie prototypu mechatronicznej dłoni wykorzystując technologie druku 3D, a także uzyskanie konstrukcji możliwie jak najbardziej zbliżonej do organicznej kończyny pod względem budowy i zakresu ruchu.

2. Przegląd dostępnych rozwiązań

Rozwój technologii druku 3D, w szczególności drukarek desktopowych, znacząco wpłynął na rynek protetyczno-ortetyczny. Technologie przyrostowe umożliwiają szybkie i ekonomiczne wytwarzanie elementów o skomplikowanej geometrii. Modyfikacja kształtu wyprodukowanych komponentów nie powoduje istotnego wzrostu kosztów prototypów, co jest szczególnie ważne przy projektowaniu różnego rodzaju urządzeń. Ma to znaczenie zwłaszcza w kontekście konstrukcji odwzorowujących funkcje ruchowe ludzkiej dłoni, takich jak mechaniczne protezy tej części ciała [8].

Na rynku pojawiło się wiele projektów typu DIY (do it yourself), które mogą być wykonywane w warunkach domowych oraz dostosowywane do indywidualnych wymiarów kończyny pacjenta. Wśród najbardziej rozpoznawalnych rozwiązań tego typu można wyróżnić projekty Robohand [5], Cyborg Beast [11] oraz Raptor Hand [7].

Robohand, będąca rozwinięciem projektu Robofinger, jest to mechaniczna proteza dłoni, w której zginanie palców realizowane jest poprzez ruch nadgarstka. Zastosowane cięgna napinają się podczas jego zgięcia, powodując zamknięcie dłoni i umożliwiając chwyt przedmiotów. Warunkiem poprawnego działania protezy jest zachowana ruchomość nadgarstka na poziomie co najmniej 30°. Konstrukcja ta przeznaczona była głównie dla dzieci w wieku przedszkolnym i wczesnoszkolnym [8, 5].

Kolejnym popularnym rozwiązaniem typu open source jest Cyborg Beast. W porównaniu do Robohand charakteryzuje się bardziej zwartą i estetyczną konstrukcją oraz niemal całkowitym wykorzystaniem elementów drukowanych w technologii 3D. W projekcie zastosowano dwuliniowy system cięgien, składający się z elementów sztywnych i elastycznych, co pozwala na aktywne zginanie palców oraz ich pasywne prostowanie. Rozwój projektu doprowadził do powstania wielu wariantów konstrukcyjnych, różniących się m.in. budową kciuka oraz stopniem ruchomości nadgarstka. Proteza Cyborg Beast, dzięki kompatybilności

z aplikacją Handomatic, umożliwia parametryczne dopasowanie geometrii poszczególnych elementów do indywidualnych potrzeb użytkownika [8, 11].

Raptor Hand stanowi rozwinięcie wcześniejszych koncepcji i łączy ich najważniejsze zalety. Projekt został zoptymalizowany pod kątem procesu drukowania oraz montażu. Zredukowano liczbę elementów wymagających podpór, uproszczono konstrukcję oraz zastąpiono śruby drukowanymi kołkami. W efekcie niemal wszystkie elementy konstrukcyjne, z wyjątkiem cięgien, mogą zostać wykonane w technologii druku 3D, co skraca czas produkcji i zmniejsza koszty [8, 7].

Zarówno Cyborg Beast, jak i Raptor Hand, pokazują potencjał technologii druku 3D w tworzeniu funkcjonalnych, indywidualnie dopasowanych prototypów dłoni. W przypadku obu tych projektów możliwe jest parametryczne dopasowanie protezy do warunków anatomicznych użytkownika z wykorzystaniem dedykowanego oprogramowania. Pozwala to na modyfikację geometrii poszczególnych elementów oraz dostosowanie konstrukcji do indywidualnych potrzeb pacjenta [8]. W kontekście rozwijanego projektu mechatronicznej dłoni, doświadczenia te potwierdzają, że druk 3D pozwala na wystarczająco dokładne odwzorowanie anatomicznych cech dłoni, umożliwia testowanie zakresu ruchu poszczególnych palców oraz przygotowanie konstrukcji modułowej, która może być dalej rozwijana w kierunku pełnej mechatroniki.

3. Założenia projektowe

Głównymi założeniami projektowymi jest opracowanie mechatronicznej dłoni, która będzie jak najbardziej odzwierciedlała anatomię ludzkiej dłoni. Szczególną uwagę zwrócono na budowę palców, które miały posiadać możliwość zginania w stawach, naśladując naturalny ruch paliczków.

Do podstawowych założeń należały:

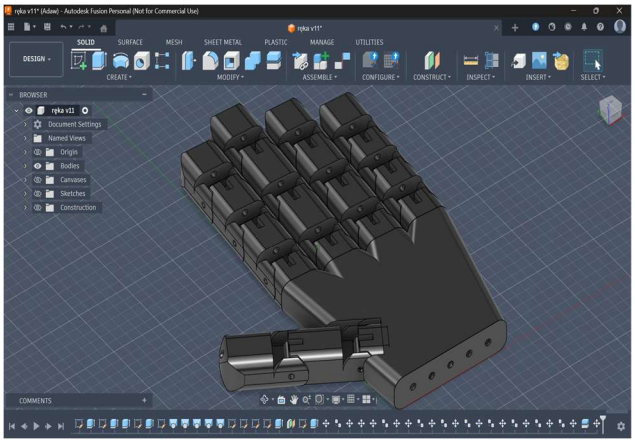
1. Odwzorowanie proporcji i segmentacji palców (paliczki dalsze, środkowe oraz bliższe).
2. Zapewnienie ruchomości poszczególnych segmentów.
3. Umożliwienie niezależnego zginania palców.
4. Zastosowanie technologii druku 3D jako głównej metody wytwarzania elementów konstrukcyjnych.
5. Otrzymanie modelu z niskim nakładem kosztów.

Projekt miał charakter konstrukcyjny, dlatego priorytetem była funkcjonalność mechanizmu, a nie pełna integracja z układami elektronicznymi.

4. Opracowanie modelu 3D

Trójwymiarowy model mechatronicznej dłoni opracowano w środowisku CAD Autodesk Fusion, umożliwiającym przestrzenne modelowanie elementów mechanicznych oraz analizę ich wzajemnego dopasowania (rys. 2.) [9]. Model

został następnie wyeksportowany w formacie STL, umożliwiając jego dalsze przetworzenie w oprogramowaniu do druku 3D i przygotowanie zgodnie z workflow technologii FDM.



Rys. 2. Model dłoni w Autodesk Fusion

Fig. 2. Hand model in Autodesk Fusion

Projekt obejmował:

1. segmentację dłoni na śródręcze i palce, zgodnie z anatomią ludzkiej ręki,
2. konstrukcję przegubów umożliwiających ruch zginania paliczków,
3. przygotowanie otworów montażowych do prowadzenia żyłek pełniących funkcję sztucznych ścięgien oraz gumek działających jako prostowniki,
4. zaplanowanie odpowiednich luzów konstrukcyjnych umożliwiających swobodny ruch elementów po wydruku.

Model został zaprojektowany w sposób modułowy, co umożliwia łatwą modyfikację poszczególnych części w przyszłości, np. w celu dodania napędu elektrycznego.

5. Wydruk 3D i montaż

Opracowany model wydrukowano w technologii osadzania topionego materiału (ang. fused deposition modeling, FDM) przy użyciu drukarki 3D Creality CR-6 SE. W tej technologii filament termoplastyczny jest podgrzewany i nanoszony warstwa po warstwie, aż do uzyskania pełnego kształtu obiektu. FDM jest jedną z najpopularniejszych metod druku 3D ze względu na prostotę, niskie koszty materiałów oraz możliwość wykonywania części o złożonej geometrii [10]. Poszczególne elementy protezy wykonano z polilaktydu (PLA), który należy do materiałów termoplastycznych. Materiał ten wybrano ze względu na łatwość obróbki, niski koszt oraz odpowiednią wytrzymałość mechaniczną do zastosowań prototypowych.

Proces druku przebiegał zgodnie z typowym workflow FDM [10], obejmującym następujące etapy:

1. Slicing albo inaczej plasterkowanie - import modelu do specjalnego oprogramowania, podział modelu na warstwy i wygenerowanie kodu (G-code), określającego ścieżki ruchu głowicy drukującej, prędkość i parametry ekstrudera.
2. Drukowanie - nakładanie podgrzanego filamentu PLA warstwa po warstwie.
3. Post-processing - usunięcie ewentualnych podpór i przygotowanie powierzchni elementów do montażu.

Analiza jakości wydruku nie wykazała istotnych odkształceń ani błędów wpływających na późniejszy montaż. Wszystkie elementy pasowały do siebie i nie wymagały dodatkowej obróbki mechanicznej.

Proces montażu rozpoczęto od połączenia poszczególnych paliczków w obrębie każdego palca. Segmenty połączono za pomocą fragmentów filamentu, które pełnią funkcję osi przegubów. Końce filamentu zostały zgrzane, w celu utworzenia trwałego połączenia zapobiegającego ich wysunięciu.

Następnie przez odpowiednio przygotowane otwory przeprowadzono żyłki, które pełnią funkcję sztucznych ścięgien. Ich napięcie umożliwia zginanie palców. Zostały również przymocowane gumki, które pełnią funkcję prostowników.

Wszystkie wydrukowane elementy pasowały do siebie i nie wymagały dodatkowej obróbki mechanicznej. Proces montażu przebiegł zgodnie z założeniami projektowymi i nie napotkano istotnych problemów technicznych.

6. Wynik

W wyniku przeprowadzonych prac projektowych opracowano funkcjonalny model mechatronicznej dłoni, wykonany w technologii FDM z filamentu PLA (rys. 3.). Konstrukcja umożliwia niezależne zginanie poszczególnych palców, co świadczy o skutecznej realizacji założeń projektowych.



Rys. 3. Finalny rezultat
Fig. 3. Final result

Model poprawnie odwzorowuje podstawowe działanie ludzkiej dłoni w zakresie ruchu zginania. Uzyskane rezultaty potwierdzają poprawność przyjętych rozwiązań konstrukcyjnych oraz możliwość dalszego rozwoju projektu.

7. Kierunki dalszego rozwoju projektu

W przyszłości, w celu udoskonalenia projektu, możliwe jest zastosowanie miniaturowych silników elektrycznych zintegrowanych w obrębie dłoni, podobnie jak zaprezentowano w pracy Mohamaddiego i in. [4].

Możliwe jest także zastosowanie innych materiałów w druku 3D, aby uzyskać bardziej wytrzymałą konstrukcję, a dzięki modularnej budowie elementów FDM łatwo wprowadzać modyfikacje geometryczne lub rozmiarowe poszczególnych części.

8. Wnioski

Zastosowanie technologii przyrostowych pozwoliło na niskokosztowe opracowanie prototypu mechatronicznej dłoni. Druk 3D pozwolił na precyzyjne odwzorowanie zaprojektowanych elementów oraz łatwą modyfikację konstrukcji. Opracowany model pozwala na niezależne zginanie palców, co potwierdza zasadność przyjętych rozwiązań konstrukcyjnych. Przyjęta konstrukcja spełnia założone cele projektowe. Przedstawiony projekt stanowi solidną podstawę do dalszych prac rozwojowych, w tym integracji napędu elektrycznego oraz zastosowania bardziej wytrzymałych materiałów. Projekt potwierdza, że połączenie mechatroniki oraz technologii druku 3D stwarza realne możliwości tworzenia funkcjonalnych protez kończyn w warunkach edukacyjnych.

Literatura

[1] Bednarczyk K., Mikołajczyk T.: Model dłoni sterowany mięśniami pneumatycznymi, w: CAX'2010: Komputerowe wspomaganie nauki i techniki. VII warsztaty naukowe, Wydaw. Uczelniane Uniwersytetu Technologiczno-Przyrodniczego, Bydgoszcz 2011, s. 119–124, dostęp online: https://dlibra.pbs.edu.pl/dlibra/publication/326/edition/331/content (dostęp: 23 lutego 2026 r.).

[2] Belter J.T., Segil J.L., Dollar A.M., Weir R.F.: Mechanical design and performance specifications of anthropomorphic prosthetic hands: A review, Journal of Rehabilitation Research and Development, vol. 50, no. 5, 2013, pp. 599–618, DOI: 10.1682/jrrd.2011.10.0188

[3] Blausen.com staff: Medical gallery of Blausen Medical 2014, WikiJournal of Medicine, vol. 1, no. 2, 2014, DOI: 10.15347/wjm/2014.010, ISSN 2002-4436, dostęp online: https://commons.wikimedia.org/wiki/File:Blausen_0440_HandBones.png (dostęp: 23 lutego 2026 r.).

[4] Mohammadi A., Lavranos J., Zhou H., Mutlu R., Alici G., Tan Y., Choong P., Oetomo D.: A practical 3D-printed soft robotic prosthetic hand with multi-articulating capabilities, PLoS ONE, vol. 15, no. 5, 2020, art. e0232766, DOI: 10.1371/journal.pone.0232766

[5] RoboHand: projekt protezy dłoni drukowany w 3D w otwartym dostępie, dostęp online: http://www.robohand.net (dostęp: 23 lutego 2026 r.).

[6] Siegel J.R., Harwood J.K., Lau A.C., Widmer J., Parsons J., Zuniga J.M.: A performance evaluation of commercially available and 3D-printable prosthetic hands using the anthropomorphic hand assessment protocol, BMC Biomedical Engineering, vol. 6, 2024, art. 86, DOI: 10.1186/s42490-024-00086-w

[7] The Raptor Hand – projekt e-NABLE, dostęp online: https://enablingthefuture.org/upper-limb-prosthetics/the-raptor-hand/ (dostęp: 23 lutego 2026 r.).

[8] Wróbel E.: Wielofunkcyjna orteza rehabilitacyjna dłoni z możliwością blokowania chwytu przeznaczona do wytwarzania w technologii przyrostowej. W: Inżynieria biomedyczna. Metody przyrostowe w technice medycznej. Wrocław: Politechnika Wrocławska; 2017. s. 47-64.

[9] Young K., Pierce J., Zuniga J.: Assessment of body-powered 3D printed partial finger prostheses: a case study, 3D Printing in Medicine, vol. 5, no. 1, 2019, art. 7.

[10] Zhou L., Miller J., Vezza J., Mayster M., Raffay M., Justice Q., Al Tamimi Z., Hansotte G., Sunkara L.D., Bernat J.: Additive Manufacturing: A Comprehensive Review, Sensors, vol. 24, nr 9, 2024, s. 2668, DOI: 10.1186/s41205-019-0044-0

[11] Zuniga J., Katsavelis D., Peck J., Stollberg J., Petrykowski M., Carson A., Fernandez C.: Cyborg Beast: a low-cost 3D-printed prosthetic hand for children with upper-limb differences, BMC Research Notes, vol. 8, 2015, DOI:10.1186/s13104-015-0971-9

CONSTRUCTION DESIGN OF A MECHATRONIC HAND

Summary

The main goal of the project was to design and execute a construction project for a mechatronic hand using 3D printing technology, and to assess the possibility of replication of the basic motor functions of the human hand. The goal was to obtain a structure that would allow the fingers to bend and maintain a segmented structure corresponding to the anatomy. The research part used a design-experimental method. A three dimensional model was created in a CAD environment using Autodesk Fusion, where joints and mounting holes were designed. Then, a physical prototype was made using an additive method with a Creality CR-6 SE printer and PLA filament. After printing, mechanical assembly was carried out, connecting the phalanges with filament made axes and using strings to act as tendons and rubber bands to return to the initial position. The result of the project was a smoothly functioning model of a mechatronic hand that allows independent bending of individual fingers. The design retained its mechanical stability and correct fit of components without the need for additional processing. The conclusions indicate that 3D printing technology is an effective and low cost tool for prototyping mechatronic devices. The developed model provides a solid basis for further development work, including the integration of an electric drive and the use of more durable construction materials.

Keywords: mechatronics, 3D printing, artificial hand, CAD, 3D modeling, prosthesis.

Daria Ciura
Ewelina Nycz
Studenckie Koło Naukowe Komunikacji Marketingowej „Brief"

dr Damian Gębarowski
opiekun Koła Naukowego

ORIENTACJA NA KLIENTA JAKO DETERMINANT PRZEWAGI KONKURENCYJNEJ WSPÓŁCZESNYCH PRZEDSIĘBIORSTW

Streszczenie

Celem rozdziału jest wykazanie, że orientacja na klienta stanowi kluczowy determinant przewagi konkurencyjnej współczesnego przedsiębiorstwa w warunkach globalizacji i intensywnej rywalizacji rynkowej. W pracy zastosowano metodę analizy literatury z zakresu zarządzania jakością i zarządzania strategicznego oraz analizę danych wtórnych dotyczących zależności między poziomem orientacji rynkowej a konkurencyjnością przedsiębiorstw. Rozważania osadzono w kontekście systemów zarządzania jakością, w szczególności norm ISO 9001, siedmiu zasad zarządzania jakością oraz koncepcji Total Quality Management, które podkreślają centralną rolę klienta w kształtowaniu procesów organizacyjnych. Analiza wskazuje, że trwała przewaga konkurencyjna wynika przede wszystkim z umiejętności systematycznego rozpoznawania i zaspokajania obecnych oraz przyszłych potrzeb klientów. Wysoka orientacja na klienta sprzyja wzrostowi konkurencyjności, poprawie rentowności oraz budowaniu lojalności nabywców, stabilizując pozycję rynkową organizacji. Ma ona charakter strategiczny, wymaga zaangażowania całej organizacji i stanowi fundamentalny warunek jej trwałego rozwoju.

Słowa kluczowe: zarządzanie jakością, TQM, lojalność klienta, strategia organizacji, procesy biznesowe.

1. Wprowadzenie

Współczesny rynek, charakteryzujący się wysoką zmiennością oraz intensywną rywalizacją między podmiotami, wymusił na przedsiębiorstwach zmianę dotychczasowych strategii rozwoju. Globalizacja sprawiła, że konkurencja stała się szersza i bardziej międzynarodowa. Powszechny dostęp do technologii spowodował, że wiele firm dysponuje zbliżonymi możliwościami produkcyjnymi.

W konsekwencji przewaga oparta wyłącznie na optymalizacji procesów wytwórczych przestała być wystarczająca.

Jak słusznie zauważa Anna Rakowska, kluczowym wyzwaniem dla firm nie jest dziś sam proces produkcji, lecz unikalna zdolność do projektowania rozwiązań nadążających za dynamicznie ewoluującymi oczekiwaniami nabywców [12]. Z tego też wynika, że organizacje powinny koncentrować się na innych obszarach, takich jak budowanie i utrzymywanie relacji z klientami. W efekcie punkt ciężkości w zarządzaniu przesunął się z wnętrza organizacji w stronę jej otoczenia rynkowego.

Konsekwencją tej transformacji jest zasadnicza zmiana postrzegania roli konsumenta w działalności przedsiębiorstwa. W obliczu dynamicznych zmian rynkowych klient przestał być jedynie biernym odbiorcą oferowanych produktów, a stał się kluczowym zasobem organizacji. Obecnie to właśnie trwałe relacje z nabywcami stanowią fundament budowania przewagi konkurencyjnej oraz główny czynnik napędzający rozwój firmy [18]. Takie podejście prowadzi do odmiennego spojrzenia na proces zarządzania. W nowoczesnej gospodarce rynkowej centralne miejsce zajmuje klient, który pełni rolę kluczowego interesariusza, wyznaczając kierunek działań przedsiębiorstwa poprzez swoje potrzeby i oczekiwania [16, s. 46]. W konsekwencji podmioty gospodarcze są zmuszone traktować każdego odbiorcę jako istotny kapitał ekonomiczny, wymagający szczególnej uwagi oraz profesjonalnego zarządzania.

Przyjęcie orientacji na klienta nie jest zatem jedynie wyborem marketingowym, ale nadrzędnym warunkiem przetrwania na rynku. Jak zauważa Adam Hamrol, organizacja, która nie potrafi skutecznie przyciągnąć i zatrzymać klientów, zaczyna działać w próżni, co w perspektywie długofalowej pozbawia ją jakichkolwiek szans na przyszłość [2, s. 93]. Dlatego strategicznym priorytetem każdego nowoczesnego przedsiębiorstwa powinno być dogłębne zrozumienie oczekiwań rynku oraz elastyczne dostosowywanie do nich kierunków własnego rozwoju.

Celem rozdziału jest wykazanie, że orientacja na klienta stanowi kluczowy determinant przewagi konkurencyjnej współczesnych przedsiębiorstw. Analiza została przeprowadzona w kontekście systemowego zarządzania jakością, roli klienta w procesach biznesowych oraz znaczenia budowania lojalności dla wzmacniania pozycji rynkowej współczesnych organizacji. Podjęta problematyka i postawiony cel zdeterminowały wybór metod badawczych takich jak przegląd źródeł literaturowych i wnioskowanie. Rozdział ma charakter przeglądowo-teoretyczny.

2. Orientacja na klienta a zarządzanie jakością

2.1. Znaczenie jakości i rola klienta

Współczesne przedsiębiorstwo nie może funkcjonować w oderwaniu od oczekiwań swoich klientów. Tylko wtedy, gdy organizacja jest w stanie zapewnić wysoką jakość, może skutecznie rywalizować na rynku. Dlatego niezwykle istotne

jest ciągłe monitorowanie procesów technologicznych, identyfikowanie niedoskonałości oraz eliminowanie problemów, które mogłyby prowadzić do utraty klientów [16, s. 39]. Utrata klienta w warunkach silnej konkurencji oznacza nie tylko spadek przychodów, lecz również osłabienie przewagi konkurencyjnej.

Każdy aspekt działalności przedsiębiorstwa, który dotyczy jakości, powinien być podporządkowany klientowi, gdyż to on jest ostatecznym weryfikatorem efektów pracy organizacji [6, s. 42]. Istnienie firmy w dużej mierze zależy od jej klientów – to ich zakupy i ocena wyrobów stanowią najlepszy miernik jakości. Z tego powodu głównym celem każdej organizacji powinno być dążenie do pełnego zadowolenia klienta [16, s. 45].

Podstawą skutecznej orientacji na klienta jest jego dokładna identyfikacja oraz poznanie zarówno obecnych, jak i przyszłych potrzeb. Organizacja powinna nie tylko spełniać wymagania klientów, ale także starać się wykraczać ponad ich oczekiwania, wprowadzając rozwiązania tworzące dodatkową wartość [2, s. 93--94]. W praktyce oznacza to, że przedsiębiorstwo musi aktywnie pozyskiwać informacje zwrotne od rynku i monitorować, w jaki sposób jego produkty są postrzegane przez odbiorców. Prowadzi to do długofalowej lojalności klientów, a co za tym idzie – utrzymania konkurencyjności.

Według A. Hamrola tradycyjne metody oceny wyników przedsiębiorstwa, oparte głównie na wskaźnikach ekonomicznych, ustępują miejsca ocenie dokonywanej przez rynek, co umożliwia ciągłe doskonalenie jakości i lepsze dostosowanie oferty do potrzeb klientów [2, s. 65]. Szczególnie istotne w tym kontekście jest otoczenie rynkowe, w którym funkcjonuje organizacja. Wprowadzenie systemu zarządzania jakością wiąże się z uzyskaniem certyfikatów jakości oraz wyróżnień prestiżowych, które nie tylko potwierdzają jego sprawność, ale także wyróżniają firmę wśród konkurencji [2, s. 75]. Wielu klientów, dla których jakość ma kluczowe znaczenie, wykazuje większą skłonność do ponownego skorzystania z oferty.

Wdrażanie skutecznej orientacji na klienta wymaga również przestrzegania norm i systemów zarządzania jakością, takich jak ISO 9000:2015 i ISO 9001:2015. Podstawowym celem tych systemów jest spełnienie wymagań klienta oraz dążenie do przekraczania jego oczekiwań, co przyczynia się do trwałego sukcesu organizacji poprzez zdobycie i utrzymanie zaufania klientów oraz innych stron zainteresowanych [9]. Najwyższe kierownictwo organizacji powinno wykazywać przywództwo i zaangażowanie w tym obszarze, zapewniając, że wymagania klientów oraz obowiązujące przepisy prawne i regulacyjne są nie tylko określone, ale także konsekwentnie spełniane. Należy przy tym uwzględniać ryzyka i szanse wpływające na zgodność wyrobów i usług oraz na zdolność zwiększania zadowolenia klientów [10].

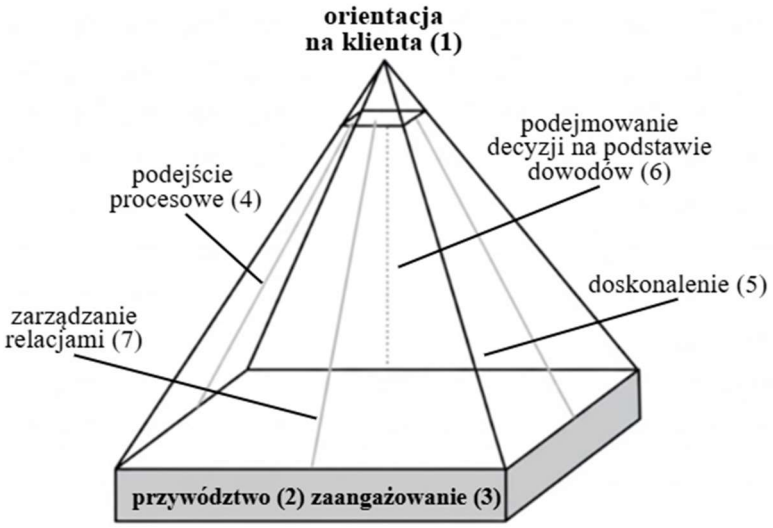
Odpowiedzialność kierownictwa obejmuje także przekazywanie wszystkim pracownikom świadomości, że spełnienie potrzeb klientów wymaga nieustannego poznawania ich oczekiwań, przekształcania ich w wymagania organizacyjne oraz wdrażania procesów pozwalających na ich realizację [2, s. 156-157].

2.2. Orientacja na klienta w siedmiu zasadach zarządzania jakością

Międzynarodowa norma ISO 9001 jest oparta na zasadach zarządzania jakością, szczegółowo opisanych w ISO 9000. Można wyróżnić siedem zasad zarządzania jakością, takich jak:

1. orientacja na klienta,
2. przywództwo,
3. zaangażowanie ludzi,
4. podejście procesowe,
5. doskonalenie,
6. podejmowanie decyzji na podstawie dowodów,
7. zarządzanie relacjami [9].

W modelu siedmiu zasad zarządzania jakością przyjęto, że wiodącą zasadą jest orientacja na klienta. Jej skuteczne wdrożenie wymaga procesowego podejścia, podejmowania decyzji na podstawie dowodów, a także utrzymywania partnerskich relacji z dostawcami, którzy mają kluczowy wpływ na spełnienie wymagań klienta [2, s. 93]. Choć każda z zasad jest istotna, model przedstawiony na rysunku 1. wyraźnie wskazuje na ich hierarchiczny i procesowy charakter.



Rys. 1. Siedem zasad zarządzania jakością – ujęcie modelowe
Fig. 1. Seven Quality Management Principles – A Model-Based Approach
Źródło: opracowanie własne na podstawie [2, s. 92].

Na samym szczycie piramidy znajduje się orientacja na klienta, która jest zwieńczeniem wszystkich działań – celem nadrzędnym, ku któremu zbiegają się pozostałe procesy. Należy zauważyć, że skuteczne ukierunkowanie na klienta

wymaga spójności całego układu. Szeroka postawa w postaci przywództwa i zaangażowania ludzi stanowi niezbędne oparcie dla wznoszących się ścian procesów i doskonalenia. Taka konstrukcja sugeruje, że satysfakcja odbiorcy nie jest odizolowanym elementem, lecz bezpośrednią wypadkową stabilności wszystkich pozostałych fundamentów jakości [2, s. 93].

Analiza przedsiębiorstwa BUDPOL autorstwa A. Hamrola wskazuje, że kluczowym celem strategicznym organizacji powinno być pełne zaspokojenie potrzeb i oczekiwań klientów, któremu nadaje się najwyższy priorytet [2, s. 96]. Skuteczna realizacja tej zasady wymaga od organizacji przejścia z modelu reaktywnego na model proaktywny. Praktycznym odzwierciedleniem zasady orientacji na klienta jest podejście zaprezentowane w tabeli 1.

Tabela 1. Znaczenie zasady: Orientacja na klienta w osiąganiu wybranych celów strategicznych firmy BUDPOL

Table 1. The importance of customer orientation in reaching specific strategic objectives of the BUDPOL company

Cel strategiczny	Organizacja zna i stosuje wytyczne zasady	W organizacji zasada nie jest stosowana
Pełne zaspokojenie potrzeb i oczekiwań klientów	Z klientem poszukującym określonej usługi jest przeprowadzany pełny wywiad (cel zakupu, metoda aplikacji, specjalne wymagania techniczne itp.). Klient otrzymuje najbardziej mu odpowiadające rozwiązanie	Potrzeby klienta są ignorowane. Dominuje nastawienie na szybki efekt w postaci sprzedaży kolejnego standardowego produktu. Powoduje to utratę szansy na poznanie potrzeb klienta, co ogranicza możliwości doskonalenia oferty produktowej

Źródło: Opracowanie własne na podstawie [2, s. 96].

W organizacji, która w pełni stosuje wytyczne zarządzania jakością, kontakt z klientem wykracza poza ramy prostej transakcji, przyjmując formę partnerstwa proaktywnego. Kontakt opiera się tu na pogłębionym wywiadzie dotyczącym celu zakupu, metod aplikacji czy specyficznych wymagań technicznych. Takie podejście pozwala na dostarczenie rozwiązania „szytego na miarę", co nie tylko bezpośrednio przekłada się na wzrost lojalności klienta, ale buduje trwałą przewagę konkurencyjną. Z kolei podejście reaktywne zorientowane na szybki efekt i oferowanie wyłącznie standardowych produktów prowadzi do utraty cennych informacji o rynku. Brak orientacji na klienta ogranicza możliwości doskonalenia oferty, co w dłuższej perspektywie drastycznie obniża konkurencyjność przedsiębiorstwa, czyniąc je podatnym na działania sprawniejszych rywali [2, s. 96].

Orientacja na klienta, jako jedna z siedmiu zasad zarządzania jakością, stanowi niezbędne narzędzie diagnostyczne i rozwojowe każdego nowoczesnego biznesu [2, s. 96]. Do kluczowych korzyści stosowania orientacji na klienta można więc zaliczyć:

- wzrost efektywności rynkowej i udziału w rynku, dzięki elastycznemu reagowaniu na zmieniające się potrzeby nabywców,
- optymalizację wykorzystania zasobów poprzez koncentracje procesów na wartościach kluczowych dla odbiorcy,
- budowanie trwałej lojalności oraz poprawę reputacji, co przekłada się na powtarzalność transakcji i stabilność biznesową organizacji [9].

W praktyce operacyjnej zastosowanie zasady orientacji na klienta prowadzi zwykle do:

- identyfikacji grup klientów (bezpośrednich i pośrednich),
- ciągłego badania i zrozumienia potrzeb i oczekiwań klienta,
- upowszechniania wiedzy o potrzebach klientów wewnątrz całej struktury,
- integracji celów organizacji z wymaganiami rynkowymi,
- orientacji procesów projektowych i produkcyjnych na pełne spełnienie wymagań odbiorców,
- monitorowania poziomu satysfakcji oraz szybkiego reagowania na sygnały z rynku,
- zarządzania relacjami z klientami oraz innymi stronami zainteresowanymi [9].

2.3. TQM jako ewolucja orientacji na klienta

Kolejnym, bardziej zaawansowanym etapem wdrażania orientacji na klienta jest koncepcja Total Quality Management (TQM), czyli kompleksowego zarządzania jakością. Istotę tego podejścia najlepiej oddaje dekompozycja samej nazwy, którą należy rozumieć jako:

a) Total – działanie kompleksowe, obejmujące całą strukturę organizacji,
b) Quality – skoncentrowanie na jakości rozumianej jako zaspokajanie zidentyfikowanych potrzeb nabywców,
c) Management – zarządzanie oparte na ciągłym dążeniu do doskonałości procesów [6, s. 140].

TQM nie jest jedynie zbiorem procedur, lecz głęboką zmianą kultury organizacyjnej. Jej głównym celem jest satysfakcja klienta – zarówno w ujęciu bieżącym, jak i przyszłościowym [6, s. 143]. W tym ujęciu to właśnie klient staje się „mechanizmem", który spaja wszystkie procesy wewnątrz firmy. Każdy pracownik, bez względu na zajmowane stanowisko, staje się ogniwem w łańcuchu relacji dostawca-klient, co wymusza pełną koncentrację na jakości pracy na każdym etapie [4, s. 7].

Z punktu widzenia strategicznego, TQM postrzega rentowność nie jako cel sam w sobie, ale jako naturalny, długofalowy skutek właściwej obsługi rynku. Choć zysk pozostaje ważny, to siłą napędową firmy staje się dostarczanie wartości klientowi [4, s. 8]. Takie podejście bezpośrednio przekłada się na wzrost konkurencyjności, ponieważ angażuje wszystkich członków organizacji w proces ciągłego doskonalenia usług i produktów [4, s. 5].

Wdrożenie tej filozofii staje się we współczesnej gospodarce kwestią „być albo nie być". Jak zauważył prekursor nowoczesnego podejścia do jakości, William Edwards Deming [6, s. 113]: „W przyszłości będą tylko dwa typy przedsiębiorstw – te, które wprowadziły jakość totalną i te, które wypadły z biznesu. Nie musicie tego robić – przetrwanie nie jest obowiązkowe, w przyszłości na rynku utrzymają się jedynie te przedsiębiorstwa, które postawią na jakość totalną, podczas gdy pozostałe po prostu z niego wypadną".

W praktyce TQM wymaga zacieśnienia więzi z rynkiem poprzez:

- aktywne poszukiwanie sygnałów od nabywców i angażowanie ich w proces projektowania produktów [6, s. 147],
- koncentrację na procesach i dostawcach, gdzie każde działanie jest ukierunkowane na wymagania rynku, a zasoby są pozyskiwane wyłącznie od wiarygodnych partnerów [2, s. 70],
- zastąpienie suchych wskaźników ekonomicznych „oceną rynkową", która staje się nadrzędnym miernikiem sukcesu firmy [6, s. 140].

TQM to filozofia, która podporządkowuje każdą płaszczyznę funkcjonowania organizacji idei totalnej jakości. W efekcie prowadzi to nie tylko do wzrostu satysfakcji klienta, ale przede wszystkim do zwiększenia elastyczności i siły konkurencyjnej całego przedsiębiorstwa [17].

3. Orientacja na klienta jako kluczowy element strategii konkurencyjnej

3.1. Klient jako centralny punkt procesów biznesowych i rozpoznawanie jego wymagań

Fundamentem nowoczesnego przedsiębiorstwa jest przyjęcie perspektywy, w której to klient, wraz ze swoimi potrzebami, staje się punktem wyjścia dla wszystkich działań operacyjnych. Rozpoznanie i zaspokojenie tych oczekiwań decyduje o sukcesie rynkowym, gdyż to właśnie w centrum uwagi organizacji musi znajdować się nabywca [6, s. 41]. W warunkach silnej konkurencji klient jest traktowany jako specyficzny „zasób", o który należy dbać poprzez oferowanie warunków korzystniejszych niż u rywali, co bezpośrednio przekłada się na wzrost udziałów w rynku i poprawę rentowności [16, s. 46].

Skuteczna orientacja na klienta wymaga jednak odejścia od działań doraźnych na rzecz systematycznego procesu. Rozpoznawanie wymagań nie może być okazjonalne – musi mieć charakter systematyczny, oparty na zestawie sprawdzonych narzędzi, takich jak:

- pogłębione wywiady i ankiety,
- aktywny udział w targach branżowych,
- stały monitoring satysfakcji, w tym analiza reklamacji i raportów rynkowych [2, s. 95].

Kluczowe jest przy tym, aby kierownictwo organizacji potrafiło terminowo rozpoznać te oczekiwania i przekształcić je w wyraźnie zdefiniowane standardy oraz wymogi prawne [6, s. 197].

Współczesnym wsparciem dla tak pojętej orientacji są systemy klasy Customer Relationship Management. Koncepcja ta zmierza do optymalizacji procesów biznesowych poprzez zintegrowanie wszystkich informacji o kliencie pozyskiwanych przez różne ogniwa firmy [18]. Wdrożenie CRM pozwala na budowanie stabilnych relacji z najlepszymi klientami dzięki reorganizacji firmy i zaawansowanej segmentacji [1]. Jak zauważa Anetta Pukas, to właśnie prawidłowo zarządzane informacje o kliencie stają się współczesną determinantą trwałej przewagi konkurencyjnej [11].

Należy jednak podkreślić, że orientacja na klienta zewnętrznego jest niemożliwa do zrealizowania bez uwzględnienia roli klienta wewnętrznego. W codziennym zarządzaniu jakością nie można zapominać o pracownikach, którzy mają kontakt z produktem na każdym etapie jego powstawania. To od ich kompetencji zależy, w jakim stopniu produkt końcowy spełni wymagania nabywcy [2, s. 94]. Współczesna praktyka rynkowa dowodzi, że innowacyjność jest domeną organizacji inwestujących w kapitał ludzki. Ciągły rozwój pracowników i specjalistyczne szkolenia nie tylko podnoszą jakość wykonania, ale przede wszystkim wyostrzają zdolność zespołu do identyfikowania ukrytych potrzeb klientów oraz przewidywania posunięć konkurencji [7].

Przedsiębiorstwo zorientowane na klienta musi stać się organizacją opartą na wiedzy i specyficznej kulturze organizacyjnej, gdzie wspólny zbiór wartości wszystkich pracowników jest nakierowany na tworzenie partnerskich relacji z rynkiem [18]. Dopiero tak przygotowany fundament procesowy i ludzki pozwala na osiągnięcie wysokich wskaźników konkurencyjności.

3.2. Wpływ orientacji na klienta na poziom konkurencyjności

Współczesna analiza strategiczna wskazuje, że konkurencyjność nie jest stanem stałym, lecz procesem ciągłego dążenia do maksymalizacji wartości użytkowej dla klienta. Przedsiębiorstwo osiąga przewagę konkurencyjną wtedy, gdy potrafi wykorzystać własne zasoby i sygnały płynące z otoczenia w sposób bardziej efektywny od swoich rywali [8, s. 13]. W praktyce oznacza to posiadanie unikalnej zdolności do oferowania produktów i usług, które odpowiadają oczekiwaniom odbiorców w stopniu wyższym niż oferty konkurencji [5].

Kluczowym dowodem na bezpośrednią korelację między podejściem rynkowym a sukcesem firmy są badania statystyczne przeprowadzone na próbie 334 przedsiębiorstw produkcyjnych. Wykazano w nich, że wraz ze wzrostem poziomu orientacji rynkowej znacząco rośnie indeks konkurencyjności jednostki [15]. Wyniki te prezentuje poniższa tabela 2.

Tabela 2. Orientacja rynkowa a konkurencyjność przedsiębiorstw produkcyjnych

Table 2. Market orientation and competitiveness of manufacturing companies

Orientacja rynkowa	Wskaźnik konkurencyjności	Różnica	Wartość p testu Mann-Whitneya U-testu	Liczba firm	Udział (%)
≤3.0	2.52	-	-	49	15
>3.0	3.22	0.70***	<0.001	285	85
>3.5	3.38	0.85***	<0.001	188	56
>4.0	3.57	1.05***	<0.001	76	23
>4.5	4.01	1.49***	<0.001	17	5

Uwagi: *** p<0.01

Źródło: Opracowanie własne na podstawie [15, s. 209].

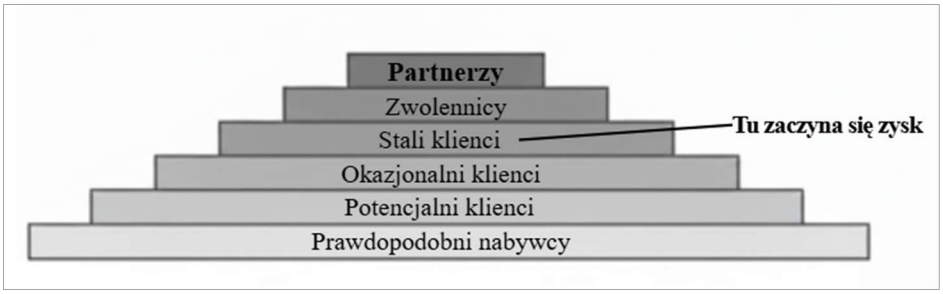
Z powyższych danych wynika, że orientacja na klienta jest szczególnie istotna w sektorach o niskim i średnim poziomie technologicznym. W tych obszarach zrozumienie potrzeb nabywcy, a nie tylko innowacji produktowych, decyduje o przetrwaniu firmy. Potwierdza to koncepcję, że przedsiębiorstwo posiada przewagę, gdy robi coś lepiej od innych lub dysponuje unikalną wiedzą o rynku, której konkurenci nie są w stanie szybko skopiować [5].

3.3. Ekonomiczne i relacyjne aspekty lojalności

Wymiernym skutkiem skutecznej orientacji na klienta jest budowanie jego lojalności. Jak wskazuje Stanisława Wilmańska-Sosnowska, zdolność do tworzenia grupy lojalnych partnerów jest znacznie bardziej efektywna niż ciągłe pozyskiwanie nowych klientów [18]. Ekonomiczne uzasadnienie tej strategii znajduje potwierdzenie w klasycznych badaniach nad retencją: wzrost wskaźnika utrzymania klientów o zaledwie 5% może przełożyć się na wzrost zysku w przedziale od 25% do nawet 100% [13]. Wynika to z faktu, że utrzymanie dotychczasowego klienta jest kilkakrotnie tańsze niż zdobycie nowego. Lojalny nabywca staje się więc darmowym i najbardziej wiarygodnym źródłem promocji firmy [14]. Proces budowania więzi nie jest jednak jednolity. Można go przedstawić w formie gradacji poziomów lojalności, co pozwala na wyodrębnienie sześciu kluczowych grup nabywców:

- prawdopodobni nabywcy, którzy jeszcze nie dokonali zakupu,
- potencjalni klienci, wykazując realne zainteresowanie ofertą,
- okazjonalni klienci, dokonujący zakupów sporadycznie bez przywiązania do marki,
- stali klienci, regularnie korzystając z usług,
- zwolennicy, klienci emocjonalni związani z firmą i polecający ją innym,
- partnerzy, najwyższy stopień, oparty na wzajemnych korzyściach [3, s. 29].

Wizualną interpretację tej ścieżki rozwoju relacji prezentuje rys. 2.



Rys. 2. Piramida lojalności
Fig. 2. Customer Loyalty Hierarchy
Źródło: Opracowanie własne na podstawie [1, s. 99; 3, s. 29].

Budowanie lojalności wykracza poza sam fakt doprowadzenia do powtarzalnych zakupów przez nabywców. Fundamentem myślenia o zysku w perspektywie długofalowej jest posiadanie stabilnej bazy stałych odbiorców. Choć regularne transakcje są niezbędnym pierwszym krokiem, ostatecznym celem organizacji powinno być przekształcenie tych kontaktów w pełnowartościowe relacje partnerskie, oparte na wzajemnej współpracy oraz otwartości. Takie podejście pozwala firmie wyjść poza tradycyjny model sprzedażowy i budować trwałą przewagę opartą na głębokim zaufaniu z klientami [1, s. 99].

4. Podsumowanie

Współczesna orientacja na klienta wyewoluowała z prostej techniki sprzedażowej do rangi nadrzędnego determinantu strategii każdego przedsiębiorstwa. W obliczu globalizacji i niemal identycznych możliwości technologicznych wielu firm, to właśnie klient stał się jednym z najważniejszych zasobów, źródłem i siłą napędową rynkowych zmagań [18]. Jak wykazano, przewaga konkurencyjna nie jest dziś budowana wyłącznie na optymalizacji produkcji, lecz na unikalnej zdolności organizacji do rozpoznawania i zaspokajania dynamicznie zmieniających się potrzeb nabywców, co stanowi istotę nowoczesnego zarządzania jakością.

Wdrożenie systemowych rozwiązań, takich jak normy ISO 9001 czy filozofia TQM, pozwala na trwałe osadzenie klienta w centrum wszystkich procesów biznesowych. Dzięki takiemu podejściu, jakość przestaje być parametrem technicznym, a staje się stopniem zadowolenia odbiorcy, który jest ostatecznym weryfikatorem działań firmy [16, s. 46]. Strategiczne nastawienie na budowanie lojalności klientów-partnerów jest znacznie bardziej efektywne i rentowne niż agresywna ekspansja nakierowana na doraźne pozyskiwanie nowych rynków. Potwierdzają to dane ekonomiczne, wskazujące na drastyczny wzrost zysków nawet przy niewielkiej poprawie wskaźnika utrzymania dotychczasowych nabywców [13].

Ostatecznie należy stwierdzić, że orientacja na klienta to proces ciągły, wymagający zaangażowania całej struktury organizacyjnej – od najwyższego kierownictwa po pracowników liniowych, czyli klientów wewnętrznych [2, s. 94]. Przedsiębiorstwa, które potrafią skutecznie przekształcić sygnały płynące z otoczenia w konkretną wartość rynkową, osiągają wyższy indeks konkurencyjności i budują trwałą odporność na działania rywali [15]. W obecnej gospodarce rynkowej inwestycja w relacje z klientem nie jest zatem wyborem, lecz fundamentalnym warunkiem sukcesu i przetrwania.

Literatura

[1] Daszkiewicz M., Waniowski P.: Relacje z klientami w tworzeniu przewagi konkurencyjnej przedsiębiorstw, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 33, 2008, s. 94-16, https://www.dbc.wroc.pl (dostęp: 20.02.2026 r.).

[2] Hamrol A.: Zarządzanie jakością z przykładami, Wydawnictwo Naukowe PWN, Warszawa 2013.

[3] Hill N., Alexander J.: Pomiar satysfakcji i lojalności klientów, Oficyna Ekonomiczna, Kraków 2003.

[4] Kordos J.: TQM – zarys teorii i praktyki, Zarządzanie. Teoria i Praktyka, nr 2 (20), 2017, s. 3-10, https://bazhum.muzhp.pl (dostęp: 16.02.2026 r.).

[5] Kwiecień A., Żak K.: Relacje z klientami w tworzeniu przewagi konkurencyjnej przedsiębiorstwa, Zarządzanie i Finanse, t. 4, nr 3, 2013, s. 263-276, https://www.academia.edu (dostęp: 15.02.2026 r.).

[6] Mroczko F.: Zarządzanie jakością, Prace Naukowe Wałbrzyskiej Wyższej Szkoły Zarządzania i Przedsiębiorczości, Seria: Zarządzanie, wydanie specjalne, Wałbrzych 2012, https://wwszip.pl (dostęp: 16.02.2026 r.).

[7] Pawłowska A.: Kompetencja „orientacja na klienta" u osób zakładających działalność gospodarczą oraz jej znaczenie dla procesu inwestowania w rozwój pracowników, Problemy Zarządzania 2015, t. 13, nr 1, https://bazhum.muzhp.pl (dostęp: 19.02. 2026 r.).

[8] Piotrowska A.: Konkurencyjność przedsiębiorstw – główne aspekty i uwarunkowania, Przegląd Nauk Ekonomicznych, nr 32, 2019, s. 9-21, https://dspace.uni.lodz.pl (dostęp: 18.02.2026 r.).

[9] PN-EN ISO 9000:2015-10 – Systemy zarządzania jakością – Podstawy i terminologia, Polski Komitet Normalizacyjny, Warszawa 2016.

[10] PN-EN ISO 9001:2015-10 – Systemy zarządzania jakością – Wymagania, Polski Komitet Normalizacyjny, Warszawa 2016.

[11] Pukas A.: Zarządzanie relacjami z klientem w tworzeniu przewagi konkurencyjnej przedsiębiorstwa – ujęcie dynamiczne, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2019.

[12] Rakowska A.: Przewaga konkurencyjna i kompetencje polskich przedsiębiorstw w kontekście wymagań stawianych nowoczesnym organizacjom, Organizacja i Zarządzanie: kwartalnik naukowy, nr 4, 2008, s. 7-20, https://oamquarterly.polsl.pl (dostęp: 15.02. 2026 r.).

[13] Reichheld F., Sasser W. E.: Zero Defections: Quality Comes To Service, Harvard Business Review, nr 9-10, 1990.

[14] Schulz M.: Wpływ satysfakcji na lojalność klientów na rynku B2B, Ekonomiczne Problemy Usług, nr 72, 2011, s. 107-116, https://bazhum.muzhp.pl (dostęp: 18.02.2026 r.).

[15] Soniewicki M.: Market orientation and competitiveness of Polish manufacturing companies, International Business and Global Economy, nr 36, 2017, pp. 209–221, https://ejournals.eu/czasopismo/ibage (dostęp: 18.02.2026 r.).

[16] Strojek-Filus M., Krynke M., Zasadzień M.: Wybrane aspekty funkcjonowania przedsiębiorstwa – koszty, konkurencyjność, jakość, PA NOVA S.A., Gliwice 2018, http://www.stegroup.pl (dostęp: 16.02.2026 r.).

[17] Sudolska A.: Orientacja na klienta – klient wewnętrzny i zewnętrzny organizacji, w: Czytelnik czy klient?, Toruń 4-6 grudnia 2003 roku, Stowarzyszenie Bibliotekarzy Polskich, https://www.ebib.pl (dostęp: 17.02.2026 r.).

[18] Wilmańska-Sosnowska S.: Lojalność klientów jako cel marketingowy współczesnego przedsiębiorstwa, Acta Universitatis Lodziensis. Folia Oeconomica, t. 195, 2005, s. 13-21, http://bazekon.icm.edu.pl/bazekon/element/bwmeta1.element.ekon-element-000109857945 (dostęp: 15.02.2026 r.).

CUSTOMER ORIENTATION AS A DETERMINANT OF COMPETITIVE ADVANTAGE IN MODERN ENTERPRISES

Summary

The aim of the article is to demonstrate that customer orientation constitutes a key determinant of competitive advantage in contemporary enterprises operating under conditions of globalization and intense market competition. The study employs a literature review in the fields of quality management and strategic management, as well as an analysis of secondary data concerning the relationship between the level of market orientation and enterprise competitiveness. The discussion is framed in the context of quality management systems, in particular ISO 9001 standards, the seven quality management principles, and the Total Quality Management concept, all of which emphasize the central role of the customer in shaping organizational processes. The analysis indicates that sustainable competitive advantage results primarily from the ability to systematically identify and meet current and future customer needs. A high level of customer orientation supports increased competitiveness, improved profitability, and the development of customer loyalty, which over the long term stabilizes the enterprise's market position. Customer orientation has a strategic character, requires the involvement of the entire organization, and constitutes a fundamental condition for its long-term development.

Keywords: quality management, TQM, customer loyalty, organizational strategy, business processes.

Julia Ziąbek, Natalia Klocek, Maja Jaszowska,
Dominik Dworak, Sebastian Cwynar
Koło Naukowe Elektroniki i Technologii Informacyjnych

dr inż. Bartosz Pawłowicz
opiekun Koła Naukowego

ANALIZA BEZPIECZEŃSTWA STRON WYGENEROWANYCH ZA POMOCĄ AI

Streszczenie

Celem pracy jest określenie poziomu bezpieczeństwa stron internetowych wygenerowanych przy użyciu kilku dużych modeli językowych, a następnie analiza wpływu treści i stopnia szczegółowości promptów na jakość uzyskanych zabezpieczeń. W procesie weryfikacji wykorzystane zostało darmowe, otwartoźródłowe narzędzie do testowania bezpieczeństwa aplikacji webowych – OWASP ZAP. Prompty wykorzystywane do generowania kodu były konstruowane w sposób iteracyjny i stopniowo zwiększały poziom złożoności oraz precyzji wymagań dotyczących zarówno funkcjonalności, jak i aspektów bezpieczeństwa aplikacji. Zaczynając od podejścia jakie może mieć zwykły użytkownik, kończąc na poziomie seniora w kwestii bezpieczeństwa. W zależności od stopnia doprecyzowania poleceń oraz liczby wprowadzonych poprawek możliwe było wygenerowanie aplikacji o zróżnicowanym poziomie ochrony – od stron podatnych na podstawowe zagrożenia, takie jak brak walidacji danych wejściowych, podatność na ataki typu XSS czy niewłaściwa konfiguracja nagłówków HTTP, po rozwiązania spełniające kluczowe zalecenia standardów bezpieczeństwa. Analizie poddano wpływ sposobu formułowania promptów na eliminację podatności zgodnie z wytycznymi OWASP Top 10. Projekt zakłada wieloetapowe podejście badawcze, w którym początkowa wersja aplikacji nie uwzględnia zasad bezpiecznego programowania, a następnie podlega kolejnym iteracjom generowania kodu z rosnącym naciskiem na usuwanie wykrytych podatności. Takie podejście pozwala ocenić, w jakim stopniu modele językowe są w stanie reagować na wskazane problemy bezpieczeństwa oraz czy możliwe jest wygenerowanie aplikacji spełniającej podstawowe wymagania bezpieczeństwa z ich użyciem.

Słowa kluczowe: Podatności, model językowy, generowanie kodu.

1. Wstęp

Dynamiczny rozwój dużych modeli językowych (LLM – *Large Language Models*) sprawił, że coraz częściej są one wykorzystywane do automatycznego generowania kodu aplikacji internetowych, obejmujących zarówno warstwę frontendową, backendową, jak i bazodanową. Choć rozwiązania te znacząco przyspieszają proces tworzenia oprogramowania i obniżają próg wejścia dla programistów, rodzą jednocześnie istotne pytania dotyczące bezpieczeństwa generowanych aplikacji. W szczególności problematyczne może być powstawanie luk bezpieczeństwa wynikających z braku uwzględnienia dobrych praktyk oraz standardów ochrony danych już na etapie generowania kodu.

Celem niniejszego rozdziału jest analiza bezpieczeństwa stron internetowych wygenerowanych przy użyciu różnych promptów przekazywanych do modeli językowych. Ocena ta zostanie przeprowadzona w oparciu o wytyczne organizacji OWASP (Open Web Application Security Project). OWASP stanowi powszechnie uznany punkt odniesienia w procesie identyfikacji i klasyfikacji podatności występujących w aplikacjach webowych.

Analiza nie będzie ograniczać się wyłącznie do identyfikacji podatności, lecz uwzględni również ocenę wpływu zastosowanych mechanizmów bezpieczeństwa na użyteczność aplikacji. Skuteczne zabezpieczenia powinny zapewniać odpowiedni poziom ochrony przy jednoczesnym zachowaniu wysokiej dostępności i wygody użytkowania. W związku z tym przeanalizowane zostaną między innymi wymagania dotyczące procesu uwierzytelniania, złożoności haseł oraz innych mechanizmów kontroli dostępu pod kątem ich proporcjonalności i potencjalnego wpływu na doświadczenia użytkowników. Celem jest ocena, czy wygenerowane aplikacje zachowują właściwy kompromis pomiędzy bezpieczeństwem a funkcjonalnością.

2. Generowanie kodu przez AI

Sztuczna inteligencja analizuje duże zbiory danych w tym kodu lub dokumentacji, w celu identyfikowania wzorców, tworzenia prognoz i generowania treści. Jest to swojego rodzaju dostosowywanie treści pod użytkownika tak, aby ten był jak najbardziej zadowolony z otrzymanego wyniku. Przyspiesza to generowanie kodu ale nie daje użytkownikowi pełnej kontroli nad tym co zostanie wygenerowane. Proces ten można określić często jako zgadywanie tego o co chodzi tworzącemu prompt, często z pominięciem kluczowych funkcjonalności lub bez uwzględnienia pewnych zasad bezpieczeństwa. Obecnie bezpieczniej jest używać sztucznej inteligencji jako asystenta do wykonywania najbardziej powtarzalnych zadaniań. AI nie jest w stanie projektować całych architektur oprogramowania ani podejmować strategicznych decyzji, może przejąć czasochłonne zadania, takie jak refaktoryzacja kodu lub generowanie testów jednostkowych.

3. OWASP Top Ten

Jest to globalna organizacja non-profit, której celem jest poprawa bezpieczeństwa aplikacji internetowych. Jej działalność obejmuje tworzenie darmowych narzędzi, dokumentacji oraz najlepszych praktyk związanych z cyberbezpieczeństwem. Będziemy się stosować do listy kryteriów stworzonej w 2021 roku.

3.1. Awaria kontroli dostępu (Broken Access Control)

Awaria kontroli dostępu występuje wtedy, gdy aplikacja nie egzekwuje prawidłowo reguł uprawnień, przez co użytkownik może wykonywać działania wykraczające poza przyznane mu uprawnienia. Skutkiem takich błędów może być nieautoryzowany dostęp do danych, ich odczyt, modyfikacja lub usunięcie, a także wykonywanie operacji administracyjnych przez osoby nieuprawnione. Typowym przykładem jest manipulacja adresem URL lub parametrami żądania w celu uzyskania dostępu do zasobów należących do innego użytkownika albo do danych poufnych.

3.2. Błędy kryptograficzne (Cryptographic Failures)

Błędy kryptograficzne obejmują nieprawidłowe stosowanie mechanizmów kryptograficznych lub ich całkowity brak, co może prowadzić do ujawnienia danych wrażliwych. Do najczęstszych problemów należą stosowanie haseł zapisanych na stałe w kodzie, używanie przestarzałych lub niebezpiecznych algorytmów kryptograficznych oraz wykorzystywanie niewystarczającej entropii przy generowaniu kluczy lub tokenów. Podstawowym przykładem takiej podatności jest przechowywanie poufnych danych, takich jak hasła, w postaci jawnego tekstu.

3.3. Wstrzyknięcie (Injection)

Podatność typu wstrzyknięcie polega na możliwości przekazania do aplikacji złośliwego kodu lub polecenia, które następnie zostaje wykonane przez interpreter, bazę danych lub inny komponent systemu. Ryzyko występuje szczególnie wtedy, gdy dane wejściowe od użytkownika nie są odpowiednio walidowane, filtrowane lub oczyszczane, a aplikacja wykorzystuje dynamicznie budowane zapytania lub polecenia bez bezpiecznej parametryzacji. Skutkiem ataku może być nieautoryzowany dostęp do danych, ich modyfikacja, usunięcie, a także przejęcie kontroli nad częścią systemu. Przykładem jest SQL Injection.

3.4. Niebezpieczny projekt (Insecure Design)

Niebezpieczny projekt odnosi się do błędów na poziomie koncepcji, architektury lub logiki bezpieczeństwa aplikacji. Kategoria ta obejmuje sytuacje, w których już na etapie projektowania nie przewidziano odpowiednich mechanizmów ochronnych albo zaprojektowano je w sposób nieskuteczny. Przykładami mogą być brak ograniczeń dla działań użytkownika, niewłaściwe podejście do ochrony

danych uwierzytelniających czy ujawnianie zbyt szczegółowych informacji w komunikatach błędów. W odróżnieniu od błędów implementacyjnych, niebezpieczny projekt nie może zostać w pełni naprawiony wyłącznie przez poprawne zakodowanie rozwiązania, ponieważ problem leży w samych założeniach projektowych.

3.5. Nieprawidłowa konfiguracja zabezpieczeń (Security Misconfiguration)

Nieprawidłowa konfiguracja zabezpieczeń występuje wtedy, gdy elementy środowiska aplikacji, serwera, usług chmurowych lub samej aplikacji nie zostały właściwie zabezpieczone. Do najczęstszych przyczyn należą pozostawienie domyślnych kont i haseł, włączone niepotrzebne usługi, porty lub funkcje, błędnie skonfigurowane uprawnienia, a także zbyt szczegółowe komunikaty błędów ujawniające informacje techniczne o systemie. Ryzyko wzrasta również wtedy, gdy oprogramowanie nie jest aktualizowane. Tego rodzaju błędy mogą ułatwiać atakującemu rozpoznanie środowiska i wykorzystanie kolejnych podatności.

3.6. Komponenty podatne na ataki i przestarzałe (Vulnerable and Outdated Components)

Podatność ta dotyczy wykorzystywania komponentów, bibliotek, frameworków lub innych zależności zawierających znane luki bezpieczeństwa albo nieutrzymywanych przez dostawcę. Problemem może być zarówno używanie przestarzałych wersji oprogramowania, jak i brak pełnej wiedzy o wszystkich komponentach stosowanych w aplikacji, w tym o zależnościach pośrednich. Ryzyko zwiększa brak regularnego skanowania podatności, brak aktualizacji oraz nierealizowanie testów zgodności po wdrożeniu poprawek. W efekcie atakujący może wykorzystać publicznie znane luki do naruszenia bezpieczeństwa systemu.

3.7. Błędy identyfikacji i uwierzytelniania (Identification and Authentication Failures)

Błędy identyfikacji i uwierzytelniania dotyczą nieprawidłowości w potwierdzaniu tożsamości użytkownika, obsłudze logowania, zarządzaniu sesją oraz ochronie danych uwierzytelniających. Do typowych problemów należą dopuszczanie ataków siłowych, credential stuffing, stosowanie słabych lub domyślnych haseł, brak skutecznych mechanizmów odzyskiwania dostępu do konta oraz niewłaściwe przechowywanie haseł, na przykład w postaci jawnej lub z użyciem słabego haszowania. Istotnym czynnikiem ryzyka jest również brak uwierzytelniania wieloskładnikowego lub jego nieprawidłowe wdrożenie. Skutkiem takich błędów może być przejęcie kont użytkowników.

3.8 Awarie oprogramowania i integralności danych (Software and Data Integrity Failures)

Awarie integralności oprogramowania i danych dotyczą sytuacji, w których aplikacja lub infrastruktura nie zapewnia odpowiedniej ochrony przed nieautoryzowaną modyfikacją kodu, aktualizacji, bibliotek albo danych. Ryzyko występuje m.in. wtedy, gdy system korzysta z komponentów pochodzących z niezaufanych źródeł, nie weryfikuje integralności aktualizacji lub posiada niewłaściwie zabezpieczony proces CI/CD. Zagrożeniem jest również niebezpieczna deserializacja danych, umożliwiająca manipulowanie strukturami danych przez atakującego. Konsekwencją może być uruchomienie złośliwego kodu, przejęcie kontroli nad aplikacją lub naruszenie integralności przetwarzanych danych.

3.9. Rejestrowanie i monitorowanie błędów zabezpieczeń (Security Logging and Monitoring Failures)

Błędy rejestrowania i monitorowania zdarzeń bezpieczeństwa występują wtedy, gdy aplikacja lub infrastruktura nie zapisuje istotnych zdarzeń, nie analizuje ich lub nie reaguje odpowiednio na wykryte incydenty. Brak skutecznego logowania i monitorowania utrudnia wykrycie prób ataku, naruszeń bezpieczeństwa oraz działań nieautoryzowanych użytkowników. Problemem może być zarówno brak rejestracji krytycznych zdarzeń, jak i brak mechanizmów alarmowania, korelacji zdarzeń oraz procedur reagowania. W rezultacie incydenty mogą pozostać niezauważone przez długi czas, co zwiększa skalę potencjalnych szkód.

3.10. Fałszowanie żądań po stronie serwera (Server-Side Request Forgery (SSRF))

Podatność SSRF występuje wtedy, gdy aplikacja pobiera zasoby zdalne na podstawie danych wejściowych użytkownika bez odpowiedniej walidacji lub ograniczeń. Umożliwia to atakującemu wymuszenie na serwerze wysłania żądania do nieoczekiwanego zasobu, w tym do systemów wewnętrznych, usług administracyjnych lub zasobów chronionych przez zaporę sieciową, VPN albo listy kontroli dostępu. W efekcie możliwe staje się obejście zabezpieczeń sieciowych, rozpoznanie infrastruktury wewnętrznej, a w niektórych przypadkach również dostęp do poufnych danych lub dalsza eskalacja ataku.

4. Pierwszy prompt

Pierwszy prompt, został napisany na wzór osoby, która programowała mało lub nie programowała wcale. Główne założenie: strona ma działać i ładnie wyglądać. Jest to tak zwane zjawisko *vibe codingu*, które z samym programowaniem często nie ma wiele wspólnego, oddajemy wszystkie decyzje - nawet te strategicznie - sztucznej inteligencji co może prowadzić do wielu niedociągnięć. Do tego promptu nie dodajemy żadnych informacji o bezpieczeństwie, chcemy zobaczyć

jak wiele luk w zabezpieczeniach może się pojawić w takim przypadku. Postanowiono użyć frameworka Flask w Pythonie, ponieważ jest on prosty i popularny do tworzenia małych aplikacji webowych oraz SQLite jako bazy danych ze względu na jego lekkość i łatwość użycia w małych projektach.

Cześć, uczę się programowania i chcę stworzyć swoją pierwszą aplikację do portfolio. Napisz dla mnie kod prostej aplikacji webowej w Pythonie (używając frameworka Flask). Aplikacja ma być 'Osobistym Dziennikiem'.

Wymagania funkcjonalne:

1. Użytkownik może założyć konto i się zalogować.
2. Po zalogowaniu widzi listę swoich wpisów i może dodać nowy wpis (tytuł i treść).
3. Dodaj prostą wyszukiwarkę, żeby można było znaleźć wpis po tytule.
4. Użyj bazy danych SQLite.

Wymagania wizualne:

1. Zrób to w jednym pliku app.py (z HTML w środku jako string lub oddzielnie, jak wolisz), żeby było mi łatwo to uruchomić.
2. Dodaj ładny styl CSS, żeby strona wyglądała nowocześnie, miała kolory i wyśrodkowane elementy

Zależy mi, żeby kod po prostu działał po skopiowaniu.

Treść pierwszego użytego promptu.

4.1. Chat GPT 5 - analiza bezpieczeństwa

Co ciekawe po wygenerowaniu kodu przez AI, zapytał się on czy dodać hashowanie haseł. Oznacza to, że jest w stanie sam określić pewne luki w zabezpieczeniach i zaproponować ich poprawę, natomiast nie dodał tego domyślnie. Został użyty kod wygenerowany pierwotnie bez żadnych poprawek. Poniżej znajduje się analiza bezpieczeństwa wygenerowanego kodu.

Analiza wykazuje, że aplikacja w obecnym stanie jest krytycznie podatna na ataki. Zidentyfikowano wiele krytycznych podatności, które pozwoliłyby atakującemu (zarówno uwierzytelnionemu, jak i nieuwierzytelnionemu) na całkowite przejęcie serwera, kradzież danych użytkowników i zakłócenie dostępności usług. Najpoważniejszym znaleziskiem jest podatność Server-Side Template Injection (SSTI), wynikająca z niewłaściwego użycia "f-stringów" w Pythonie wewnątrz funkcji render_template_string. W kodzie pokazuje to ta linia:

```
<input type='text' name='search' placeholder='Szukaj po tytule...'
    value='{search}'>
```

Listing 1. Podatność Server-Side Template Injection

Luka ta umożliwia zdalne wykonanie złośliwego kodu, dając atakującemu pełną kontrolę nad serwerem.

Dodatkowo, aplikacja cierpi na podatności Cross-Site Scripting (XSS), ponieważ treść wpisu pobrana z bazy jest wklejana bezpośrednio do HTML. Jeśli ktoś doda notatkę ze złośliwym skryptem, wykona się on u każdego, kto wejdzie na Dashboard:

```
entries_html = "".join([f"<div
    class='entry'><h4>{e[1]}</h4><p>{e[2]}</p></div>" for e in entries])
```

Listing 2. Podatność Cross-Site Scripting

Jednak ze względu na to, że każdy użytkownik ma dostęp tylko do swoich wpisów to podatność znaleziona przez OWASP ZAP to tak naprawdę Self-XSS, która jest mniej groźna, ale nadal niepożądana. Jeśli istniałby panel administratora, który mógłby przeglądać wpisy wszystkich użytkowników, to atakujący mógłby wykorzystać tę lukę do wykonania złośliwego kodu w kontekście administratora i przejęcia kontroli nad całą aplikacją. Można też wykorzystać tę lukę do kradzieży ciasteczek sesyjnych innych użytkowników, jeśli atakujący jest w stanie nakłonić ich do odwiedzenia spreparowanego linku lub strony zawierającej złośliwy skrypt.

Jak wspomniano na początku analizy, hasła użytkowników są przechowywane w bazie danych w postaci zwykłego tekstu, co stanowi poważne naruszenie zasad bezpieczeństwa:

```
c.execute("INSERT INTO users(username, password) VALUES(?, ?)", (username,
password))
```

Listing 3. Podatność związana z przechowywaniem haseł jawnym tekstem

Za to zostało wykorzystane mechanizm zapytań parametryzowanych, co chroni przed atakami SQL Injection w tym miejscu. Dzięki temu sterownik bazy danych sqlite3 traktuje dane wejściowe wyłącznie jako wartości literalne, a nie jako wykonywalny kod SQL.

Klucz służący do podpisywania sesji jest wpisany na stałe w kodzie. Jeśli kod wycieknie, każdy może sfałszować sesję administratora, model językowy nie zaproponował jego zmiany.

```
app.secret_key = "supersecretkey"
```

Listing 4. Podatność związana z zapisem klucza służącego do podpisywania sesji wpisany na stałe w kodzie

Formularze nie mają ukrytego pola z tokenem CSRF. Inna strona może wysłać żądanie POST do aplikacji w imieniu zalogowanego użytkownika, co może prowadzić do nieautoryzowanych działań.

```
<form method='POST'>
<input type='text' name='title'...>
```

Listing 5. Podatność związana z jawnym polem z tokenem CSRF

Brak mechanizmów ograniczających próby logowania sprawia, że aplikacja jest podatna na ataki siłowe. Atakujący może wielokrotnie próbować zgadnąć hasło użytkownika bez żadnych konsekwencji.

Aplikacja ma włączony Debug mode, więc w razie błędu wyświetli interaktywny debugger, który pozwala na wykonanie kodu na serwerze i ujawnia kod źródłowy. Jest to jednak tylko w trybie deweloperskim, więc ta luka zostanie pominięta w dalszej analizie.

Inną podatnością jest wyciek informacji poprzez komunikaty o błędach. Na przykład, podczas rejestracji, jeśli nazwa użytkownika jest już zajęta, aplikacja zwraca szczegółowy komunikat o błędzie, który może pomóc atakującemu w identyfikacji istniejących kont:

```
except Exception:
msg = "Ta nazwa użytkownika jest już zajęta!"
```

Listing 6. Podatność związana z jawną identyfikacją istniejących kont

Można też znaleźć informację o serwerze w nagłówkach HTTP, co może pomóc atakującemu w identyfikacji potencjalnych podatności specyficznych dla używanego oprogramowania serwera. Ta konkretna podatność będzie występować na każdej stronie ze względu na sposób ich testowania - lokalny serwer Flask domyślnie nadpisuje tę informację podczas uruchamiania aplikacji. Dlatego będzie to sprawdzane manualnie w kodzie.

Oprócz tego znaleziono kilka mniejszych podatności, takich jak brak zabezpieczeń nagłówków HTTP (np. Content-Security-Policy, X-Content-Type-Options), które mogą pomóc w ochronie przed różnymi atakami webowymi. Dodatkowo ciasteczka sesyjne nie mają ustawionych flag Secure i HttpOnly, co zwiększa ryzyko kradzieży sesji poprzez ataki XSS lub przechwycenie ruchu sieciowego. Brak CAPTCHA pozwala na zautomatyzowane ataki na formularze logowania i rejestracji, natomiast CAPTCHA może być uciążliwa dla użytkowników, więc nie jest to krytyczna luka. Przy funkcji strony należy się też zastanowić jak wiele zabezpieczeń jest potrzebnych, aby nie utrudniać korzystania z aplikacji, żeby nie zniechęcić użytkowników.

Podsumowanie znalezionych podatności znajduje się w tabeli 1.

Tabela 1. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez ChatGPT

Table 1. Summary of detected vulnerabilities for the application generated by ChatGPT

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Server-Side Template Injection (SSTI)	A03:2021 - Injection	Niewłaściwe użycie f-stringów w funkcji.	Możliwość zdalnego wykonania kodu, co daje pełną kontrolę nad serwerem (podatność krytyczna).
Przechowywanie haseł jawnym tekstem	A02:2021 - Cryptographic Failures	Hasła zapisywane są w bazie danych bez hashowania.	Natychmiastowe ujawnienie danych logowania wszystkich użytkowników w przypadku wycieku bazy.
Trwale wpisany klucz sekretu	A02:2021 - Cryptographic Failures	Klucz app.secret_key wpisany na stałe w kodzie.	Możliwość sfałszowania sesji i przejęcia konta dowolnego użytkownika.
Self-Cross-Site Scripting (Stored XSS)	A03:2021 - Injection	Brak sanityzacji treści wpisów przy wyświetlaniu (bezpośrednie wstrzyknięcie do HTML).	Wykonanie złośliwego kodu JS u autora. Groźne w połączeniu z atakiem CSRF.
Brak ochrony CSRF	A01:2021 - Broken Access Control	Brak tokenów anty-CSRF w formularzach POST	Możliwość wymuszenia akcji w imieniu zalogowanego użytkownika bez jego wiedzy
Brak Rate Limitingu	A07:2021 – Identification and Authentication Failures	Brak ograniczeń liczby prób logowania.	Podatność na ataki siłowe na hasła użytkowników.
User Enumeration	A07:2021 – Identification and Authentication Failures	Różne komunikaty błędów przy logowaniu lub rejestracji.	Umożliwia weryfikację istnienia konkretnych nazw użytkowników w systemie.
Niezabezpieczone ciasteczka	A05:2021 – Security Misconfiguration	Brak flag HttpOnly i Secure dla ciasteczek sesyjnych.	Zwiększone ryzyko kradzieży sesji przez XSS oraz podsłuch sieciowy.
Brak nagłówków bezpieczeństwa	A05:2021 – Security Misconfiguration	Brak CSP, X-ContentType-Options	Zmniejszona odporność na ataki webowe.
Ujawnienie wersji serwera	A05:2021 – Security Misconfiguration	Nagłówki HTTP ujawniają używane oprogramowanie serwera.	Ułatwiony rekonesans dla atakującego.
Brak ochrony przed Clickjackingiem	A05:2021 – Security Misconfiguration	Brak nagłówka X-Frame-Options lub odpowiedniego CSP.	Umożliwia osadzenie aplikacji w ramce (iframe) i wykonanie ataku UI Redressing.

4.2. Gemini 3 (Pro) – analiza bezpieczeństwa

W przeciwieństwie do kodu wygenerowanego przez ChatGPT, głównie zostały użyte odpowiednie biblioteki do obsługi bezpieczeństwa, takie jak Flask-Login do zarządzania sesjami użytkowników oraz Werkzeug do bezpiecznego hashowania haseł. Dzięki temu hasła nie są przechowywane w bazie danych w postaci zwykłego tekstu, co znacznie zwiększa bezpieczeństwo aplikacji.

Jedną z najpoważniejszych luk w aplikacji jest brak ochrony przed atakami Cross-Site Request Forgery (CSRF). Formularze w aplikacji nie zawierają tokenów CSRF, co oznacza, że złośliwa strona mogłaby wysłać żądania w imieniu zalogowanego użytkownika bez jego wiedzy. To może prowadzić do nieautoryzowanych działań, takich jak dodawanie wpisów do dziennika:

```
DASHBOARD_HTML = LAYOUT.replace("{% block content %}{% endblock %}",
"""
  ...
  <div class="card">
    <h3>Nowy wpis</h3>
    <form method="POST">
      <input type="text" name="title" placeholder="Tytuł wpisu..." required>
       <textarea name="content" rows="4" placeholder="Co u Ciebie słychać?" re-
quired></textarea>
      <button type="submit">Zapisz w dzienniku</button>
    </form>
  </div>
  ...
""")
```

Listing 7. Podatność na ataki Cross-Site Request Forgery

Brak weryfikacji po stronie serwera sprawia, że jeśli zalogowany użytkownik wejdzie na złośliwą stronę przygotowaną przez atakującego, przeglądarka może w tle wysłać żądanie POST do /dashboard, dodając niechciany wpis bez wiedzy ofiary.

Klucz służący do kryptograficznego podpisywania ciasteczek sesyjnych został wpisany na stałe bezpośrednio w kodzie źródłowym. Natomiast w tym przypadku klucz wygląda inaczej:

```
app.config['SECRET_KEY'] = 'twoj-bardzo-tajny-klucz-zmien-go' # Klucz do zabez-
pieczenia sesji
```

Listing 8. Podatność związana z wpisaniem na stałe w kodzie kryptograficznego klucza służącego do podpisywania ciasteczek sesyjnych

Ma on w sobie bezpośrednio informację, że powinien zostać zmieniony, co sugeruje, że AI zdaje sobie sprawę z ryzyka związanego z hardcodowaniem sekretów w kodzie, dlatego w tym przypadku nie zostanie to uznane za lukę w zabezpieczeniach.

Tak jak w poprzednim kodzie, tutaj również nie zastosowano żadnych mechanizmów ograniczających liczbę prób logowania. Brak rate limiting sprawia, że aplikacja jest podatna na ataki siłowe, gdzie atakujący może wielokrotnie próbować zgadnąć hasło użytkownika bez żadnych konsekwencji.

Aplikacja nie implementuje żadnych nagłówków bezpieczeństwa HTTP, takich jak Content-Security-Policy (CSP) czy X-Content-Type-Options. Brak tych nagłówków może ułatwić ataki typu Cross-Site Scripting (XSS) oraz inne ataki webowe. Nie ma też ustawionych flag Secure i HttpOnly dla ciasteczek sesyjnych, co zwiększa ryzyko kradzieży sesji poprzez ataki XSS lub przechwycenie ruchu sieciowego.

W tym przypadku również aplikacja ujawnia informacje o serwerze w nagłówkach HTTP, co może pomóc atakującemu w identyfikacji potencjalnych podatności specyficznych dla używanego oprogramowania serwera oraz ujawnia informacje o użytkownikach podczas rejestracji:

```
@app.route('/register', methods=['GET', 'POST'])
def register():
    if request.method == 'POST':
        ...
        if User.query.filter_by(username=username).first():
            flash('Taki użytkownik już istnieje!')
            return redirect(url_for('register'))
```

Listing 9. Podatność związana z ujawnieniem informacji o serwerze w nagłówkach HTTP

OWASP ZAP wykrył podatność SQL Injection w formularzu logowania. Manualna weryfikacja kodu sugeruje jednak, że jest to False Positive, ponieważ kod wykorzystuje bibliotekę SQLAlchemy, która domyślnie stosuje zapytania parametryzowane:

```
user = User.query.filter_by(username=username).first()
```

Listing 10. Zapytania parametryzowane w SQLAlchemy

W tym przypadku SQLAlchemy bezpiecznie obsługuje dane wejściowe użytkownika, zapobiegając wstrzyknięciom SQL.

Podsumowanie znalezionych podatności znajduje się w tabeli 2.

Tabela 2. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Gemini

Table 2. Summary of detected vulnerabilities for the application generated by Gemini

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Brak ochrony CSRF	A01:2021 - Broken Access Control	Brak tokenów anty-CSRF w formularzach POST.	Możliwość wymuszenia akcji w imieniu zalogowanego u»ytkownika bez jego wiedzy.
Brak Rate Limitingu	A07:2021 - Identification and Authentication Failures	Brak ograniczeń liczby prób logowania.	Podatność na ataki siłowe (Brute Force) na hasła użytkowników.
User Enumeration	A07:2021 - Identification and Authentication Failures	Komunikat "Taki użytkownik już istnieje!"przy rejestracji.	Umożliwia weryfikację istnienia konkretnych nazw użytkowników w systemie.
Niezabezpieczone ciasteczka	A05:2021 - Security Misconfiguration	Brak wymuszenia flag HttpOnly i Secure.	Zmniejszona odporność na ataki webowe (XSS, Clickjacking).
Brak nagłówków bezpieczeństwa	A05:2021 - Security Misconfiguration	Brak CSP, X-ContentType-Options	Zmniejszona odporność na ataki webowe (XSS, Clickjacking).
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Nagłówki HTTP ujawniają używane oprogramowanie serwera.	Ułatwiony rekonesans dla atakującego.
Brak ochrony przed Clickjackingiem	A05:2021 - Security Misconfiguration	Brak nagłówka X-Frame-Options lub odpowiedniego CSP.	Umożliwia osadzenie aplikacji w ramce (iframe) i wykonanie ataku UI Redressing.

4.3. Claude (Sonnet 4.5) – analiza bezpieczeństwa

Podatności w tym przypadku są bardzo podobne do tych z poprzedniego kodu, dlatego tym razem zostaną wyliczone bez szczegółowego opisu i ze wskazaniem na odpowiednie fragmenty kodu, jeśli jest taka możliwość.

Brak ochrony przed atakami CSRF - formularze nie zawierają tokenów CSRF.

```
<form method="POST" action="{{ url_for('delete_entry', entry_id=entry.id) }}" ...>
    <button type="submit" class="btn btn-danger">Usuń</button>
</form>
```

Listing 11. Podatność związana z brakiem tokenów CSRF

Trwale wpisany klucz sesji

```
app.config['SECRET_KEY'] =
    'twoj-tajny-klucz-zmien-to-w-produkcji'
```

Listing 12. Podatnosć związana z trwale wpisanmym kluczem sesji

Natomiast tutaj AI dodało informację, że klucz powinien zostać zmieniony w środowisku produkcyjnym, co sugeruje świadomość ryzyka związanego z hardcodowaniem sekretów, więc nie będzie to uznane za lukę.

Enumeracja użytkowników:

```
if User.query.filter_by(username=username).first():
    flash('Ta nazwa użytkownika jest już zajęta!', 'error')
    return redirect(url_for('register'))
```

Listing 13. Podatność zwiazana z enumeracją użytkowników

Podsumowanie znalezionych podatności znajduje się w tabeli 3.

Tabela 3. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Claude

Table 3. Summary of detected vulnerabilities for the application generated by Claude

Podatność	**Kategoria OWASP**	**Opis problemu**	**Wpływ i zagrożenia**
Brak ochrony CSRF	A01:2021 - Broken Access Control	Formularze nie zawierają tokenów anty-CSRF.	Możliwość wymuszenia akcji w imieniu zalogowanego użytkownika bez jego wiedzy.
Enumeracja użytkowników	A07:2021 - Identification and Authentication Failures	Komunikat błędu ujawnia, czy nazwa użytkownika jest już zajęta.	Pozwala atakującemu na weryfikację istnienia konkretnych kont w systemie.
Brak nagłówków bezpieczeństwa	A05:2021 - Security Misconfiguration	Brak nagłówków CSP.	Zmniejszona odporność na ataki XSS.
Niezabezpieczone ciasteczka	A05:2021 - Security Misconfiguration	Brak flag HttpOnly i Secure dla ciasteczek sesyjnych.	Ryzyko kradzieży sesji przez skrypty JS (XSS) lub przechwycenia w nieszyfrowanym ruchu.
Brak limitu prób logowania	A07:2021 - Identification and Authentication Failures	Aplikacja nie ogranicza liczby nieudanych prób logowania.	Wysoka podatność na ataki siłowe (Brute Force) na hasła użytkowników.
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Nagłówki HTTP ujawniają używane oprogramowanie serwera.	Ułatwiony rekonesans dla atakującego.
Brak ochrony przed Clickjackingiem	A05:2021 - Security Misconfiguration	Brak nagłówka X-Frame-Options lub odpowiedniego CSP.	Umożliwia osadzenie aplikacji w ramce (iframe) i wykonanie ataku UI Redressing.

4.4. Deepseek (V 3.2) – analiza bezpieczeństwa

Podatności w tym przypadku są bardzo podobne do tych z poprzedniego kodu, dlatego tym razem zostaną wyliczone bez szczegółowego opisu i ze wskazaniem na odpowiednie fragmenty kodu, jeśli jest taka możliwość.

Brak ochrony przed atakami CSRF - formularze nie zawierają tokenów CSRF.

```
<form action="{{ url_for('delete_post', post_id=post.id) }}" method="POST" ...>
   <button type="submit" ...>Usuń</button>
</form>
```

Listing 14. Podatnosć związana z brakiem tokenów CSRF

Trwale wpisany klucz sesji

```
app.config['SECRET_KEY'] = 'twoj-tajny-klucz-tutaj'
# Zmień to w środowisku produkcyjnym!
```

Listing 15. Podatnosć związana z trwale wpisanym kluczem sesji

Natomiast tutaj AI dodało informację, że klucz powinien zostać zmieniony w środowisku produkcyjnym, co sugeruje świadomość ryzyka związanego z hardcodowaniem sekretów, więc nie będzie to uznane za lukę.

Enumeracja użytkowników:

```
@app.route('/register', methods=['GET', 'POST'])
def register():
   # ...
   existing_user = User.query.filter_by(username=username).first()
   if existing_user:
      flash('Nazwa użytkownika jest już zajęta!', 'danger')
```

Listing 16. Podatność związana z enumeracją użytkowników

Podsumowanie znalezionych podatności znajduje się w tabeli 4.

Tabela 4. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Deepseek

Table 4. Summary of detected vulnerabilities for the application generated by Deepseek

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Brak ochrony CSRF	A01:2021 - Broken Access Control	Formularze nie zawierają tokenów anty-CSRF.	Możliwość wymuszenia akcji w imieniu zalogowanego użytkownika bez jego wiedzy.
Enumeracja użytkowników	A07:2021 - Identification and Authentication Failures	Komunikat błędu ujawnia, czy nazwa użytkownika jest już zajęta.	Pozwala atakującemu na weryfikację istnienia konkretnych kont w systemie.
Brak nagłówków bezpieczeństwa	A05:2021 - Security Misconfiguration	Brak nagłówków CSP.	Zmniejszona odporność na ataki XSS.
Niezabezpieczone ciasteczka	A05:2021 - Security Misconfiguration	Brak flag HttpOnly i Secure dla ciasteczek sesyjnych.	Ryzyko kradzieży sesji przez skrypty JS (XSS) lub przechwycenia w nieszyfrowanym ruchu.
Brak limitu prób logowania	A07:2021 - Identification and Authentication Failures	Aplikacja nie ogranicza liczby nieudanych prób logowania.	Wysoka podatność na ataki siłowe (Brute Force) na hasła użytkowników.
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Nagłówki HTTP ujawniają używane oprogramowanie serwera.	Ułatwiony rekonesans dla atakującego.
Brak ochrony przed Clickjackingiem	A05:2021 - Security Misconfiguration	Brak nagłówka X-Frame-Options lub odpowiedniego CSP.	Umożliwia osadzenie aplikacji w ramce (iframe) i wykonanie ataku UI Redressing.

4.5. Porównanie wyników

Analiza bezpieczeństwa przeprowadzona dla czterech modeli pozwoliła na wyodrębnienie wyraźnych różnic w podejściu do generowania kodu przez sztuczną inteligencję, gdy w prompcie nie zawarto wymagań dotyczących bezpieczeństwa. W szczególności wyróżnił się model ChatGPT-5, który wykazał się największą liczbą luk bezpieczeństwa w wygenerowanym kodzie. Model ten wygenerował kod oparty na zapytaniach SQL oraz niebezpiecznej funkcji renderowania, co skutkowało powstaniem krytycznych podatności, które nie wystąpiły w kodzie wygenerowanym przez pozostałe modele:

1. **Server-Side Template Injection (SSTI):** Podatność ta wystąpiła wyłącznie w kodzie wygenerowanym przez ChatGPT. Wynikała ona z użycia f-stringów do wstrzykiwania danych użytkownika bezpośrednio do funkcji render_template_string, co umożliwia zdalne wykonanie kodu.

2. **Przechowywanie haseł jawnym tekstem:** ChatGPT jako jedyny zapisał hasła w bazie danych w postaci otwartego tekstu. Pozostałe modele domyślnie użyły bibliotek do bezpiecznego hashowania haseł, mimo braku takiej instrukcji w prompcie.
3. **Brak sanityzacji (XSS):** Kod ChatGPT był podatny na Stored XSS, ponieważ wklejał treść wpisów bezpośrednio do HTML. Inne modele wykorzystały mechanizmy Object-Relational Mapping i szablonów, które zapewniły podstawową ochronę.

Reszta podatności powtarzała się we wszystkich wygenerowanych kodach, co sugeruje, że bez wyraźnych instrukcji dotyczących bezpieczeństwa modele AI mają trudności z automatycznym uwzględnieniem najlepszych praktyk bezpieczeństwa w generowanym kodzie:

1. **Brak ochrony przed atakami CSRF:** Żaden z modeli nie zaimplementował tokenów anty-CSRF. Wszystkie formularze były podatne na wymuszenie żądania w imieniu zalogowanego użytkownika.
2. **Trwale wpisany klucz sesji:** We wszystkich kodach klucz SECRET_KEY został wpisany na sztywno. Natomiast oprócz ChatGPT-5, pozostałe modele dodały komentarz sugerujący zmianę klucza w środowisku produkcyjnym.
3. **Enumeracja użytkowników:** Każda z aplikacji posiadała mechanizm rejestracji, który poprzez komunikaty błędów ujawniał istnienie kont w systemie, ułatwiając rekonesans.
4. **Brak nagłówków bezpieczeństwa**.
5. **Niezabezpieczone ciasteczka sesyjne.**
6. **Brak limitu prób logowania.**
7. **Ujawnienie wersji serwera w nagłówkach HTTP.**

Modele Gemini, Claude i Deepseek wykazały się lepszym poziomem bezpieczeństwa na poziomie technologicznym poprzez użycie ORM i hashowania, eliminując błędy krytyczne klasy Injection czy Cryptographic Failures. ChatGPT wygenerował kod działający, ale fundamentalnie niebezpieczny. Żaden z modeli nie zadbał jednak samodzielnie o bezpieczeństwo logiki biznesowej jak CSRF lub Rate Limiting bez wyraźnego polecenia w prompcie.

5. Drugi prompt - poprawa luk w zabezpieczeniach

Drugi prompt ma na celu wskazanie na problemy z bezpieczeństwiem na stronie, ale nie podajemy konkretnych rozwiązań. Chcemy zobaczyć jak dobrze AI samo poradzi sobie z poprawą bezpieczeństwa na podstawie swojej analizy, czyli czy jest w stanie zidentyfikować luki i zaproponować odpowiednie poprawki bez podpowiedzi. Drugi prompt zostanie dodany na tej samej konwersacji co pierwszy, aby AI miało kontekst wygenerowanego kodu.

```
Mam prostą aplikację w Pythonie (Flask), którą napisałeś dla mnie wcześniej.
Oto mój obecny kod: [KOD]
Ten kod ma luki w zabezpieczeniach. Zanalizuj go i przepisz tak, aby był bezpieczny
i gotowy do wrzucenia na serwer. Chcę dostać cały poprawiony kod.
```

Treść drugiego promptu

5.1. ChatGPT (GPT-5) – analiza bezpieczeństwa

Pierwsze co można zauważyć to znaczna poprawa bezpieczeństwa w wygenerowanym kodzie w porównaniu do poprzedniej wersji. Podatności krytyczne takie jak zapisywanie haseł jawnym tekstem:

```
hashed = generate_password_hash(password)
check_password_hash(row["password"], password)
```

Listing 17. Hashowanie haseł

czy Server-Side Template Injection:

```
<input type="text" name="search" placeholder="Szukaj po tytule..." value="
{{ search|e }}">
response.headers.setdefault("X-Content-Type-Options", "nosniff")
```

Listing 18. Zabezpieczenie przed Server-Side Template Injection

zostały całkowicie wyeliminowane.

Oprócz tego dodano ochronę przed CSRF poprzez generowanie unikalnego tokenu kryptograficznego, który jest dołączany do formularzy i weryfikowany przy każdym żądaniu POST. W ten sposób ataki CSRF są skutecznie blokowane.

```
def validate_csrf_token(form_token):
    token = session.get("_csrf_token")
    if not token or not form_token or not secrets.compare_digest(token, form_token):
        return False
    return True
<input type="hidden" name="csrf_token" value="{{ csrf_token }}">
if request.method == "POST":
    form_csrf = request.form.get("csrf_token", "")
    if not validate_csrf_token(form_csrf):
        abort(400, description="CSRF token invalid")
```

Listing 19. Zabezpieczenie przed CSRF

Dodatkowo wprowadzono zabezpieczenia sesji poprzez ustawienie flag HttpOnly i SameSite dla ciasteczek sesyjnych, co ogranicza ryzyko kradzieży sesji przez skrypty lub ataki CSRF.

```
app.config.update(
    SESSION_COOKIE_HTTPONLY=True,
    SESSION_COOKIE_SAMESITE="Lax",
)
```

Listing 20. Poprawione zabezpieczenie sesji

Klucz sekretu sesji nie jest już na stałe wpisany w kodzie, lecz generowany losowo lub pobierany z zmiennej środowiskowej, co zwiększa bezpieczeństwo aplikacji.

```
SECRET_KEY = os.environ.get("SECRET_KEY") or secrets.token_urlsafe(32)
app.secret_key = SECRET_KEY
```

Listing 21. Losowe generowanie klucza sekretu sesji

Dodano ochronę przed Clickjackingiem poprzez ustawienie nagłówka X-Frame-Options na DENY, co zapobiega osadzaniu aplikacji w ramkach innych stron.

```
@app.after_request
def set_security_headers(response):
    response.headers.setdefault("X-Frame-Options", "DENY")
    return response
```

Listing 22. Ochrona przed Clickjackingiem

Zostały jednak nadal pewne luki bezpieczeństwa, które warto wymienić, choć ich liczba i krytyczność zostały znacznie zredukowane. Jedną z nich jest nadal możliwość enumeracji użytkowników podczas rejestracji:

```
except sqlite3.IntegrityError:
    flash("Ta nazwa użytkownika jest już zajęta.")
```

Listing 23. Podatność związana z enumeracją użytkowników

Wciąż istnieje brak limitu prób logowania, co nadal naraża aplikację na ataki brute-force.

Nadal ujawniana jest wersja serwera w nagłówkach HTTP, co może ułatwić atakującym rekonesans.

W tym kodzie też brak nagłówków bezpieczeństwa takich jak Content Security Policy (CSP), które mogłyby dodatkowo chronić przed atakami XSS. Jest to jednak świadomy wybór, ponieważ szablon używa stylów inline, co utrudnia zastosowanie ścisłej polityki CSP bez refaktoryzacji kodu front-endowego. Zostało to nawet wspomniane w komentarzu w kodzie:

```
# Content-Security-Policy is omitted because template uses inline styles;
# if you refactor to external CSS, add a strict CSP.
```

Listing 24. Podatność związana z brakiem nagłówków bezpieczeństwa takich jak Content Security Policy

Podsumowanie znalezionych podatności znajduje się w tabeli 5.

Tabela 5. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez ChatGPT

Table 5. Summary of detected vulnerabilities for the application generated by ChatGPT

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Enumeracja użytkowników	A07:2021 - Identification and Authentication Failures	Komunikat "Ta nazwa u»ytkownika jest już zajęta" przy rejestracji.	Pozwala atakującemu na weryfikację istnienia konkretnych kont w systemie.
Brak Rate Limitingu	A07:2021 - Identification and Authentication Failures	Brak mechanizmów ograniczających liczbę prób logowania.	Podatność na ataki siłowe (Brute Force) na hasła użytkowników.
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Nagłówki HTTP ujawniają wersję używanego oprogramowania.	Ułatwiony rekonesans dla atakującego (dobór exploitów pod wersję).
Brak Content Security Policy (CSP)	A05:2021 - Security Misconfiguration	Brak nagłówka CSP (świadomy wybór ze względu na użycie stylów inline).	Zmniejszona odporność na ataki XSS (brak dodatkowej warstwy ochrony w przeglądarce).

5.2. Gemini (3 Pro) – analiza bezpieczeństwa

W przeciwieństwie do ChataGPT, Gemini wykorzystuje framework Flask-WTF oraz WTForms do obsługi formularzy, co automatycznie wprowadza ochronę przed CSRF oraz walidację danych wejściowych. Dzięki temu ryzyko ataków CSRF jest skutecznie zminimalizowane, a dane wejściowe są sprawdzane pod kątem poprawności i długości.

```
class LoginForm(FlaskForm):
   username = StringField('Login', validators = [...])
```

Listing 25. Ochrona przed atakami CSRF

Kod wymusza bezpieczne flagi dla ciasteczek, co utrudnia ich kradzież i wykorzystanie w innych atakach.

```
app.config['SESSION_COOKIE_HTTPONLY'] = True
app.config['SESSION_COOKIE_SAMESITE'] = 'Lax'
```

Listing 26. Korzystanie z flag bezpiecznych dla plików cookies

Nie oznacza to jednak, że kod jest wolny od wad. Podobnie jak w przypadku ChataGPT, nadal istnieje możliwość enumeracji użytkowników podczas rejestracji:

```
class RegisterForm(FlaskForm):
   # ...
   def validate_username(self, username):
      user = User.query.filter_by(username=username.data).first()
      if user:
         raise ValidationError('Taki użytkownik już istnieje. Wybierz inny login.')
```

Listing 27. Podatność związana z enumeracją użytkowników

Brak jest również mechanizmów ograniczających liczbę prób logowania, co naraża aplikację na ataki brute-force.

Dodatkowo, podobnie jak w poprzednim przypadku, nie zastosowano nagłówków bezpieczeństwa takich jak Content Security Policy (CSP), które mogłyby dodatkowo chronić przed atakami XSS, brakuje sekcji dodającej nagłówki HTTP dla każdej odpowiedzi. Wygenerowany kod również nie ukrywa wersji serwera w nagłówkach HTTP, co może ułatwić atakującym rekonesans.

Podsumowanie znalezionych podatności znajduje się w tabeli 6.

Tabela 6. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Gemini

Table 6. Summary of detected vulnerabilities for the application generated by Gemini

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Enumeracja użytkowników	A07:2021 - Identification and Authentication Failures	Komunikat "Ta nazwa u»ytkownika jest już zajęta" przy rejestracji.	Pozwala atakującemu na weryfikację istnienia konkretnych kont w systemie.
Brak Rate Limitingu	A07:2021 - Identification and Authentication Failures	Brak mechanizmów ograniczających liczbę prób logowania.	Podatność na ataki siłowe (Brute Force) na hasła użytkowników.
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Nagłówki HTTP ujawniają wersję używanego oprogramowania.	Ułatwiony rekonesans dla atakującego (dobór exploitów pod wersję).
Brak nagłówków bezpieczeństwa (CSP)	A05:2021 - Security Misconfiguration	Brak Content Security Policy oraz sekcji dodaj¡cej nagłówki dla każdej odpowiedzi.	Zmniejszona odporność na ataki XSS (brak dodatkowej warstwy ochrony).

5.3. Claude (Sonnet 4.5) – analiza bezpieczeństwa

Tym razem wygenerowany kod nie posiada tokenów zebezpieczających przed atakiem Cross-Site Request Forgery:

```
<form method="POST" autocomplete="on">
  <div class="form-group">
    </div>
  <button type="submit" ...>Zaloguj</button>
</form>
```

Listing 28. Brak zabezpieczeń przed atakiem Cross-Site Request Forgery

Implementacja nagłówków bezpieczeństwa są ustawiane za pomocą tagów meta w szablonie bazowym. Natomiast X-Frame-Options oraz X-Content-Type-Options musi być wysłany jako nagłówek HTTP przez serwer, a nie jako meta tag w HTML, więc ta implementacja nie jest skuteczna.

```
<head>
  <meta http-equiv="X-Content-Type-Options" content="nosniff">
  <meta http-equiv="X-Frame-Options" content="DENY">
  <meta http-equiv="Content-Security-Policy" content="...">
</head>
```

Listing 29. Nieskuteczna implementacja nagłówków bezpieczeństwa

Dodatkowo można zauważyć problemy z Content Security Policy, ponieważ zezwala ona na 'unsafe-inline' dla stylów, co może być potencjalnym wektorem ataku XSS.

```
<meta http-equiv="Content-Security-Policy" content="default-src 'self';
style-src 'unsafe-inline'; script-src 'none';">
```

Listing 30. Podatność na ataki XSS

Tak jak w poprzednich przypadkach można zauważyć możliwość enumeracji użytkowników podczas rejestracji:

```
if User.query.filter_by(username=username).first():
    flash('Ta nazwa użytkownika jest już zajęta!', 'error')
    return redirect(url_for('register'))
```

Listing 31. Podatność zwiazana z enumeracją użytkowników

Oraz brak mechanizmów ograniczających liczbę prób logowania, co naraża aplikację na ataki brute-force.

Można też zauważyć, że podobnie jak w poprzednich przypadkach, nie ukryto wersji serwera w nagłówkach HTTP, co może ułatwić atakującym rekonesans.

Podsumowanie znalezionych podatności znajduje się w tabeli 7.

Tabela 7. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Claude

Table 7. Summary of detected vulnerabilities for the application generated by Claude

Podatność	**Kategoria OWASP**	**Opis problemu**	**Wpływ i zagrożenia**
Brak tokenów Anti-CSRF	A01:2021 - Broken Access Control	Formularze nie zawierają unikalnych tokenów weryfikujących intencję żądania.	Możliwość wymuszenia na zalogowanym użytkowniku niechcianych akcji (np. dodanie wpisu, zmiana hasła).
Niewłaściwa implementacja nagłówków	A05:2021 - Security Misconfiguration	Nagłówki X-Frame-Options i inne zdefiniowane jako tagi zamiast w odpowiedzi HTTP.	Przeglądarki ignorują te ustawienia w sekcji meta, co wyłącza ochronę przed Clickjackingiem.
Słaba konfiguracja CSP	A05:2021 - Security Misconfiguration	Polityka zezwala na style 'unsafe-inline' i jest zdefiniowana w tagu <meta>.	Zwiększone ryzyko skutecznego ataku XSS poprzez wstrzyknięcie złośliwego kodu CSS/JS.

Tabela 7 (cd.). Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Claude

Table 7 (cont.). Summary of detected vulnerabilities for the application generated by Claude

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Enumeracja użytkowników	A07:2021 - Identification and Authentication Failures	Komunikat błędu przy rejestracji wprost informuje, czy nazwa użytkownika jest zajęta.	Umożliwia stworzenie listy prawidłowych loginów w celu przeprowadzenia ukierunkowanych ataków.
Brak Rate Limitingu	A07:2021 - Identification and Authentication Failures	Brak ograniczeń liczby żądań do endpointów logowania i rejestracji.	Narażenie na ataki siłowe (Brute Force) oraz Credential Stuffng.
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Serwer zwraca nagłówek Server: Werkzeug/x.x.x, ujawniając technologię.	Ułatwia atakującemu dobór exploitów znanych dla konkretnej wersji oprogramowania.

5.4. Deepseek (V 3.2) – analiza bezpieczeństwa

W tym kodzie wciąż brak nagłówków bezpieczeństwa oraz można też zobaczyć informację o wersji serwera w nagłówku HTTP.

Formularz rejestracji informuje czy nazwa użytkownika jest zajęta, co umożliwia atakującemu stworzenie listy prawidłowych loginów.

```
@app.route('/register', methods=['GET', 'POST'])
# ...
      else:
         existing_user = User.query.filter_by(username=username).first()
         if existing_user:
            flash('Nazwa użytkownika jest już zajęta!', 'danger')
```

Listing 32. Podatność związana z ujawnianiem istniejacych nazw użytkowników

Ciasteczka nie mają wymuszonej flagi Secure, co w tym przypadku jest podjętą decyzją, na co wskazuje komentarz w kodzie. Jednak w środowisku produkcyjnym powinno to być ustawione na True.

```
app.config.update(
   # ...
   SESSION_COOKIE_SECURE=False,  # Ustaw na True w produkcji z HTTPS
   SESSION_COOKIE_HTTPONLY=True,
   SESSION_COOKIE_SAMESITE='Lax',
   # ...
)
```

Listing 33. Brak wymuszonej flagi Secure

Podsumowanie znalezionych podatności znajduje się w tabeli 8.

Tabela 8. Podsumowanie wykrytych podatności dla aplikacji wygenerowanej przez Deepseek

Table 8. Summary of detected vulnerabilities for the application generated by Deepseek

Podatność	Kategoria OWASP	Opis problemu	Wpływ i zagrożenia
Brak nagłówków bezpieczeństwa	A05:2021 - Security Misconfiguration	Brak implementacji nagłówków takich jak CSP, X-Frame-Options czy X-Content-TypeOptions w kodzie aplikacji.	Podatność na ataki typu Cross-Site Scripting (XSS), Clickjacking oraz MIME-sniffing.
Ujawnienie wersji serwera	A05:2021 - Security Misconfiguration	Serwer wysyła informacje o swojej wersji w nagłówku HTTP (brak mechanizmu ukrywania).	Ułatwia atakującemu rekonesans i dobranie exploitów pod konkretną wersję oprogramowania.
Enumeracja u»ytkowników	A07:2021 - Identification and Authentication Failures	Aplikacja zwraca komunikat: „Nazwa użytkownika jest już zajęta!”, gdy login istnieje w bazie.	Umożliwia atakującemu weryfikację, które loginy są zarejestrowane w systemie (budowanie listy celów).
Brak flagi Secure dla ciasteczek	A05:2021 - Security Misconfiguration	Konfiguracja SESSION_COOKIE_SECURE = False (zgodnie z komentarzem w kodzie).	Ryzyko przechwycenia ciasteczka sesyjnego i przejęcia konta (Session Hijacking), jeśli ruch nie jest szyfrowany (brak wymuszenia HTTPS).

5.5. Podsumowanie wyników

W drugim teście, gdzie kod był bezpieczniejszy, wszystkie modele wykazały się bardzo wysoką zgodnością w detekcji błędów konfiguracyjnych i logicznych. W przeciwieństwie do pierwszego testu, tutaj różnice w wynikach były minimalne. Każdy z modeli ujawnia wersję serwera w nagłówku HTTP. Wykryto również problem z komunikatem o zajętości nazwy użytkownika, co umożliwia atakującemu enumerację użytkowników. Modele radzą sobie dobrze z wykrywaniem problemów bezpieczeństwa, jednak mają problem z nagłówkami bezpieczeństwa, co jest też trochę wymuszone przez prompt pierwszy, który nakazuje zrobić wszystko w jednym pliku, co utrudnia implementację nagłówków na poziomie serwera.

Tym razem ChatGPT dorównał do wyników reszty modeli, mając podobną ilość luk bezpieczeństwa co reszta. Gorzej wypadł Claude, który nie zaimplementował tokenów CSRF, co jest poważnym błędem bezpieczeństwa w aplikacjach webowych. Implementacja nagłówków bezpieczeństwa też poszła mu gorzej, ze względu na wykorzystanie tagów meta w HTML zamiast nagłówków HTTP.

6. Trzeci prompt, ekspert od bezpieczeństwa aplikacji

Ten prompt jest modyfikacją drugiego promptu, gdzie dodano rolę eksperta od bezpieczeństwa aplikacji i standardów OWASP oraz zadanie refaktoryzacji kodu w celu usunięcia wszystkich problemów bezpieczeństwa.

```
Role: Działaj jako Senior Application Security Engineer i ekspert od standardów OWASP.
Kontekst: Poniżej wklejam kod prostej aplikacji webowej napisanej we Flasku (Python).
Zadanie: Twoim celem jest przeprowadzenie refaktoryzacji tego kodu, aby stał się bezpieczny i odporny na ataki
Wymagania:
1. Struktura: Możesz rozdzielić projekt na inne pliki jeśli jest to wymagane.
2. Kod: Wypisz cały, kompletny kod po poprawkach (nie tylko fragmenty).
Mój obecny kod: [KOD]
```

Treść trzeciego promptu

6.1. ChatGPT (GPT-5) – analiza bezpieczeństwa

Polityka haseł jest, ale bardzo podstawowa - tylko długość minimum 8 znaków. Można by dodać bardziej zaawansowane reguły, aby wymusić silniejsze hasła.

Ze wszystkiego zostały tylko dwie luki bezpieczeństwa, które nie mają dużej krytyczności. Jedną z nich jest wciąż ujawnianie wersji serwera w nagłówku HTTP. Drugą z nich jest brak dyrektywy form-action w nagłówku CSP. Teoretycznie pozwalałoby to na wysłanie formularza na zewnętrzną domenę, ale skoro atakujący nie może wstrzyknąć własnego formularza (brak XSS), ryzyko jest minimalne.

6.2. Gemini (3 Pro) – analiza bezpieczeństwa

Polityka haseł została zaostrzona do minimum 12 znaków, ale wciąż brakuje bardziej zaawansowanych reguł (wielkie litery, cyfry, symbole).

Ten kod ma dokładnie takie same podatności jak poprzedni. Brak dyrektywy form-action w CSP oraz ujawnianie wersji serwera w nagłówku HTTP. Reszta zabezpieczeń jest na dobrym poziomie.

6.3. Claude (Sonnet 4.5) – analiza bezpieczeństwa

Polityka haseł jest znacznie bardziej zaawansowana - minimum 12 znaków oraz wymóg spełnienia co najmniej trzech z czterech kategorii znaków (wielkie litery, małe litery, cyfry, znaki specjalne). Dodatkowo sprawdzane są popularne hasła.

Tak samo jak poprzednie wygenerowane kody, w tym też wycieka informacja o wersji serwera w nagłówku HTTP. Dodatkowo polityka bezpieczeństwa pozwala na style wpisane bezpośrednio w HTML (inline styles), co nie jest zalecane, ale jest konieczne, ponieważ szablony HTML generowane są dynamicznie i nie ma zewnętrznego pliku CSS. Natomiast dzięki używaniu biblioteki Bleach do oczyszczania danych wejściowych użytkownika, ryzyko XSS jest zminimalizowane.

System aktywnie zwraca HTTP/1.1 429 TOO MANY REQUESTS w przypadku przekroczenia limitów żądań, co jest dobrą praktyką zabezpieczającą przed atakami siłowymi.

6.4. Deepseek (V3.2) – analiza bezpieczeństwa

Polityka haseł jest podobna do poprzedniego kodu - minimum 12 znaków oraz wymóg spełnienia co najmniej trzech z czterech kategorii znaków (wielkie litery, małe litery, cyfry, znaki specjalne).

Główną podatnością w tym przypadku są brakujące nagłówki bezpieczeństwa w odpowiedziach HTTP. Tak jak na poprzednich stronach wycieka informacja o serwerze.

7. Czwarty prompt, wskazanie na problemy bezpieczeństwa

Celem tego promptu jest wygenerowanie kodu, który będzie bezpieczny i będzie unikał problemów z poprzednich wersji. Wskazujemy na konkretne problemy bezpieczeństwa, które wystąpiły w poprzednich wersjach i prosimy o ich naprawę, oraz aby nie zmieniać innych askpektów jeśli nie ma takiej potrzeby.

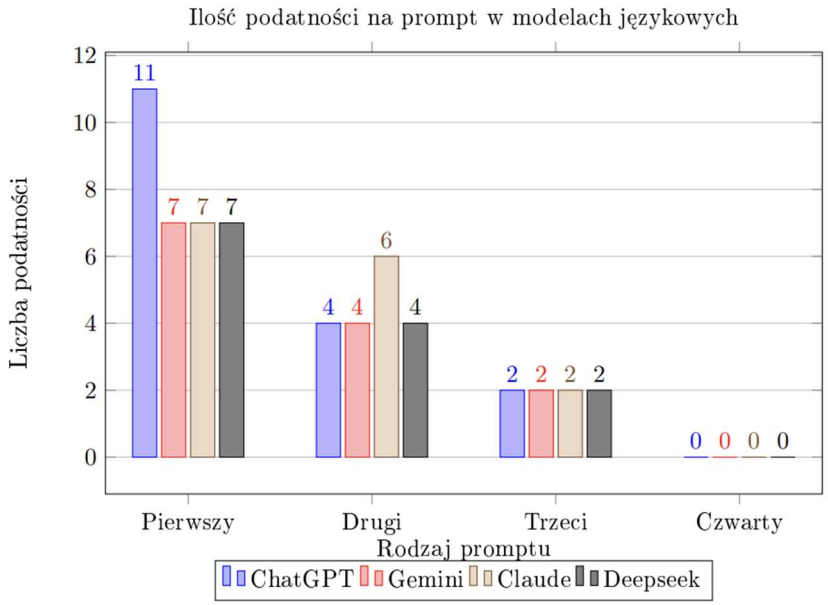
Mam aplikację w Pythonie (Flask), którą napisałeś dla mnie wcześniej.
Oto mój obecny kod: [KOD]
Wykryto w nim podane podatności:
|| Podatności, które zostały wykryte w ostatniej iteracji ||
Zanalizuj kod i przepisz tak, aby wziął pod uwagę zabezpieczenia dla tych podatności, jeśli nie ma takiej potrzeby to postaraj się nie zmieniać innych rzeczy. Kod ma być gotowy do wrzucenia na serwer. Chcę dostać cały poprawiony kod (nie fragmenty).

Treść czwartego promptu

Po użyciu tego promptu, kod, który został wygenerowany ponownie, nie różni się bardzo od wcześniejszej iteracji i udało się pozbyć wszystkich wykrytych wcześniej podatności.

7.1. Podsumowanie

Przebieg ilości podatności można zauważyć na poniższym wykresie.



Rys. 1. Podsumowanie ilości podatności po każdym prompcie

Fig. 1. Summary of vulnerability counts after each prompt

Wycieku informacji o serwerze w nagłówkach HTTP nie da się pozbyć przez to, że jest to domyślnie nadpisywane. Można za to manualnie sprawdzić czy na nagłówek serwera jest coś wpisywane, żeby uznać, że tej podatności już nie ma. Dodatkowo udało się dodać nagłówki bezpieczeństwa do odpowiedzi HTTP, co znacznie podnosi bezpieczeństwo aplikacji. Skoro żadna ze stron nie ma już wykrytych żadnych ważnych podatności po takiej samej liczbie iteracji to w jaki sposób określić, która z nich jest najlepsza pod kątem bezpieczeństwa? Można to zrobić przez porównanie kodu źródłowego i sprawdzenie, która z nich ma lepsze praktyki bezpieczeństwa w kodzie.

Między modelami można zauważyć różnice w polityce haseł. ChatGPT ma najprostszą, ponieważ wymaga jedynie 8 znaków w haśle i nawet nie ma pola na wpisanie powtórzonego hasła, więc nie można go uznać za najlepszy pod kątem bezpieczeństwa. Gemini wymaga już 12 znaków, ale dalej jest to za mało. Widać większą różnicę w podejściu Claude, który nie tylko wymaga 12 znaków, ale także wymaga spełnienia co najmniej trzech z czterech kategorii znaków (wielkie litery, małe litery, cyfry, znaki specjalne). Dodatkowo ma warunki w nazwie użytkownika, mają one mieć od 3 do 50 znaków i mogą zawierać tylko litery, cyfry i znaki _ oraz -.

Deepseek ma podobną politykę do Claude, też wymaga od 12 znaków w haśle, ale wymaga też spełnienia wszystkich czterech kategorii znaków, ale podaje tylko określone znaki specjalne z których można korzystać. Nazwa użytkownika może mieć od 3 do 20 znaków i może zawierać tylko litery, cyfry i podkreślniki.

Na pierwszy rzut oka można by uznać, że to Deepseek zaplementował najlepszą politykę haseł, ponieważ wymaga spełnienia wszystkich czterech kategorii. Natomiast dla użytkownika może to być uciążliwe, zwłaszcza z połączeniem tylko niektórych z znaków specjalnych. Takie rozwiązanie może doprowadzać manadżery haseł do generowania haseł, których nie można użyć. Dodatkowo Claude zaimplementował więcej znaków dla nazwy użytkownika, co pozwala na większą elastyczność i komfort.

8. Wnioski

Analiza bezpieczeństwa aplikacji webowych wygenerowanych przez modele LLM (ChatGPT-5, Gemini 3 Pro, Claude Sonnet 4.5, Deepseek V3.2) pozwala na sformułowanie kluczowych spostrzeżeń dotyczących zagrożeń związanych z programowaniem wspomaganym przez AI oraz skuteczności inżynierii promptów.

Generowanie kodu przy użyciu prostych promptów, nastawionych wyłącznie na funkcjonalność, czyli żeby działało, niesie ze sobą krytyczne ryzyko bezpieczeństwa. Modele, bez otrzymania wyraźnych instrukcji dotyczących zabezpieczeń, priorytetyzują poprawność składniową i wizualną kosztem bezpieczeństwa. ChatGPT w pierwszej iteracji wygenerował kod fundamentalnie niebezpieczny, zawierający podatności krytyczne, takie jak Server-Side Template Injection, czy przechowywanie haseł jawnym tekstem. Pozostałe modele wykazały wyższy poziom bezpieczeństwa, domyślnie stosując mechanizmy ORM oraz hashowanie haseł. Wspólnym mianownikiem dla wszystkich modeli w pierwszej fazie był brak zabezpieczeń warstwy logicznej i konfiguracyjnej, takie jak brak ochrony przed CSRF, brak nagłówków bezpieczeństwa, brak limitów prób logowania oraz enumeracja użytkowników dająca informacje, czy użytkownik jest w bazie.

Zastosowanie techniki iteracyjnego ulepszania promptów od ogólnego, przez wskazanie błędów, po wcielenie się w rolę eksperta ds. bezpieczeństwa przyniosło lepsze rezultaty. Wprowadzenie promptu "eksperta", czyli wcielenie się w rolę Senior Application Security Engineer, pozwoliło zredukować liczbę wykrywanych podatności do zera lub wartości minimalnych we wszystkich badanych modelach. Modele wykazały zdolność do autokorekty i potrafiły zidentyfikować wiele luk jak np. brak tokenów anty-CSRF i zaimplementować odpowiednie biblioteki po otrzymaniu odpowiedniego kontekstu.

Modele prezentują odmienne podejście do balansu między bezpieczeństwem a użytecznością, co jest szczególnie widoczne w implementacji polityk haseł. Deepseek implementuje najbardziej restrykcyjną politykę (wymóg 4 kategorii znaków, specyficzne znaki specjalne), co z perspektywy technicznej jest bezpieczne,

ale może prowadzić do problemów z użytecznością i niekompatybilności z niektórymi menedżerami haseł. Claude oferuje najbardziej zrównoważone podejście, zapewniając wysoki poziom bezpieczeństwa przy zachowaniu elastyczności dla użytkowników (większa liczba znaków w nazwie użytkownika, spełnienie 3 z 4 kategorii znaków w haśle).

Eksperyment potwierdza, że AI nie zastępuje kompetentnego programisty, a używanie metody "kopiuj-wklej", czyli skopiowanie całego kodu bez sprawdzenia jego poprawności, jest niebezpieczny. O ile AI radzi sobie z błędami technicznymi jak wstrzyknięcia, o tyle błędy logiki jak np. enumeracja użytkowników poprzez komunikaty błędów, brak Rate Limitingu często pozostają nienaprawione bez odpowiednio przygotowanego promptu.

Podsumowując, wykorzystanie LLM w procesie wytwarzania oprogramowania znacząco przyspiesza pracę, jednak bez weryfikacji kodu przez osobę posiadającą wiedzę z zakresu cyberbezpieczeństwa, prowadzi do powstawania aplikacji podatnych na ataki. AI powinno być traktowane jako asystent wykonawczy, a nie architekt bezpieczeństwa.

Literatura

[1] CyberFolks „Czym jest OWASP”? URL: https : / /cyberfolks.pl/slownik/owasp/ Dostęp: 26 października 2025.

[2] dr inż. Mirosław Mazurek. „Eksploatacja i bezpieczeństwo systemów.” Dostęp: 18 listopada 2025. URL: https://e-learning.prz.edu.pl/pluginfile.php/95179/mod_resource/content/1/wyklad_02_OWASP.pdf

[3] OWASP. OWASP Top Ten 2025. Dostęp: 26 października 2025. URL: https://owasp.org/www-project-top-ten/

[4] SAP. Na czym polega projektowanie aplikacji z AI? Dostęp: 30 października 2025. URL: https://www.sap.com/poland/resources/what-is-ai-application-development

THE SECURITY OF AI-GENERATED WEBSITES

Summary

The aim of this work is to determine the security level of websites generated using several large language models, and then analyze the impact of prompt content and level of detail on the quality of the resulting security. OWASP ZAP, a free, open-source web application security testing tool, was used in the verification process. The prompts used for code generation were constructed iteratively, gradually increasing the complexity and precision of requirements regarding both the functionality and security aspects of the application. The approach ranged from a typical user's perspective to a more advanced level of security. Depending on the level of detail of the prompts and the number of corrections implemented, applications with varying levels of security could be generated – from sites susceptible to basic threats such as lack of input validation, susceptibility to XSS attacks, or incorrect HTTP header configuration, to solutions that meet key security standard recommendations. We analyzed the impact of prompt formulation on eliminating vulnerabilities consistent with the OWASP Top 10 guidelines. The project assumes a multi-stage research approach, in which the initial version of the application

does not incorporate secure programming principles, and then undergoes subsequent iterations of code generation with increasing emphasis on eliminating detected vulnerabilities. This approach allows us to assess the degree to which language models are able to respond to identified security issues and whether it is possible to generate an application that meets basic security requirements using these models.

Keywords: vulnerabilities, language model, code generation.

Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Krystian Sołtys, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST

dr hab. inż. Tomasz Kapuściński, prof. PRz, dr inż. Dawid Kalandyk
opiekunowie Koła Naukowego

UNIWERSALNY INTERFEJS KOMUNIKACJI MIĘDZY KOMPUTEREM A WIELOMA PROJEKTAMI BAZUJĄCYMI NA RASPBERRY PI

Streszczenie

W rozdziale przedstawiono projekt i implementację uniwersalnego interfejsu komunikacyjnego do zdalnego przesyłania oraz sterowania plikami symulacyjnymi uruchamianymi na wielu urządzeniach Raspberry Pi. Celem było stworzenie jednolitego, łatwego w konfiguracji rozwiązania, szczególnie przydatnego w zastosowaniach dydaktycznych i prototypowych. Architekturę oparto na modelu klient–serwer z wykorzystaniem protokołu WebSocket, zapewniającego dwukierunkową komunikację w czasie rzeczywistym. Centralny serwer w języku C++ (z użyciem bibliotek Boost i Boost.Beast) zarządza klientami działającymi na urządzeniach docelowych. Komunikacja między programem zarządzającym a symulacją odbywa się przez potoki nazwane, bez ingerencji w logikę aplikacji. Wymiana danych opiera się na formacie JSON i obejmuje komendy sterujące, transmisję plików, komunikację z aplikacją oraz przesył obrazu. System rozszerzono o mechanizm logowania i centralnego monitoringu.

Słowa kluczowe: Raspberry Pi, WebSocket, C++, Python, interfejs komunikacji.

1. Wprowadzenie

Mimo rosnącej popularności [1] platformy Raspberry Pi jako niskokosztowego i energooszczędnego rozwiązania wykorzystywanego w edukacji, prototypowaniu oraz systemach wbudowanych, nadal brakuje prostych i zunifikowanych mechanizmów umożliwiających łatwe przesyłanie plików symulacyjnych pomiędzy środowiskiem użytkownika a urządzeniem docelowym. Istniejące metody wymagają zazwyczaj stosowania dodatkowej konfiguracji, narzędzi zewnętrznych

lub wiedzy administracyjnej, co może stanowić barierę dla użytkowników początkujących oraz w zastosowaniach dydaktycznych. Dodatkowym problemem jest brak mechanizmów umożliwiających zdalne sterowanie przebiegiem symulacji z poziomu komputera użytkownika. W szczególności nie istnieje prosta i ustandaryzowana metoda pozwalająca na wywoływanie funkcji lub modyfikowanie stanu pliku symulacyjnego działającego na platformie Raspberry Pi w sposób kontrolowany i synchroniczny. W praktyce oznacza to, że komunikacja pomiędzy aplikacją kliencką a uruchomioną symulacją jest ograniczona lub wymaga implementacji złożonych rozwiązań opartych na protokołach sieciowych.

2. Cel pracy

Celem niniejszego rozdziału jest przedstawienie rozwiązania umożliwiającego uproszczone przesyłanie oraz zdalne sterowanie plikami symulacyjnymi uruchamianymi na platformie Raspberry Pi. Zaproponowane podejście pozwala na kontrolę przebiegu symulacji z poziomu komputera użytkownika poprzez zastosowanie pośredniej warstwy zarządzającej oraz mechanizmów komunikacji międzyprocesowej, co eliminuje konieczność stosowania złożonych narzędzi konfiguracyjnych. Rozwiązanie zostało zaprojektowane z myślą o łatwej integracji z istniejącymi środowiskami symulacyjnymi oraz minimalizacji nakładu pracy po stronie użytkownika. W rozdziale omówiono architekturę systemu, zastosowane mechanizmy komunikacji oraz sposób realizacji zdalnego sterowania. Skuteczność zaproponowanego podejścia została zweryfikowana na przykładzie wybranej symulacji uruchamianej na platformie Raspberry Pi.

3. Architektura komunikacji

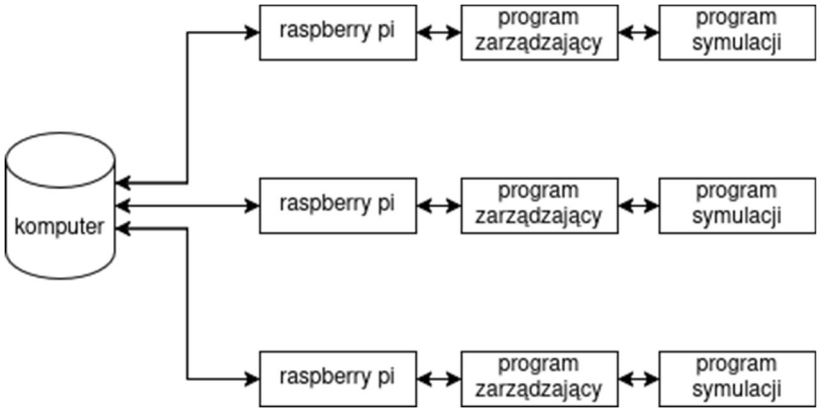
Zaproponowana architektura komunikacji opiera się na modelu klient–serwer [6], w którym komputer użytkownika pełni rolę centralnego węzła sterującego, natomiast wiele urządzeń Raspberry Pi (odtąd zwanych klientami) realizuje zadania związane z obsługą i uruchomieniem symulacji. Komunikacja pomiędzy komputerem a poszczególnymi klientami realizowana jest za pomocą protokołu WebSocket, co umożliwia dwukierunkową, asynchroniczną wymianę danych w czasie rzeczywistym.

Każdy klient uruchamia lokalny program zarządzający, który odpowiada za obsługę połączenia sieciowego z komputerem oraz pośredniczy w komunikacji z programem symulacji. Program zarządzający pełni rolę warstwy pośredniej, izolując logikę komunikacji sieciowej od logiki samej symulacji.

Wymiana danych pomiędzy programem zarządzającym a programem symulacji realizowana jest z wykorzystaniem mechanizmów komunikacji międzyprocesowej, w szczególności potoków nazwanych. Takie rozwiązanie umożliwia przekazywanie poleceń sterujących oraz danych wejściowych do symulacji bez

konieczności modyfikowania jej struktury ani implementowania dodatkowych mechanizmów sieciowych w samym programie symulacyjnym.

Zastosowanie potoków nazwanych pozwala na synchroniczną i uporządkowaną komunikację pomiędzy procesami [5] działającymi na tym samym urządzeniu, a jednocześnie umożliwia zdalne sterowanie przebiegiem symulacji z poziomu komputera użytkownika. W rezultacie architektura ta pozwala na skalowalne zarządzanie wieloma instancjami symulacji uruchomionymi na różnych urządzeniach Raspberry Pi przy zachowaniu prostoty i modularności systemu.



Rys. 1. Schemat komunikacji między serwerem a poszczególnymi klientami.
Źródło: opracowanie własne.

Serwer komunikacyjny został zaimplementowany w języku C++ z wykorzystaniem bibliotek Boost oraz Boost.Beast [2]. W celu obsługi wielu jednoczesnych połączeń zastosowano model wielowątkowy, w którym pojedynczy wątek roboczy może obsługiwać wiele połączeń sieciowych. Takie podejście pozwala na efektywne zarządzanie komunikacją z wieloma klientami oraz zapewnia skalowalność systemu przy zachowaniu niskiego narzutu obliczeniowego (rys. 1.).

Program zarządzający oraz program symulacji uruchamiane na urządzeniach Raspberry Pi zostały zaimplementowane w języku Python, a warstwa komunikacji została zrealizowana za pomocą modułu websockets [3] oraz asyncio [4]. Takie rozwiązanie umożliwia szybkie tworzenie i modyfikowanie logiki symulacyjnej oraz upraszcza integrację z mechanizmami systemowymi, w szczególności komunikacją międzyprocesową opartą na potokach nazwanych. Jednocześnie zachowana zostaje separacja pomiędzy warstwą sieciową realizowaną przez serwer a logiką symulacji wykonywaną lokalnie na urządzeniu.

4. Konfiguracja urządzeń

Prawidłowe działanie zaproponowanego systemu wymaga odpowiedniej konfiguracji urządzeń Raspberry Pi oraz uruchamianych na nich programów symulacyjnych. Konfiguracja ta została rozdzielona na dwa poziomy: konfigurację urządzenia oraz konfigurację pojedynczej symulacji.

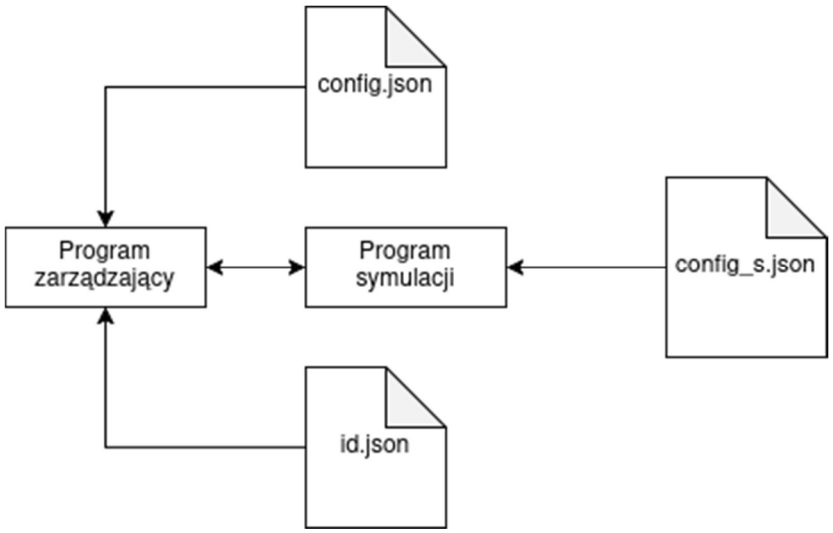
Konfiguracja urządzenia Raspberry Pi realizowana jest za pomocą pliku konfiguracyjnego config.json, który jest odczytywany przez program zarządzający podczas uruchamiania systemu. Plik ten zawiera informacje niezbędne do nawiązania połączenia z komputerem sterującym, takie jak adres serwera, port komunikacyjny oraz podstawowe parametry pracy programu zarządzającego. Dzięki temu możliwe jest dynamiczne dołączanie wielu klientów do systemu bez konieczności modyfikowania kodu aplikacji.

Dodatkowo każdy klient posiada plik identyfikacyjny id.json, którego celem jest jednoznaczna identyfikacja urządzenia w systemie. Plik ten jest przesyłany do komputera sterującego w momencie nawiązywania połączenia WebSocket i zawiera informacje opisujące urządzenie, takie jak jego nazwa, typ, wersja programu zarządzającego oraz inne dane identyfikacyjne. Na podstawie zawartości pliku id.json podejmowana jest decyzja o zaakceptowaniu urządzenia w systemie lub zerwaniu połączenia. Mechanizm identyfikacji umożliwia kontrolę dostępu urządzeń do systemu oraz weryfikację ich zgodności z wymaganiami środowiska, w szczególności pod względem wersji oprogramowania zarządzającego. Pozwala to na eliminację urządzeń nieautoryzowanych lub niekompatybilnych oraz zwiększa stabilność i bezpieczeństwo całego systemu.

Drugim poziomem konfiguracji jest plik konfiguracyjny przypisany do każdego programu symulacji. Plik ten określa parametry uruchomieniowe symulacji oraz zestaw funkcji sterujących udostępnionych do zdalnego wywołania z poziomu komputera użytkownika. Wyeksponowane funkcje mogą obejmować m.in. inicjalizację symulacji, zmianę jej parametrów, zatrzymanie lub ponowne uruchomienie.

Program zarządzający, na podstawie plików konfiguracyjnych, pośredniczy w przekazywaniu poleceń otrzymanych z komputera sterującego do odpowiednich funkcji programu symulacji. Takie podejście umożliwia elastyczne zarządzanie wieloma instancjami symulacji oraz ich zachowaniem bez konieczności ingerencji w logikę komunikacyjną systemu, a jednocześnie pozwala na łatwe rozszerzanie funkcjonalności poprzez modyfikację plików konfiguracyjnych.

Do obsługi zdalnie wywoływanych funkcji po stronie programu symulacji zastosowano autorski moduł clientProcessesCommunication. Moduł ten stanowi warstwę pośrednią pomiędzy logiką aplikacji symulacyjnej a programem zarządzającym uruchomionym na Raspberry Pi. Jego zadaniem jest rejestrowanie funkcji możliwych do wywołania z poziomu serwera, odbieranie komend przekazanych przez program zarządzający oraz uruchamianie odpowiednich funkcji w programie symulacji.



Rys. 2. Schemat konfiguracji klienta.
Źródło: opracowanie własne.

Dzięki temu program symulacji nie musi samodzielnie implementować obsługi komunikacji sieciowej ani bezpośrednio zarządzać połączeniem WebSocket. Wystarczające jest zaimportowanie modułu clientProcessesCommunication, zarejestrowanie funkcji sterujących oraz cykliczne odpytywanie modułu o nowe zdarzenia. Pozwala to zachować separację pomiędzy logiką symulacji a mechanizmem komunikacyjnym systemu.

5. Architektura komunikatu

W projektowanym interfejsie komunikacyjnym zastosowano strukturę komunikatów opartą na formacie JSON (rys. 2.). Wybór tego formatu wynika z jego uniwersalności, czytelności oraz niewielkiego narzutu danych w porównaniu z bardziej rozbudowanymi rozwiązaniami serializacji. Komunikaty w tym formacie pełnią rolę podstawowego nośnika informacji wymienianej pomiędzy poszczególnymi komponentami systemu, umożliwiając jednoznaczną interpretację przesyłanych danych niezależnie od platformy sprzętowej i języka implementacji. Zdefiniowana struktura komunikatu pozwala na przekazywanie zarówno poleceń sterujących, jak i danych konfiguracyjnych oraz informacji o stanie systemu. Obiekt komunikatu posiada następujące pola, których znaczenie i sposób wykorzystania zostały przedstawione w dalszej części rozdziału.

Tabela 1. Pola komunikatu, opracowanie własne

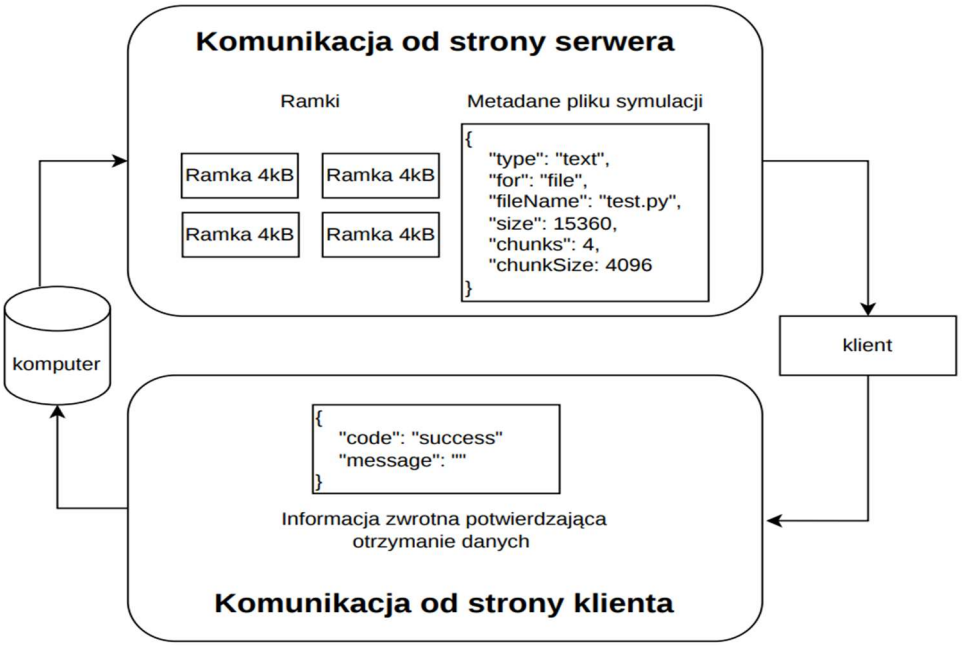
Pole	Opis
Type	Dozwolona zawartość: text, binary. Opisuje format zawartości pola content, które może być albo tekstem zakodowanym w formacie UTF-8 albo heksadecymalną reprezentacją bajtów w pliku.
For	Dozwolona zawartość: communication, frame, application, file. Opisuje typ komunikatu (w ww. kolejności): komunikacja pomiędzy serwerem a programem zarządzającym; za przesyłanie obrazu; za komunikację pomiędzy serwerem a programem symulacji; za przesyłanie pliku symulacji.
Content	Zawartość komunikatu.
Size	Rozmiar pola **content** w bajtach.
Chunks	**Specyficzny dla typu: frame i file.** Liczba ramek należących do danego pliku.
Chunksize	**Specyficzny dla typu: frame i file.** Wielkość pojedynczej ramki w bajtach.

Rozgranicza się cztery kluczowe grupy komunikatów, niezbędne do samodzielnego funkcjonowania programu symulacji oraz programu zarządzającego (tabela 2.), przy jednoczesnym zachowaniu komunikacji w trybie full-duplex. Każda z wyróżnionych grup odpowiada za inny aspekt wymiany danych, obejmując m.in. przesyłanie poleceń sterujących, danych konfiguracyjnych, informacji o stanie symulacji oraz komunikatów zwrotnych. Takie rozdzielenie komunikatów pozwala na uporządkowanie logiki komunikacyjnej oraz ułatwia jej dalszą rozbudowę i utrzymanie. Największym wyzwaniem projektowym jest opracowanie lekkiego rozwiązania, które nie powoduje nadmiernego obciążenia sieci, zwłaszcza w scenariuszach obejmujących jednoczesną obsługę większej liczby klientów. W tym celu istotne jest ograniczenie rozmiaru przesyłanych komunikatów oraz minimalizacja częstotliwości transmisji przy zachowaniu wymaganej responsywności systemu.

Tabela 2. Grupy komunikatów, opracowanie własne

Grupa	Pola JSON	Opis
komenda z serwera/ odpowiedź klienta	type: text; for: communication	Pole content zawiera polecenie/odpowiedź zapisaną w standardzie UTF-8.
przesłanie pliku symulacji	type: binary; for: file	Aplikacja jest przesyłana do klienta w postaci pliku binarnego(program symulacji). Jeżeli plik jest większy niż określona wartość, plik jest dzielony na ramki, wtedy ustawiane są również pola chunks i chunkSize
przesyłanie komend i odpowiedzi programu symulacji	type: text; for: application	Pole content zawiera polecenie/odpowiedź zapisaną w standardzie UTF-8.
przesyłanie obrazu z programu zarządzającego	type: text; for: frame	Dla każdej klatki obrazu przesyłany jest plik z metadanymi (brak pola content). Ustawiane są pola chunks i chunkSize.

Komunikat przesyłania obrazu jest unikalny — podobnie jak komunikat przesyłania aplikacji w przypadku, gdy jej rozmiar przekracza podaną wartość – ponieważ nie zawiera pól **content** oraz **size** (tabela 1.). Takie rozwiązanie wynika z celowego rozdzielenia odpowiedzialności pomiędzy warstwą sygnalizacyjną a warstwą transmisji danych. Komunikat ten pełni wyłącznie funkcję informacyjną, powiadamiając odbiorcę o nadchodzącym pliku o dużym rozmiarze, którego nie jest możliwe przesłanie w całości w pojedynczej ramce. Po jego odebraniu serwer przechodzi w tryb odbioru strumieniowego i interpretuje kolejną, z góry określoną liczbę ramek jako fragmenty jednej, większej całości. Takie podejście zapobiega błędnej interpretacji danych jako niezależnych komunikatów, a jednocześnie umożliwia kontrolę integralności oraz poprawności składania przesyłanego pliku.



Rys. 3. Przykładowy schemat komunikacji pomiędzy serwerem a klientem
Źródło: opracowanie własne.

6. Architektura odpowiedzi na komunikat

W zaprojektowanym systemie przyjęto model, w którym wyłącznie klient przesyła odpowiedzi do serwera. Serwer inicjuje komunikację poprzez wysyłanie poleceń, natomiast wszelkie informacje zwrotne, wyniki operacji oraz komunikaty błędów przekazywane są przez klienta. Odpowiedzi mają taką samą strukturę jak komunikaty i są reprezentowane w formacie JSON (rys. 3.), co zapewnia spójność

oraz jednoznaczną interpretację danych po obu stronach połączenia. Każde polecenie wysłane przez serwer traktowane jest jako zdarzenie, którego rezultat klient przekazuje w postaci niezależnego komunikatu zwrotnego. Serwer nie wysyła odpowiedzi do klienta, a jedynie analizuje otrzymane dane (tabela 3.). Takie podejście upraszcza logikę komunikacyjną, zmniejsza złożoność zarządzania stanem połączeń oraz sprzyja skalowalności systemu przy obsłudze wielu klientów jednocześnie.

Tabela 3. Schemat odpowiedzi, opracowania własne

Pole	Opis
code	pole określające wynik obsługi komunikatu po stronie klienta. Przyjmuje wartości opisujące status wykonania żądania (np. success dla poprawnego wykonania lub kody błędów dla sytuacji niepoprawnych). Na podstawie code serwer rozróżnia, czy operacja zakończyła się powodzeniem oraz czy wymagane są dodatkowe działania (np. ponowienie, przerwanie procedury).
message	pole opcjonalne zawierające dodatkową informację diagnostyczną. Jest wykorzystywane przede wszystkim wtedy, gdy code wskazuje stan inny niż success i może zawierać opis przyczyny błędu, kontekst (np. nazwa operacji, brakujący zasób, niepoprawny format danych) lub wskazówkę ułatwiającą diagnostykę. Dla code = success pole message może być puste lub pominięte.

7. Rejestrowanie i przesyłanie logów

W celu zapewnienia możliwości monitorowania pracy programów symulacyjnych oraz diagnostyki błędów w systemie, w zaproponowanej architekturze zastosowano mechanizm rejestrowania logów. Program symulacji generuje komunikaty diagnostyczne oraz informacje o przebiegu działania, które są wypisywane na standardowe wyjście (stdout) oraz standardowe wyjście błędów (stderr).

Program zarządzający przechwytuje strumienie stdout oraz stderr programu symulacji i rejestruje je w lokalnych plikach logów na urządzeniu Raspberry Pi. Logi te zapisywane są w sposób ciągły, co umożliwia zachowanie pełnej historii działania symulacji oraz identyfikację ewentualnych błędów w trakcie jej wykonywania.

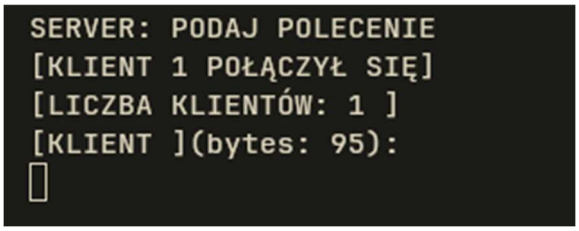
Komputer sterujący, pełniący rolę serwera, cyklicznie odpytuje podłączone urządzenia Raspberry Pi o dostępność nowych wpisów w logach. W przypadku wykrycia nowych danych program zarządzający przesyła je do serwera za pośrednictwem połączenia WebSocket. Mechanizm ten pozwala na zdalny podgląd logów w czasie zbliżonym do rzeczywistego, bez konieczności bezpośredniego dostępu do systemu plików urządzenia.

Zastosowane rozwiązanie umożliwia centralne gromadzenie logów z wielu urządzeń Raspberry Pi, co znacząco ułatwia analizę działania systemu, wykrywanie błędów oraz nadzór nad przebiegiem symulacji. Jednocześnie zachowana zostaje separacja pomiędzy logiką symulacji a mechanizmami komunikacyjnymi i diagnostycznymi.

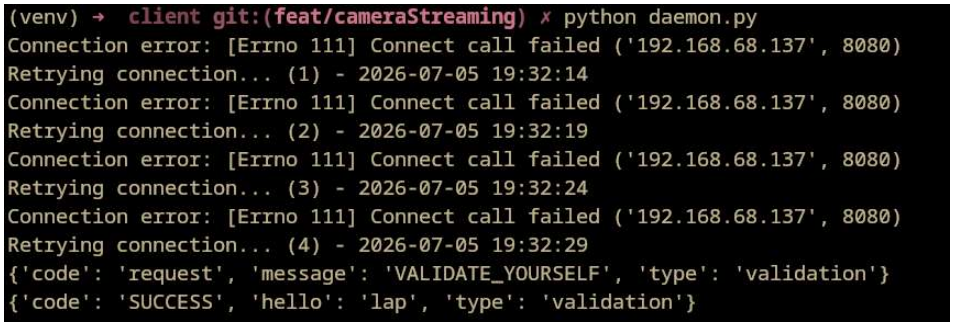
8. Weryfikacja działania systemu

W celu sprawdzenia poprawności działania zaproponowanego interfejsu przygotowano stanowisko testowe składające się z komputera sterującego oraz urządzenia klienckiego pracującego w tej samej sieci lokalnej. Na komputerze uruchomiono serwer komunikacyjny, natomiast na kliencie program zarządzający oraz przykładowy program symulacji.

Weryfikacja obejmowała nawiązanie połączenia WebSocket, przesłanie pliku symulacji, uruchomienie programu po stronie klienta, przekazanie komendy sterującej do programu symulacji za pomocą potoków nazwanych oraz odebranie odpowiedzi zwrotnej przez serwer. Dodatkowo sprawdzono przesyłanie komunikatów diagnostycznych w postaci logów.



Rys. 4. Nawiązanie połączenia WebSocket oraz walidacja klienta
Źródło: opracowanie własne.



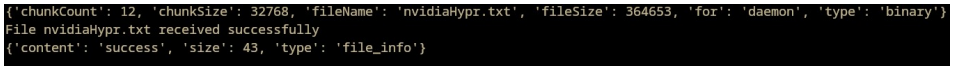
Rys. 5. Ponawianie połączenia oraz walidacja klienta po stronie programu zarządzającego
Źródło: opracowanie własne.

W pierwszym etapie weryfikacji sprawdzono poprawność nawiązania połączenia pomiędzy serwerem a urządzeniem klienckim. Po uruchomieniu programu zarządzającego po stronie klienta serwer poprawnie zarejestrował nowe połączenie oraz zaktualizował liczbę aktywnych klientów, co przedstawiono na rys. 4. Po stronie klienta zarejestrowano moment połączenia się z serwerem na rys. 5. Potwierdza to poprawne działanie podstawowego mechanizmu komunikacji WebSocket oraz możliwość identyfikacji podłączonego urządzenia.

Następnie przeprowadzono test przesyłania pliku z komputera sterującego do klienta. Serwer wysłał polecenie przesłania pliku (rys. 6.), a klient odebrał metadane opisujące nazwę pliku, jego rozmiar oraz liczbę fragmentów, na które został podzielony podczas transmisji. Przebieg tej operacji przedstawiono na rys. 7. Otrzymany komunikat potwierdzający poprawne zakończenie transferu wskazuje, że mechanizm dzielenia pliku na ramki oraz jego ponownego składania po stronie klienta działa prawidłowo.

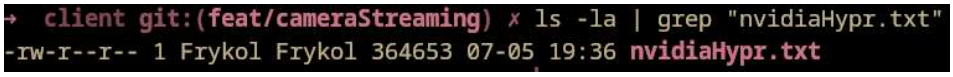


Rys. 6. Polecenie przesłania pliku z serwera do klienta
Źródło: opracowanie własne



Rys. 7. Odbiór metadanych pliku oraz potwierdzenie zakończenia transmisji po stronie klienta
Źródło: opracowanie własne.

Ostatnim etapem testu było sprawdzenie, czy przesłany plik został rzeczywiście zapisany w systemie plików urządzenia klienckiego. W tym celu wyświetlono zawartość katalogu i zweryfikowano obecność odebranego pliku wraz z jego rozmiarem, co pokazano na rys. 8. Zgodność rozmiaru pliku zapisanego po stronie klienta z wartością przesłaną w metadanych potwierdza poprawność transmisji oraz zapisania danych.



Rys 8. Potwierdzenie odbioru pliku
Źródło: opracowanie własne.

Kolejnym etapem weryfikacji było sprawdzenie możliwości zdalnego zarządzania uruchomioną symulacją. W tym celu z poziomu serwera wybrano aktywny program symulacyjny za pomocą polecenia SWITCH lap simBim.py, a następnie wysłano do niego komendy wywołujące funkcję dodaj z argumentami 15 oraz 12. Po każdej operacji pobrano logi działania programu za pomocą polecenia GET lap LOG, co pozwoliło na sprawdzenie, czy polecenia zostały poprawnie wykonane po stronie klienta.

Przebieg operacji po stronie serwera przedstawiono na rys. 9. Widoczne są kolejno polecenia przełączenia aktywnej symulacji, wywołania funkcji oraz po-

brania logów. Otrzymane komunikaty diagnostyczne potwierdzają zmianę wartości akumulatora najpierw do 15, a następnie do 27, co oznacza, że funkcja została wywołana dwukrotnie z poprawnymi argumentami.

```
SWITCH lap simBim.py
SERVER: PODAJ POLECENIE
SEND lap FUNCTION dodaj 15
SERVER: PODAJ POLECENIE
GET lap LOG
[2026-07-05 20:02:03 OUTPUT] - Dodano do akumulatora, teraz wartosc 15

SERVER: PODAJ POLECENIE
SEND lap FUNCTION dodaj 12
SERVER: PODAJ POLECENIE
GET lap LOG
[2026-07-05 20:02:03 OUTPUT] - Dodano do akumulatora, teraz wartosc 15
[2026-07-05 20:02:16 OUTPUT] - Dodano do akumulatora, teraz wartosc 27
```

Rys 9. Zdalne przełączenie symulacji, wywołanie funkcji oraz odczyt logów po stronie serwera
Źródło: opracowanie własne.

Na rysunku 10. przedstawiono komunikaty odebrane po stronie klienta. Pierwszy komunikat informuje o wyborze programu simBim.py jako aktywnej symulacji, natomiast kolejne komunikaty zawierają polecenia przekazane do aplikacji. Potwierdza to, że program zarządzający poprawnie odbiera dane z serwera i przekazuje je do programu symulacji jako komendy aplikacyjne.

```
{'fileName': 'simBim.py', 'for': 'simulation', 'type': 'application'}
{'cmd': 'dodaj 15 ', 'for': 'application', 'type': 'command'}
{'cmd': 'dodaj 12 ', 'for': 'application', 'type': 'command'}
```

Rys 10. Komunikaty odebrane przez program zarządzający po stronie klienta
Źródło: opracowanie własne.

Przeprowadzony test potwierdził, że możliwe jest zdalne przesłanie i uruchomienie programu symulacji oraz sterowanie jego działaniem bez konieczności ręcznego logowania się do urządzenia klienta. System poprawnie rozpoznawał klienta na podstawie pliku identyfikacyjnego, odbierał komunikaty zwrotne oraz umożliwiał monitorowanie działania programu.

9. Ograniczenia i kierunki dalszego rozwoju

Obecna implementacja interfejsu komunikacyjnego, mimo że umożliwia efektywne przesyłanie plików symulacyjnych i zdalne sterowanie instancjami na wielu urządzeniach Raspberry Pi, posiada pewne ograniczenia. Przesyłanie dużych

plików, zwłaszcza obrazów i aplikacji powyżej 4 kB, wiąże się z podziałem na ramki, co w scenariuszach z wieloma jednoczesnymi klientami może prowadzić do zwiększonego narzutu sieciowego i wydłużenia czasu przesyłu. Kolejnym ograniczeniem jest brak mechanizmów adaptacyjnych w zakresie przepustowości i priorytetyzacji komunikatów. W obecnej implementacji wszystkie typy komunikatów są przesyłane równorzędnie, co w przypadku obciążenia sieci może skutkować spadkiem responsywności systemu lub opóźnieniami w sterowaniu symulacjami. Dodatkowo architektura oparta na modelu klient–serwer i protokole WebSocket zakłada ciągłe połączenie z każdym klientem, co w przypadku bardzo dużej liczby urządzeń może wymagać dalszej optymalizacji wydajnościowej [7]. W ramach dalszego rozwoju systemu planowane są modyfikacje i rozszerzenia mające na celu zwiększenie jego funkcjonalności i bezpieczeństwa. Możliwe kierunki obejmują wprowadzenie optymalizacji transmisji dużych plików, np. przez kompresję lub dynamiczne zarządzanie priorytetami ramek. Dodatkowo badania nad skalowalnością w środowiskach z bardzo dużą liczbą klientów mogą pozwolić na wprowadzenie mechanizmów rozproszonego serwera lub kolejkowania komunikatów, co zwiększy elastyczność systemu w zastosowaniach laboratoryjnych i dydaktycznych.

10. Podsumowanie

Przeprowadzona analiza oraz implementacja uniwersalnego interfejsu komunikacyjnego wykazały, że możliwe jest efektywne zdalne przesyłanie i sterowanie plikami symulacyjnymi uruchamianymi na wielu urządzeniach Raspberry Pi przy zachowaniu modularnej i skalowalnej architektury. Zastosowanie protokołu WebSocket oraz pośredniej warstwy zarządzającej umożliwia dwukierunkową, asynchroniczną komunikację, przy jednoczesnym uproszczeniu integracji z istniejącymi programami symulacyjnymi.

Struktura komunikatów oparta na formacie JSON zapewnia czytelność, elastyczność oraz bezpieczeństwo interpretacji przesyłanych danych, a wydzielenie grup komunikatów i mechanizm ramkowania pozwala na przesyłanie dużych plików bez ryzyka ich nieprawidłowego odczytu. Mechanizm rejestrowania i zdalnego przesyłania logów umożliwia bieżący monitoring pracy systemu i ułatwia diagnostykę błędów. Opracowane rozwiązanie charakteryzuje się prostotą konfiguracji, modularnością i możliwością skalowania do obsługi wielu klientów, co czyni je przydatnym zarówno w zastosowaniach dydaktycznych, jak i prototypowych. Jednocześnie wskazano obszary wymagające dalszych usprawnień, w tym bezpieczeństwo komunikacji oraz optymalizację przesyłania dużych plików.

Literatura

1. Andrew Lombardi, WebSocket. Lightweight Client-Server Communications, Wydawnictwo O'Reilly Media, 2015.

2. Danny Staple, Jak zaprogramować robota. Zastosowanie Raspberry Pi i Pythona w tworzeniu autonomicznych robotów, Wydanie II, Wydawnictwo Helion, 2022.
3. Nicholas A. Solter, Scott J. Kleper, C++. Zaawansowane programowanie, Wydawnictwo Helion, 2005.

Źródła

[1] https://www.komputerswiat.pl/komputery-i-laptopy/komputery/raspberry-pi-sprzedano-juz-ponad-125-mln-minikomputerow/6jftzt8 [dostęp: 11.02.2026].
[2] https://www.boost.org/doc/libs/latest/libs/beast/doc/html/index.html [dostęp: 11.02.2026].
[3] https://websockets.readthedocs.io/en/stable [dostęp: 11.02.2026].
[4] https://docs.python.org/3/library/asyncio.html [dostęp: 11.02.2026].
[5] https://opensource.com/article/19/4/interprocess-communication-linux-channels [dostęp: 11.02.2026].
[6] https://www.researchgate.net/profile/Shakirat-Sulyman/publication/271295146_Client-Server_Model/links/5864e11308ae8fce490c1b01/Client-Server-Model.pdf [dostęp: 11.02.2026].
[7] https://dev.to/ably/challenges-of-scaling-websockets-3493 [dostęp: 11.02.2026].

UNIVERSAL COMMUNICATION INTERFACE BETWEEN A COMPUTER AND MANY RASPBERRY PI-BASED PROJECTS

Summary

This article presents the design and implementation of a universal communication interface for remotely transferring and controlling simulation files running on multiple Raspberry Pi devices. The goal was to create a unified, easily configurable solution, particularly useful for teaching and prototyping applications. The architecture is based on a client-server model using the WebSocket protocol, ensuring bidirectional, real-time communication. A central C++ server (using the Boost and Boost.Beast libraries) manages clients running on target devices. Communication between the management program and the simulation occurs via named pipes, without interfering with the application logic. Data exchange is based on the JSON format and includes control commands, file transfer, communication with the application, and image transfer. The system has been enhanced with a logging and central monitoring mechanism.

Keywords: Raspberry Pi, WebSocket, C++ , Python, communication interface.

Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Krystian Sołtys, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST

dr hab. inż. Tomasz Kapuściński, prof. PRz, dr inż. Dawid Kalandyk
opiekunowie Koła Naukowego

PROJEKT I STEROWANIE ROBOTEM SAMOBALANSUJĄCYM

Streszczenie

Niniejszy rozdział naukowy przedstawia kompleksowy proces projektowania, budowy oraz oprogramowania dwukołowego robota samobalansującego, opartego na niestabilnym modelu odwróconego wahadła. Szczegółowo omówiono konstrukcję mechaniczną w formie zwartej obudowy przeznaczonej do druku 3D, w której zastosowano podział na dwie strefy: dolną z akumulatorem dla obniżenia środka ciężkości oraz górną z jednostką sterującą Raspberry Pi . Kluczowym aspektem pracy jest implementacja systemu sterowania w czasie rzeczywistym przy użyciu języka Python. Wykorzystano sensory inercyjne oraz filtr komplementarny do precyzyjnej estymacji kąta wychylenia, eliminując błędy pomiarowe. Zaprojektowano układ sterowania, którego celem jest stabilizacja układu za pomocą dyskretnego regulatora PID, który steruje silnikami prądu stałego poprzez sygnał PWM . Robot pełni funkcję narzędzia dydaktycznego na zajęciach ze sztucznej inteligencji, a plany rozwojowe obejmują dodanie platformy transportowej do badania wpływu zmiennego obciążenia na dynamikę sterowania.

Słowa kluczowe: robotyka, robot balansujący, Raspberry Pi, odwrócone wahadło.

1. Wprowadzenie

W ostatnich latach robotyka mobilna przeżywa dynamiczny rozwój, obejmujący wiele specjalistycznych obszarów. Jednym z najbardziej interesujących, a zarazem wymagających zagadnień, jest projektowanie robotów balansujących, czyli maszyn, których konstrukcja opiera się na niestabilnym z natury modelu odwróconego wahadła [1]. Ich unikalną cechą jest zdolność do utrzymywania pionowej pozycji w punkcie równowagi chwiejnej przy jednoczesnym przemieszczaniu się,

co zapewnia im wysoką zwrotność i możliwość operowania w wąskich przestrzeniach, niedostępnych dla klasycznych platform wielokołowych.

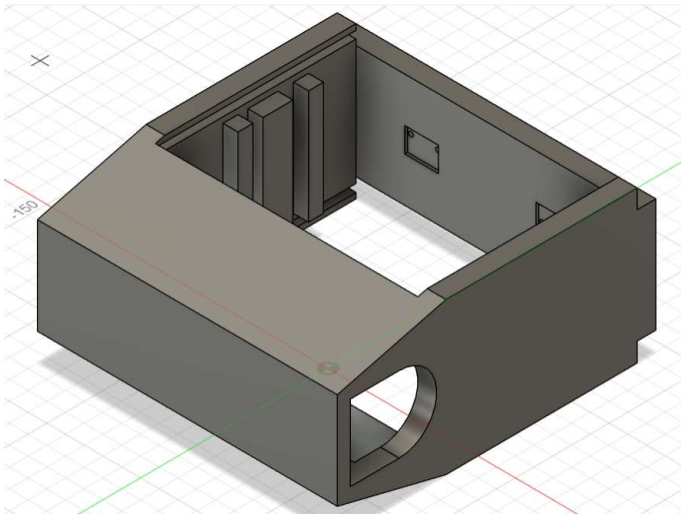
Roboty samobalansujące stanowią interesujący przykład zastosowania systemów sterowania w praktyce. Podobnie jak człowiek instynktownie balansuje ciałem, aby nie upaść, tak robot musi błyskawicznie reagować na zmiany kąta wychylenia. Taka dokładność jest możliwa dzięki wykorzystaniu inercyjnych systemów pomiarowych, łączących działanie akcelerometrów i żyroskopów. Popularnym zastosowaniem badanych możliwości maszyn jest Segway Ninebot - robot balansujący używany głównie jako środek transportu osobistego [2].

2. Konstrukcja robota

Podstawą fizycznej realizacji robota balansującego jest zaprojektowanie odpowiedniej struktury nośnej, która jest wystarczająco sztywna, chroni komponenty elektroniczne oraz, co kluczowe w przypadku modelu odwróconego wahadła, zapewnia optymalny rozkład masy. W ramach prac projektowych opracowano model obudowy, przystosowany do wydruku za pomocą drukarki 3D [3].

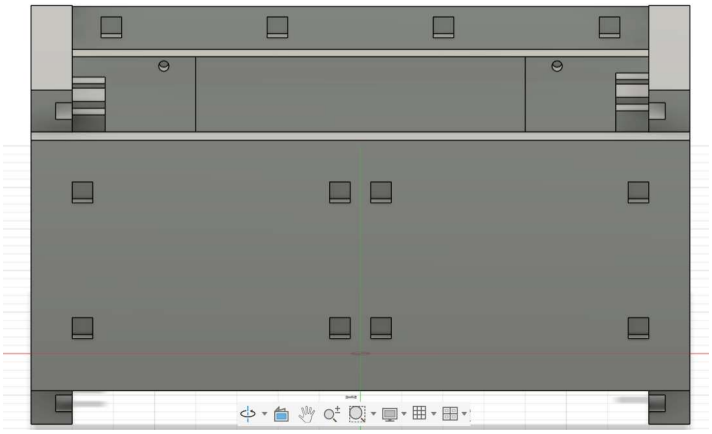
2.1. Geometria kadłuba i założenia projektowe

Główny korpus robota (przedstawiony na rys. 1. i rys. 2.) został zaprojektowany jako jednolita, zwarta konstrukcja skorupowa. Taka konstrukcja zapewnia wysoką odporność na uszkodzenia mechaniczne mogące powstać podczas ewentualnych upadków, które są nieodłącznym elementem fazy testowania algorytmów stabilizacji.



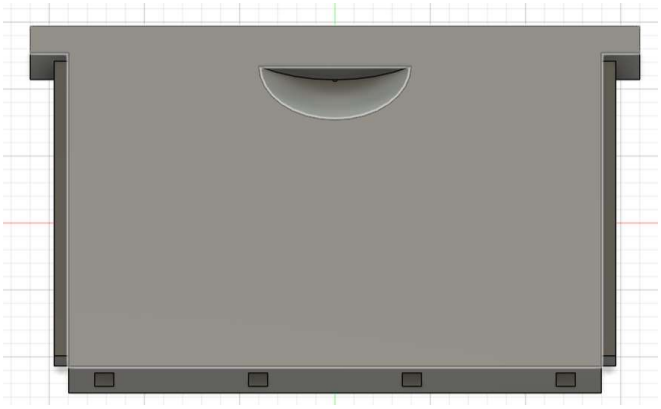
Rys. 1. Rzut boczny obudowy

Fig. 1. Side view of the case

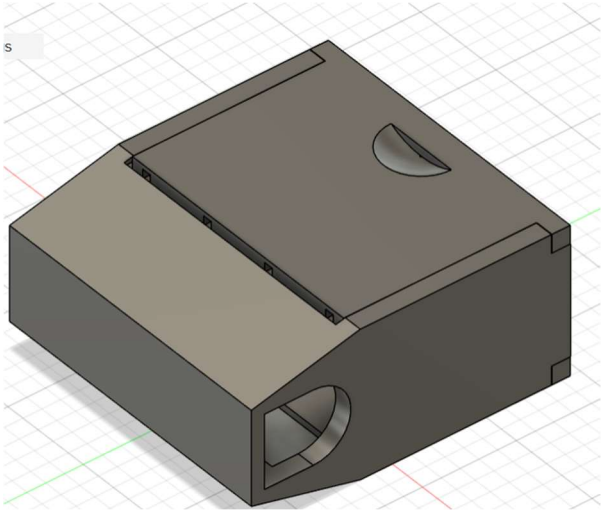


Rys. 2. Rzut górny obudowy
Fig. 2. Top view of the case

Charakterystyczne wycięcia w bocznych ścianach obudowy (widoczne na rzutach izometrycznych) dedykowane są osiom napędowym silników. Ich umiejscowienie w dolnej części bryły podyktowane jest koniecznością maksymalnego obniżenia środka ciężkości całego układu. Przednie i tylne ściany obudowy, tak samo jak platformy, będą wsuwane (jak widać na rys. 2.) z dodatkowym zabezpieczeniem w postaci magnesów (rys. 3.). Końcowy wygląd obudowy, po złożeniu, ma być solidny, a jednocześnie łatwy w modyfikacji (rys. 4.).



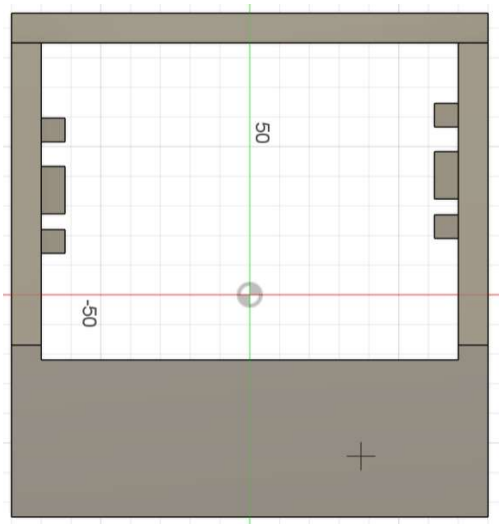
Rys. 3. Przednia i tylna ściana obudowy
Fig. 3. Front and back wall of the case



Rys. 4. Przedstawienie złożonej obudowy
Fig. 4. View of the assembled case

2.2. Wewnętrzna architektura i rozmieszczenie komponentów

Wnętrze obudowy zostało zaprojektowane z wykorzystaniem półek, widocznych na rzucie przednim (rys. 5.). Rozwiązanie to pozwala na modułowy montaż podzespołów na wysuwanych platformach, co ułatwia serwisowanie oraz wymianę komponentów bez konieczności demontażu całego robota.



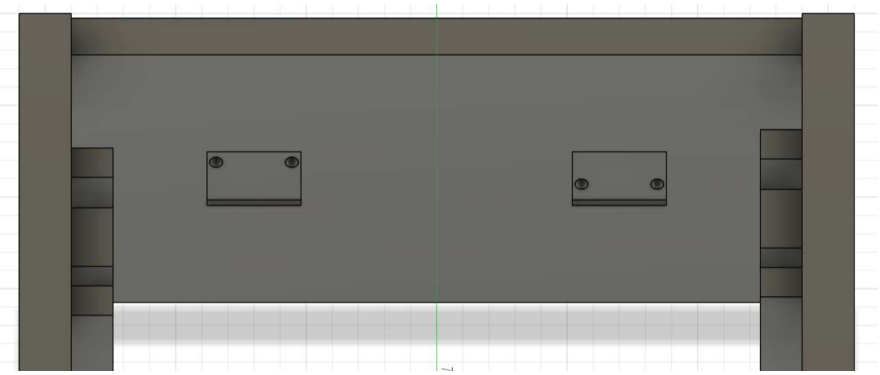
Rys. 5. Rzut przedni obudowy
Fig. 5. Front view of the case

Konfiguracja wnętrza obejmuje dwa poziomy montażowe:

- **Poziom dolny:** Dolna platforma przeznaczona jest na akumulator zasilający. Umieszczenie najcięższego elementu robota (baterii) w najniższym możliwym punkcie konstrukcji jest zabiegiem celowym, mającym na celu obniżenie momentu bezwładności względem osi kół, co pasywnie wspomaga proces stabilizacji pionowej.
- **Poziom górny:** Górna platforma dedykowana jest jednostce obliczeniowej, mikrokomputerowi Raspberry Pi. Fizyczna separacja warstwy logicznej od zasilającej minimalizuje ryzyko zakłóceń elektromagnetycznych oraz zapewnia lepsze chłodzenie procesora sterującego.

2.3. Integracja systemu sensorycznego

Kluczowym elementem dla robota balansującego jest precyzyjny pomiar orientacji w przestrzeni [4]. W górnej części obudowy zaprojektowano dedykowane gniazda montażowe dla dwóch niezależnych modułów żyroskopowych (rys. 6.).



Rys. 6. Dedykowane miejsce na żyroskopy
Fig. 6. Dedicated place for gyroscopes

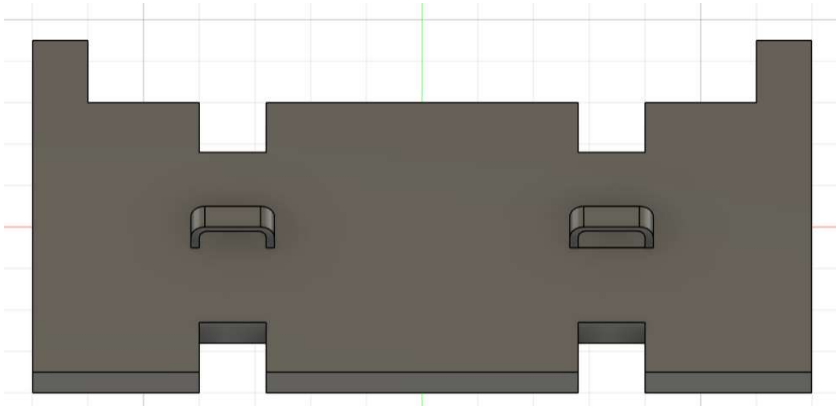
Zastosowanie sztywnych mocowań śrubowych bezpośrednio do struktury nośnej, zamiast montażu pośredniego, ma na celu eliminację drgań własnych i opóźnień mechanicznych, które mogłyby utrudniać prawidłowe działanie robota. Wykorzystanie dwóch sensorów (czujniki PCS10-BMI160, rys. 7.) pozwala na implementację algorytmów pozwalających na odczytanie dokładnej pozycji i sytuacji, w której znajduje się robot. Zamontowanie ich w górnej części konstrukcji pozwala na lepsze odczytanie przechylenia względem pionu.



Rys. 7. Czujnik PCS10-BMI160
Fig. 7. PCS10-BMI160 Sensor

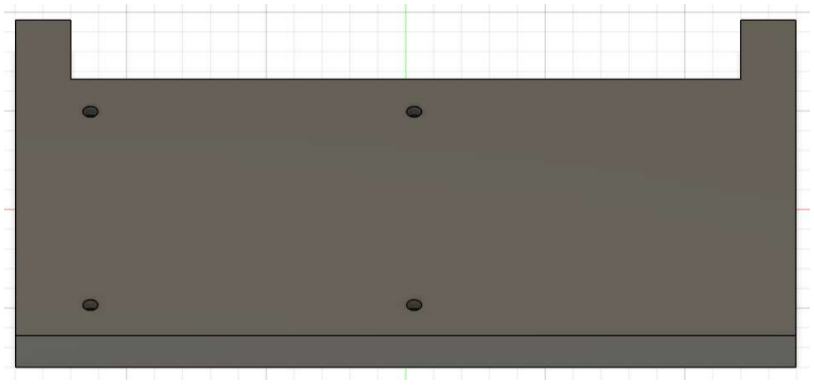
3. Układ mechaniczny

Modułem zasilającym robota jest bateria o wymiarach 3,5 cm/ 3,3 cm / 10,5 cm. Przytwierdzona do platformy za pomocą pasków mocujących, które przechodzą przez uchwyty (widoczne na rys. 8.), zatem znajduje się ona na spodzie.



Rys. 8. Dolna platforma
Fig. 8. Bottom platform

Na górnej platformie znajduje się system sterowania zrealizowany z wykorzystaniem komputera jednopłytkowego Raspberry Pi. Miejsca na gwinty potrzebne do przytwierdzenia płytki pokazane są na rys. 9. Chłodzenie modułu odbywa się za pomocą komponentu dostępnego w zestawie z płytką.



Rys. 9. Górna platforma
Fig. 9. Top platform

4. Działanie robota

4.1. Sposób pomiaru kąta odchylenia

Kąt przeliczany jest na podstawie wartości zmierzonych przez akcelerometr i żyroskop. Konieczne jest wykorzystanie obu komponentów, gdyż żyroskop mierzy prędkość kątową, przez co nie bierze on pod uwagę przyspieszenia liniowego. Akcelerometr mierzy zarówno siłę grawitacji, jak i przyspieszenie liniowe, zatem podczas poruszania się robota wynik może się różnić od tego zgodnego ze stanem rzeczywistym [5].

Sygnał akcelerometru interpretowany jest za pomocą poniższego wzoru, pozwalającego na wyznaczenie kąta pochylenia robota:

$$\phi = atan2(A_Y, A_Z) \cdot \frac{180}{\pi} \quad (1)\ [6]$$

gdzie: ϕ - kąt pochylenia wyznaczony ze wskazań akcelerometru,
A_Y , A_Z - dane odczytane z akcelerometru odpowiednio dla osi Y i Z.

Najprostszą metodą na połączenie sygnału z akcelerometru i żyroskopu jest zastosowanie filtru komplementarnego, który opisano następującym równaniem rekurencyjnym:

$$\phi = \alpha \cdot (\phi_{pop} + \omega \cdot dx) + (1 - \alpha) \cdot \phi_{acc} \quad (2)\ [7]$$

gdzie: ϕ - kąt wyjściowy,
ϕ_{pop} - poprzednia wartość kąta wyjściowego,
ω - prędkość kątowa (odczytana z żyroskopu),
ϕ_{acc} - kąt wyznaczony z zależności,
α - charakterystyka częstotliwościowa filtru, wyznaczana:

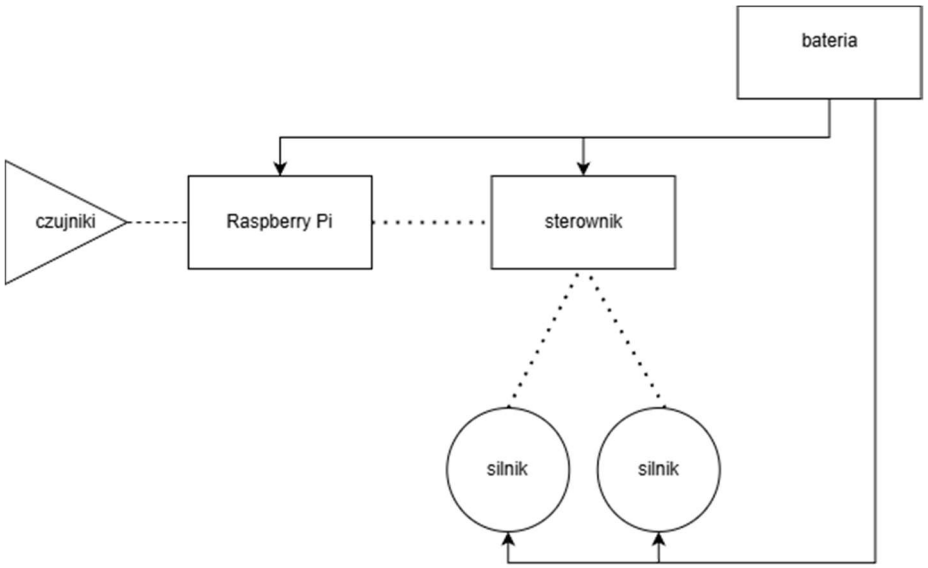
$$\alpha = \frac{\tau}{\tau + dt} \tag{3}$$

gdzie τ to stała czasowa.

4.2. Architektura programu

Oprogramowanie robota zaimplementowane jest w języku Python [7] i realizuje zadania związane z obróbką danych z żyroskopów i akcelerometrów, estymacją kąta pochylenia, obliczaniem sygnału sterującego oraz generowaniem sygnałów PWM sterujących napędem. Struktura programu ma charakter cykliczny i opiera się na pętli głównej wykonywanej z ustaloną częstotliwością próbkowania.

Po uruchomieniu systemu następuje inicjalizacja interfejsów komunikacyjnych oraz konfiguracja wyprowadzeń GPIO odpowiedzialnych za sterowanie napędu silników prądu stałego. Na tym etapie ustawiane są również parametry regulatora PID oraz zmienne pomocnicze niezbędne do obliczeń numerycznych. Planowane działanie robota przedstawione jest na rys. 10.



Rys. 10. Schemat planowanego działania robota

Fig. 10. Diagram of the robot's planned operation

Istotnym elementem fazy startowej jest kalibracja czujników, polegająca na wyznaczeniu offsetów żyroskopu w stanie spoczynku. Pozwala to ograniczyć wpływ błędu pomiarowego na estymację kąta pochylenia. W dalszej części programu, dane z akcelerometru i żyroskopu są cyklicznie odczytywane i przetwarzane w celu wyznaczenia aktualnego kąta odchylenia konstrukcji od pionu.

Ze względu na właściwości pomiarowe czujników, zastosowano filtr komplementarny. Uzyskany w ten sposób kąt stanowi podstawę do wyznaczenia regulacji, definiowanego jako różnica pomiędzy aktualnym kątem a wartością zadaną (pozycją równowagi).

W kolejnym etapie realizowany jest algorytm, który na podstawie różnicy kątów wyznacza wartość sygnału sterującego. Składowa proporcjonalna odpowiada za reakcję zależną od bieżącego błędu, składowa całkująca eliminuje różnicę kątów, natomiast składowa różniczkująca poprawia stabilność układu. Obliczony sygnał sterujący jest następnie przekształcany na odpowiednie wypełnienie sygnału PWM oraz kierunek obrotów silników. Postać ciągłą sygnału sterującego $u(t)$, który odpowiada za moment napędowy silników, opisuje równanie:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{de(t)}{dt} \quad (4)\ [8]$$

gdzie: $u(t)$ - wartość sygnału sterującego w chwili t,
$e(t)$ - uchyb regulacji, czyli różnica pomiędzy zadanym kątem (pionem) a kątem zmierzonym,
K_p - wzmocnienie proporcjonalne (reaguje na bieżący błąd),
K_i - wzmocnienie całkujące (akumulacja błędu w czasie i eliminacja ustalonej różnicy kąta),
K_d - wzmocnienie różniczkujące (estymacja trendu zmian i działanie tłumiące).

Jest to model idealny, wykorzystywany w teorii sterowania. Ponieważ program działa w sposób dyskretny (próbkowanie co określony przedział czasu), wykorzystany został wzór przybliżony, gdzie:

- Całka (aproksymacja prostokątna):

$$I_k = I_{k-1} + e_k \cdot \Delta t \quad (5)$$

- Pochodna:

$$D_k = \frac{e_k - e_{k-1}}{\Delta t} \quad (6)$$

- Sygnał sterujący:

$$u_k = K_p e_k + K_i I_k + K_d D_k \quad (7)$$

gdzie: k - numer próbki
Δt - okres próbkowania

Cały proces realizowany jest w pętli o wysokiej częstotliwości odświeżania, co zapewnia szybką reakcję na zmiany położenia środka ciężkości. Stabilność układu zależy bezpośrednio od czasu próbkowania, poprawności filtracji sygnałów oraz właściwego doboru parametrów regulatora. Poniższy schemat przedstawia teoretyczne działanie programu (rys. 11.).



Rys. 11. Schemat blokowy działania programu
Fig. 11. Program's flowchart

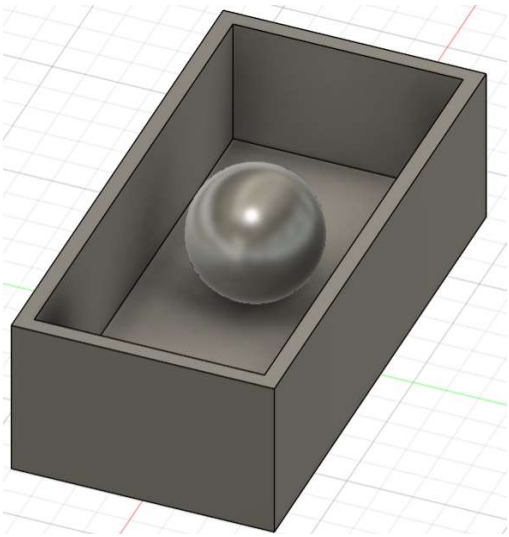
5. Wartość dydaktyczna

Opracowany robot samobalansujący będzie wykorzystywany na zajęciach laboratoryjnych w ramach przedmiotu Sztuczna Inteligencja. Konstrukcja, oparta na modelu odwróconego wahadła i wykorzystująca komputer jednopłytkowy Raspberry Pi oraz czujniki, umożliwia praktyczną analizę układów dynamicznych o charakterze niestabilnym.

Celem zajęć laboratoryjnych jest zapoznanie studentów z procesem implementacji algorytmu regulacji oraz filtracji sygnałów pomiarowych oraz strojenia parametrów sterowania w warunkach rzeczywistych. Wykorzystanie fizycznego obiektu badawczego sprzyja kształtowaniu umiejętności interpretacji wyników pomiarowych, identyfikacji problemów praktycznych [9] oraz oceny stabilności układu w różnych warunkach.

6. Planowane kierunki rozwoju

Jednym z potencjalnych kierunków rozwoju opracowanej konstrukcji jest rozbudowa o dodatkową platformę montażową umieszczoną w górnej części robota. Platforma ta miałaby pełnić funkcję powierzchni transportowej, umożliwiającej umieszczanie na niej obiektów o różnej masie i charakterystyce dynamicznej, takich jak kule, cylindry czy naczynia wypełnione cieczą. Rozszerzenie funkcjonalności w tym zakresie pozwoliłoby przekształcić układ z prostego modelu odwróconego wahadła w bardziej złożony system dynamiczny o zmiennych parametrach. Na rysunku 12. przedstawiony jest model platformy.



Rys. 12. Planowany wygląd górnej platformy
Fig. 12. Planned appearance of the top platform

Wprowadzenie dodatkowego obciążenia w górnej części konstrukcji powoduje zmianę środka masy całego układu oraz modyfikację momentu bezwładności względem osi obrotu. W konsekwencji prowadzi do zmiany parametrów dynamicznych modelu matematycznego, w szczególności współczynników występujących w równaniu ruchu. W przypadku transportu cieczy w otwartym naczyniu dochodzi dodatkowo do zjawiska przemieszczania się masy wewnątrz układu (zjawisko falowania [10]), co wprowadza elementy nieliniowości oraz dodatkowe opóźnienia dynamiczne. Tego typu rozszerzenie znacząco zwiększa złożoność zadania regulacji i stanowi interesujące zagadnienie badawcze z zakresu sterowania układami niestabilnymi.

Rozbudowana konstrukcja mogłaby zostać wykorzystana do realizacji zaawansowanych ćwiczeń laboratoryjnych obejmujących analizę wpływu zmiany

parametrów obiektu na stabilność układu regulacji. Z punktu widzenia dydaktycznego, wprowadzenie platformy transportowej zwiększa realizm, przybliżając praktyczne zastosowanie robotów mobilnych w przemyśle, logistyce czy systemach autonomicznych. Balansowanie obiektu o zmiennym rozkładzie masy stanowi praktyczny przykład modelu stabilizacji w warunkach niepewności modelu, co pozwala na rozwijanie kompetencji w zakresie identyfikacji parametrów, strojenia regulatorów oraz analizy odporności układu na zakłócenia.

7. Podsumowanie

Opracowany robot dwukołowy jest przykładem zaawansowanego układu, który realizuje problem stabilizacji dynamicznej w czasie rzeczywistym. Konstrukcja oparta na zasadzie odwróconego wahadła wykorzystująca komputer Raspberry Pi oraz czujniki, umożliwia precyzyjne monitorowanie położenia robota i aktywne sterowanie jego ruchem.

Integralną częścią systemu jest algorytm sterowania, który łączy dane z akcelerometru i żyroskopu, filtruje je i oblicza sygnały sterujące napędem silników. Zaprojektowane podejście ma docelowo umożliwiać stabilizację robota oraz szybką reakcję na zmianę położenia, co stanowi przedmiot planowanych testów eksperymentalnych. Implementacja uwzględnia zarówno teoretyczne podstawy postaci ciągłej regulatora, jak i jego dyskretną realizację w środowisku cyfrowym, co umożliwia stabilną i szybką reakcję na zmiany położenia robota .

Literatura

[1] Helman Hanna, Sunnerman Henrik Two-Wheeled Self-Balancing Robot: Design and control based on the concept of an inverted pendulum.

[2] Marcinkowska Sandra, Robot balansujący - Co to takiego?, .https://botland.com.pl/blog/robot-balansujacy-co-to-takiego/, dostęp: 04.04.2026 r.

[3] Tutorial Books, Autodesk Fusion 360 For Beginners (June 2021)

[4] Nawawi S. W., Ahmad M. N., Osman J. H. S., Real-Time Control of a Two-Wheeled Inverted Pendulum Mobile Robot.

[5] grip_rz, Budowa robota balansującego - praktyczne porady, https://forbot.pl/blog/budowa-robota-balansujacego-praktyczne-porady-id9178, dostęp: 04.04.2026 r.

[6] AHRS, Complementary filter, https://ahrs.readthedocs.io/en/latest/filters/complementary.html, dostęp: 04.04.2026 r.

[7] Simon Monk, Programming with the Raspberry Pi.

[8] Wikipedia, Proportional–integral–derivative controller, https://en.wikipedia.org/wiki/Proportional%E2%80%93integral%E2%80%93derivative_controller, dostęp: 04.04.2026 r.

[9] Córdova-Esparza Diana-Margarita , Romero-Gonzalez Julio-Alejandro , Córdova-Esparza Karen-Edith , Terven Juan , López-Martinez Rocio-Edith , Active Learning Strategies in Computer Science Education: A Systematic Review,

[10] Wikipedia, Wind wave, https://en.wikipedia.org/wiki/Wind_wave, dostęp: 04.04.2026 r.

DESIGN AND CONTROL OF A SELF-BALANCING ROBOT

Summary

This scientific article presents a comprehensive process of designing, constructing, and programming a two-wheeled self-balancing robot based on the unstable inverted pendulum model. The mechanical structure is discussed in detail, featuring a compact enclosure designed for 3D printing, divided into two sections: a lower compartment housing the battery to reduce the center of gravity, and an upper compartment containing the Raspberry Pi control unit. A key aspect of the work is the implementation of a real-time control system using the Python programming language. Inertial sensors and a complementary filter are employed for precise estimation of the tilt angle, effectively reducing measurement errors. System stability is ensured by a discrete PID controller, which drives DC motors via a PWM signal. The robot serves as an educational tool in artificial intelligence courses, and future development plans include the addition of a transport platform to investigate the impact of variable load on control dynamics.

Keywords: robotics, self-balancing robot, Raspberry Pi, inverted pendulum.

Dominik Dworak, Julia Ziąbek, Maja Jaszowska,
Natalia Klocek, Sebastian Cwynar
Koło Naukowe Elektroniki i Technologii Informacyjnych

dr inż. Bartosz Pawłowicz
opiekun Koła Naukowego

IDENTYFIKACJA UKŁADÓW SCALONYCH BEZ PEŁNYCH OZNACZEŃ

Streszczenie

W rozdziale został opisany proces identyfikacji układów scalonych znalezionych w losowych urządzeniach elektronicznych, które z różnych powodów nie posiadają pełnego oznaczenia, pozwalającego na ich szybką identyfikację. Funkcje układów zostały ustalone na podstawie pomiarów, sąsiadujących elementów, obudowy i szczątkowemu oznaczeniu. Po identyfikacji układów zostały zmierzone ich parametry, między innymi maksymalna częstotliwość pracy, pobór prądu, pojemność wejściowa i czas wzrostu napięcia na wyjściu. Opisany jest sposób testowania układów oraz jak zidentyfikowano funkcję układu. Przedstawione metody i wskazówki mogą zostać wykorzystane do rozpoznawania powszechnych komponentów używanych w urządzeniach elektronicznych. Identyfikacja nieoznaczonych układów jest wymagana do przeprowadzenia naprawy urządzenia na wypadek awarii tego układu scalonego, szczególnie przy katastrofalnej awarii, której efektem może być wypalenie układu. W takich wypadkach można posłużyć się jedynie wskazówkami pochodzącymi z sąsiednich komponentów oraz rozłożenia ścieżek na płytce drukowanej. Znanie działania układu pozwala również na jego ponowne wykorzystanie w innym urządzeniu lub dobranie kompatybilnego zamiennika dla naprawianego przedmiotu. Umiejętność identyfikowania elementów elektronicznych jest kluczowa przy naprawie sprzętu elektronicznego, który w innym wypadku stałby się odpadem pogłębiającym globalny problem elektrośmieci oraz kryzys klimatyczny. W ramach rozdziału zostało zbadane 7 komponentów wymontowanych z wyrzuconych urządzeń elektronicznych. Wszystkie zostały wstępnie zidentyfikowane, po czym przeprowadzono dokładniejsze badania mające na celu potwierdzenie ich funkcji oraz pomiar parametrów tychże komponentów.

Słowa kluczowe: elektronika, pomiary, naprawa elektroniki, ekologia, planowane postarzanie, inżynieria wsteczna.

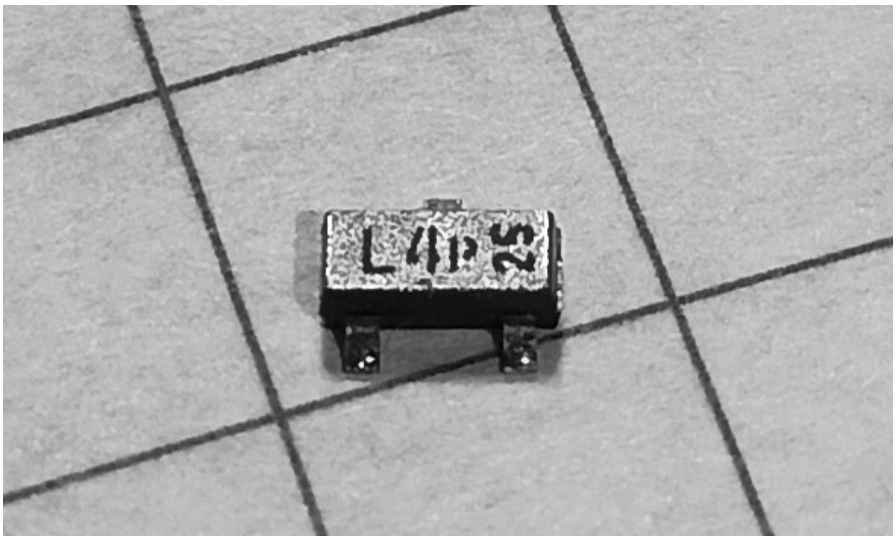
1. Wprowadzenie

Celem niniejszego rozdziału jest opisanie i wyjaśnienie metodologii oraz sposobów identyfikacji nieoznaczonych układów scalonych na podstawie ich obudowy, szczątkowego oznaczenia, pomiarów oraz sąsiadujących elementów. Znajomość modelu układu jest wymagana w przypadku wymiany uszkodzonego egzemplarza na nowy. Bez tej informacji nieoznaczony i uszkodzony układ, który w rzeczywistości może być szeroko dostępny, uniemożliwia działanie urządzenia, zmuszając właściciela do wymiany większej części (na przykład modułu) lub nawet całości urządzenia. Niektórzy producenci mogą celowo usuwać oznaczenia układów aby utrudnić lub uniemożliwić ich naprawę w sytuacji, kiedy układ ulegnie uszkodzeniu czego przykład jest widoczny na rysunku 2, a inny przypadek został opisany na blogu *Bunniestudios* [1]. Taka praktyka jest kojarzona z *planowanym postarzaniem produktów* (z ang. *Planned Obsolescence*) i może przyczyniać się do generowania elektrośmieci oraz pogłębiania kryzysu klimatycznego [2].

2. Powody nieoznaczania układów scalonych

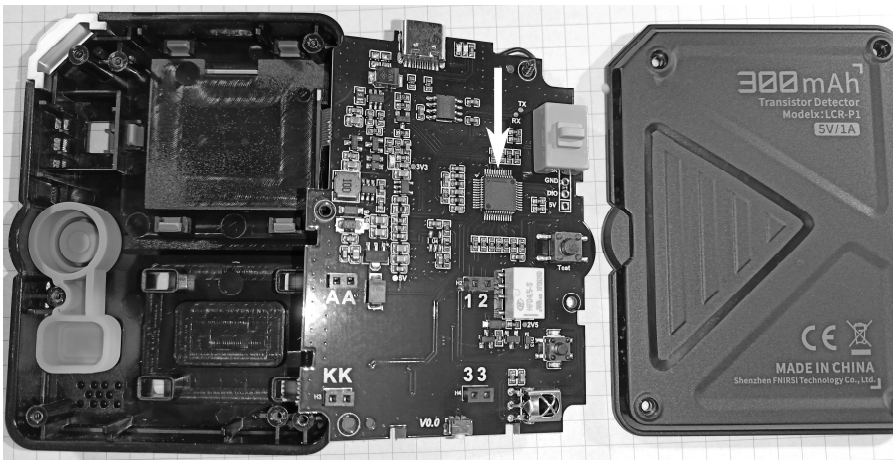
Poza wymienionym planowanym postarzaniem produktu istnieją inne powody braku oznaczenia układu, jednym z nich jest mały rozmiar. Układy w małych obudowach (np. SOT-23) posiadają bardzo małą powierzchnię, co utrudnia nadrukowanie pełnego oznaczenia modelu układu, zamiast tego producenci często wolą oznaczyć numer partii układu, bo ta informacja jest ważniejsza dla ich klientów (przykładowy układ SOT-23 pokazany na rysunku 1.). Czasami producenci urządzeń elektronicznych zamawiając duże wolumeny danego układu żądają nadrukowania wewnętrznego oznaczenia zamiast modelu układu. Takie oznaczenie ułatwia identyfikację na linii produkcyjnej, ale jest bezużyteczne dla użytkownika końcowego. Oznaczenie wewnętrzne występuje również przy układach zamawianych na specjalne zamówienie, które nie są w normalnej sprzedaży przez producenta układów. Producenci tacy jak Apple często stosują własne oznaczenia ogólnie dostępnych układów w celu łatwiejszej identyfikacji w fabryce, czego skutkiem ubocznym jest utrudnienie naprawy urządzenia.

Kolejnym powodem usuwania nadruków z układów jest próba ochrony przed kopiowaniem pomimo tego, że ta metoda nie jest efektywna, bo podmiot chcący skopiować urządzenie zazwyczaj ma dostęp do bardziej zaawansowanych narzędzi niż zwykły użytkownik (np. laser do usuwania obudowy układu, maszyna do prześwietlania) oraz wiedzy i doświadczenia pozwalających na identyfikację układu pomimo usunięcia oznaczenia. Ta praktyka jest szczególnie popularna wśród producentów taniej elektroniki (na rysunku 2. pokazano przykład usuniętego oznaczenia w urządzeniu producenta Fnirsi).



Rys. 1. Układ w obudowie SOT-23. Mały rozmiar nie pozwala na nadrukowanie pełnego oznaczenia, zamiast tego nadrukowany został kod partii lub kod wewnętrzny (L4p 25). Układ został sfotografowany na tle papieru z nadrukowaną kratką 5mm (dla skali)

Fig. 1. Device in a SOT-23 package. Small size makes it impossible to print the full marking, instead a lot number or an inside marking was printed (L4p 25). Device was photographed on paper with 5mm grid (for scale)



Rys. 2. Usunięte oznaczenie z mikrokontrolera w testerze komponentów Fnirsi LCR-P1 (największy układ scalony na środku rysunku). Równe boki, ostre rogi, niewielka ilość usuniętego materiału oraz tekstura powierzchni z której usunięto oznaczenie sugeruje usunięcie za pomocą lasera. Układ to najprawdopodobniej mikrokontroler STM32

Fig. 2. Microcontroller with removed markings inside a Fnirsi LCR-P1 component tester (biggest IC in the center). Straight edges, sharp corners, small amount of material removed and texture of the surface from which markings were removed suggest removal by laser. The IC is most propably an STM32 microcontroller

2.1. Nowoczesne sposoby identyfikacji układów

W przypadku nadrukowania numeru partii na układzie jednoznaczna identyfikacja jest stanowczo utrudniona i wymaga pomiarów. Jednakże w wielu przypadkach nadrukowane są skrócone wersje oznaczeń, co daje szansę na identyfikację układu za ich pomocą. W takim przypadku bardzo dobrym narzędziem są bazy kodów SMD. Zawierają one kody popularnych elementów pogrupowane na podstawie obudowy. Takie bazy mogą naprowadzić serwisanta lub inną osobę identyfikującą układ na prawdopodobny model, lecz z powodu ograniczonej ilości kombinacji znaków wynik z bazy układów nie zawsze będzie jednoznaczny. Przykładowymi wyszukiwarkami kodów SMD są strony internetowe *smd.yooneed.one* oraz *www.marsport.org.uk* .

Kolejną metodą, która stosunkowo niedawno stała się powszechnie dostępna jest identyfikacja układów za pomocą sztucznej inteligencji poprzez analizę zdjęcia układu lub podanie jego opisu (oznaczenie i obudowa). Jest to najprostsza do wykonania metoda, lecz halucynacje SI mogą wprowadzać w błąd, dlatego zawsze warto jest zweryfikować model układu innymi metodami, uwzględniając podpowiedź SI.

3. Narzędzia użyte do identyfikacji układów

Do wstępnej identyfikacji układów z trzema wyprowadzeniami (SOT-23) został użyty uniwersalny tester komponentów Fnirsi LCR-P1 [3] (widoczny otwarty na rysunku 2.). Tester pozwala na identyfikację takich komponentów jak:

- tranzystory bipolarne (NPN i PNP),
- tranzystory unipolarne (MOSFET, IGBT, JFET),
- diody (w tym zenera),
- tyrystory.

Tester dodatkowo może mierzyć rezystancję, pojemność i induktancję komponentów (rezystory, kondensatory i cewki). Po identyfikacji komponentu LCR-P1 mierzy również jego podstawowe parametry, między innymi:

- pojemność bramki tranzystora unipolarnego,
- wzmocnienie prądu stałego tranzystora bipolarnego,
- napięcie przewodzenia diody,
- napięcie załączenie tyrystora.

Te dane pozwalają na ocenienie zastosowania komponentu, na przykład czy tranzystor bipolarny jest ogólnego zastosowania lub jest przeznaczony do wysokich częstotliwości. Ten tester jest jednak ograniczony do prostych komponentów dyskretnych i nie może identyfikować bardziej zaawansowanych układów, a same wartości pomiarów mają nieokreśloną dokładność, gdyż producent nie udostępnia dokładnych specyfikacji urządzenia. Z tego powodu pomiary powinny być traktowane jako orientacyjne. Jeśli komponent z trzema wyprowadzeniami nie może zostać zidentyfikowany, to najczęściej jest on liniowym regulatorem napięcia.

Do dokładnego pomiaru oraz identyfikacji bardziej zaawansowanych komponentów zostało użyte stanowisko złożone z wymienionych przyrządów:

- zasilacz laboratoryjny KORAD KKG305P,
- oscyloskop cyfrowy Siglent SDS 824X HD,
- generator arbitralny UNI-T UTG932E.

Zasilacz laboratoryjny pozwala na bezpieczne zasilanie mierzonego układu dzięki jego funkcji ograniczenia prądu oraz krótkiemu czasu reakcji. Chroni to układy przed przepaleniem w przypadku odwrotnego podłączenia zasilania. Komponenty powinny pobierać maksymalnie kilka miliamperów prądu, ustawienie ograniczenia na 20 mA pozwala na działanie układu przy poprawnym podłączeniu i daje czas na wyłączenie zasilania przy odwróceniu polaryzacji. W tabeli 1. przedstawiono podstawowe parametry zasilacza KKG305P.

Tabela 1. Podstawowe parametry zasilacza KORAD KKG305P

Table 1. KORAD KKG305P power supply basic parameters

Parametr	Wartość	Jednostka
Napięcie wyjścia	0 - 30	V
Prąd wyjścia	0 - 5	A
Rozdzielczość ustawienia napięcia	10	mV
Rozdzielczość ustawienia prądu	1	mA

Źródło: nota katalogowa KKG305P [4].



Rys. 3. Zasilacz KORAD KKG305P

Fig. 3. KORAD KKG305P power supply

Do obserwacji i pomiaru przebiegów napięć użyto oscyloskopu Siglent SDS 824X HD. Posiada on przetwornik 12-bitowy o częstotliwości próbkowania 2GSa/s i szerokości pasma 200MHz (oscyloskop pokazano na rysunku 4. Oznaczenie modelu to SDS 804X HD, ale został on programowo ulepszony do wyższej wersji SDS 824X). Oscyloskop służy głównie do ustalenia dynamicznych parametrów identyfikowanych komponentów, takich jak maksymalna częstotliwość pracy i szybkość narastania zbocza napięcia. Jego główne parametry zostały przedstawione w tabeli 2.

Tabela 2. Parametry oscyloskopu Siglent SDS 824X HD

Table 2. Siglent SDS 824X HD oscilloscope parameters

Parametr	Wartość	Jednostka
Rozdzielczość przetwornika analogowego	12	Bit
Szerokość pasma (bandwidth)	200	MHz
Maksymalna szybkość próbkowania	2	GSa/s
Liczba kanałów	4	
Maksymalne przechwytywanie przebiegów	500 000	wfm/s
Maksymalna długość zapisu	100	Mpts

Źródło: oficjalna nota katalogowa producenta [5].



Rys. 4. Oscyloskop Siglent SDS 824X HD. Oznaczenie na obudowie nie zgadza się ze względu na programowe ulepszenie urządzenia do wyższej wersji

Fig. 4. Siglent SDS 824X HD oscilloscope. Markings do not match the given model number because of a software upgrade to a higher model

Do generowania przebiegów testowych użyto generatora arbitralnego UNI-T UTG932E (pokazany na rysunku 5.). Pozwala on między innymi na generowanie przebiegów sinusoidalnych do 30 MHz i prostokątnych do 20 MHz które posłużą do sprawdzenia maksymalnej częstotliwości pracy badanych komponentów. Szerokość pasma jest szczególnie ważną informacją przy projektowaniu urządzenia pracującego z wysokimi częstotliwościami (na przykład w zastosowaniach komunikacji radiowej). UTG932E posiada 2 kanały co umożliwia wykonanie zaawansowanych testów oraz wyróżnia się małym rozmiarem, zmniejszając wymagania wielkości przestrzeni laboratoryjnej. W tabeli 3. pokazano podstawowe parametry generatora.

Tabela 3. Podstawowe parametry generatora arbitralnego UNI-T UTG932E

Table 3. UNI-T UTG932E arbitrary waveform generator basic parameters

Parametr	**Wartość**	**Jednostka**
Częstotliwość próbkowania	200	MSa/s
Maksymalna częstotliwość generowanego przebiegu	30 (sinus) 20 (prostokąt) 10 (arbitralny)	MHz
Maksymalna amplituda napięcia sygnału	10 (≤ 10 MHz) 5 (≤ 30 MHz)	V
Zakres napięcia wyjściowego (sygnał + offset DC)	± 5 (50Ω) ± 10 (Hi-Z)	V
Modulacja	AM, FM, PM, FSK	

Źródło: oficjalna nota katalogowa producenta [6].



Rys. 5. Generator arbitralny UNI-T UTG932E

Fig. 5. UNI-T UTG932E arbitrary waveform generator

Poza przedstawionymi przyrządami laboratoryjnymi został również użyty uniwersalny programator XGecu T48 (pokazany na rysunku 6.) pozwalający na

odczytanie wszelkiego rodzaju układów pamięci. Został on wykorzystany, ponieważ jeden z identyfikowanych układów okazał się pamięcią EEPROM z interfejsem I^2C. Programator pozwolił na potwierdzenie funkcji identyfikowanego układu scalonego.



Rys. 6. Programator uniwersalny XGecu T48
Fig. 6. XGecu T48 universal programmer

4. Sposoby oraz wskazówki na identyfikację układów

Najbardziej oczywistym i pewnym sposobem na identyfikację układu scalonego jest wyszukanie jego oznaczenia w specjalnych wyszukiwarkach przeznaczonych dla komponentów elektronicznych (na przykład *www.alldatasheet.com*). Niestety w niektórych przypadkach nie jest to możliwe, bo oznaczenie jest wewnętrznym kodem producenta lub zostało usunięte. To jest główne zagadnienie tego rozdziału, czyli jak zidentyfikować układ bez oczywistego oznaczenia. Oto lista aspektów pomagających w identyfikacji komponentów:

- obudowa komponentu,
- sąsiadujące elementy pasywne (kondensatory, cewki, rezystory),
- większe układy w pobliżu identyfikowanego komponentu,
- połączone ścieżki i ich grubość,
- podłączenie linii zasilających,
- przebiegi napięć na nóżkach.

Najłatwiejsze do identyfikacji są komponenty w obudowach SO-8 / SOIC-8 oraz SOT-23 ze względu na niewielką liczbę różnych funkcji oraz niską liczbę nóżek. W obudowach SOT-23 zazwyczaj znajdują się:

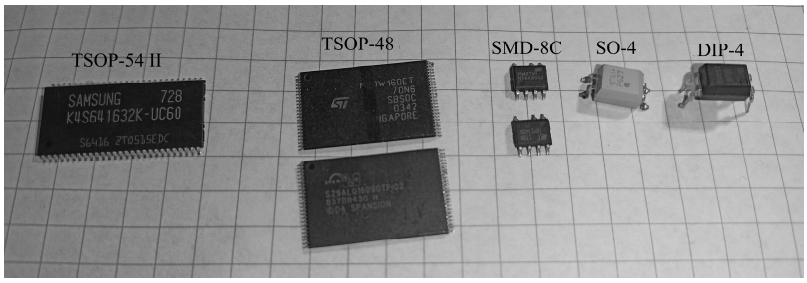
- tranzystory bipolarne i unipolarne,
- diody,
- regulatory liniowe.

Do ich identyfikacji wystarczy użycie testera komponentów (na przykład opisany wcześniej Fnirsi LCR-P1), jeśli nie zidentyfikuje on komponentu to najprawdopodobniej jest to regulator liniowy. Obudowy SO-8 / SOIC-8 mają większą różnorodność zawartych w nich układów, ale dalej są one relatywnie łatwe do identyfikacji:

- przetwornice napięcia - w ich pobliżu zazwyczaj będzie znajdowała się cewka, dioda i kondensatory,
- wzmacniacze operacyjne - jeśli układ znajduje się np. między złączem i przetwornikiem analogowo-cyfrowym; dookoła będą rezystory,
- pamięć FLASH / EEPROM - praktycznie zawsze znajdują się w pobliżu większego układu (mikroprocesor, mikrokontroler, FPGA) i nie mają w pobliżu wielu elementów pasywnych,
- tranzystory MOSFET - będą znajdować się w pobliżu złącz zasilających lub cewek, a nóżki będą połączone w 3 grupy (1 nóżka - bramka, 3 nóżki - źródło, 4 nóżki - dren).

Takie same zasady dotyczą podobnych obudów z trzema (TO-220, TO-223) i ośmioma wyprowadzeniami (TSSOP-8). Czasami w obudowach z ośmioma wyprowadzeniami mogą znajdować się bramki logiczne (NOT albo przerzutniki Schmitta), można je rozpoznać po tym, że po zasileniu ich 3 nóżki będą miały niestabilne napięcia (są to wyjścia), a pozostałe 3 nie będą miały napięcia lub będą miały zakłócenia przypominające sinusoidę o częstotliwości 50 Hz (wejścia z wysoką impedancją). Warto jeszcze wspomnieć o specjalnych obudowach (pokazane na rysunku 7.), które prawie zawsze zawierają układy o specyficznych funkcjach:

- DIP-6 / DIP-4 / SO-6 / SO-4 (obudowy 4 lub 6 pinowe) - prawie zawsze zawierają optocoupler,
- DIP-8B / DIP-8C / SMD-8C / SMD-8B (wbrew nazwie mają tylko 7 nóżek) - są to układy przetwornicy impulsowej napięcia sieciowego,
- TSOP-*x* / TSOP-*x* II (*x* to ilość nóżek) - są prawie zawsze pamięciami (RAM lub FLASH).

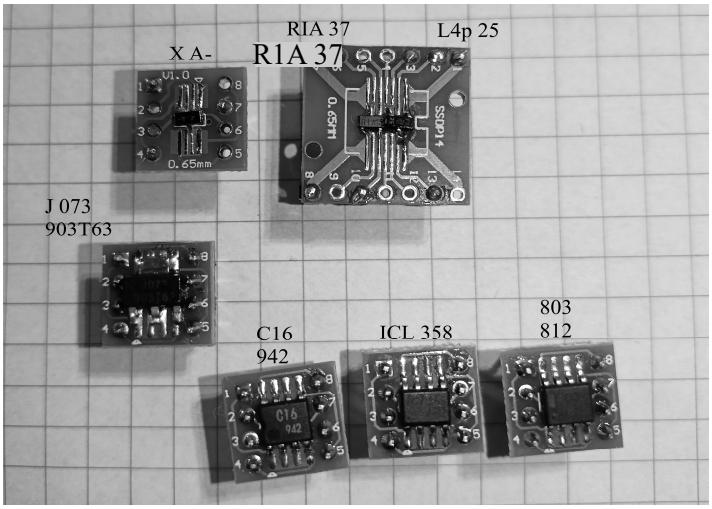


Rys. 7. Układy w unikalnych obudowach. Od lewej: pamięć RAM/FLASH w obudowie TSOP54II, pamięci RAM/FLASH w obudowach TSOP48, regulatory impulsowe w obudowach SMD-8C, optocoupler w obudowie SO-4 i optocoupler w obudowie DIP-4. TSOP i TSOP II różnią się ułożeniem nóżek wzdłuż długiej lub krótkiej krawędzi układu

Fig. 7. ICs with unique packages. From left: RAM/FLASH memory in a TSOP54II package, RAM/FLASH memories in TSOP48 packages, switching regulators in SMD-8C packages, optocouplers in SO-4 and DIP-4 packages. TSOP and TSOP II packages differ in pin placement on either the long edge or short edge of the device

5. Identyfikacja przykładowych układów

Przygotowano 7 niezidentyfikowanych układów (pokazane na rysunku 8.) bez wyraźnych oznaczeń, które na cele tego rozdziału zostaną zbadane. Najpierw na podstawie opisanych wyżej wskazówek i reguł zostaną odkryte ich funkcje, następnie układy zostaną zmierzone dokładniej aby ustalić ich parametry.



Rys. 8. Badane układy przylutowane do płytek adaptacyjnych. Dwa z nich musiały zostać przylutowane do pojedynczej płytki ze względu na brak ich wystarczającej ilości. Na rysunku zostały również podane opisy na układach

Fig. 8. Researched ICs soldered to adapter boards. Two of them had to be soldered on a single board due to a shortage of adapter boards. IC markings have been added to the photo

Ze względu na to, że wszystkie układy mają obudowy do montażu powierzchniowego (SMD) zostały one przylutowane do płytek adaptacyjnych, które pozwolą na wpięcie ich do płytki stykowej, co ułatwi pomiary.

5.1. Układ *X A-*

Pierwszym układem który zostanie zidentyfikowany jest *X A-* w obudowie SOT-23. Jest to prawie na pewno tranzystor lub dioda ze względu na jego obudowę. SOT-23 nie pozwala na odłożenie dużej mocy na elemencie, dlatego zazwyczaj są to tranzystory małych mocy, diody referencyjne (zenera) lub diody zabezpieczające, a rzadziej regulatory liniowe. Do identyfikacji zostanie użyty tester komponentów LCR-P1. Pomiar pokazano na rysunku 9.

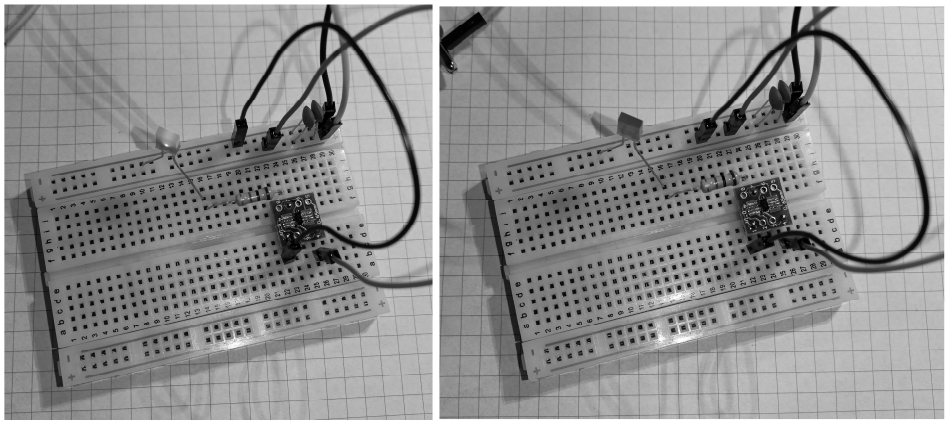


Rys. 9. Wykonanie pomiaru komponentu oznaczonego jako *X A-* testerem LCR-P1. Tester wskazał komponent jako tranzystor unipolarny (MOSFET) z kanałem P

Fig. 9. Measurement of the *X A-* marked component using the LCR-P1 device tester. The tester identified the component as a unipolar transistor (MOSFET) with a P channel

Zgodnie ze wstępnym założeniem tester oznaczył komponent jako P-MOSFET. Aby dodatkowo zweryfikować działanie komponentu został utworzony prosty obwód, gdzie MOSFET będzie sterować świeceniem diody LED. Napięcie testowe wynosi 5V, a rezystor ograniczający prąd diody ma wartość 1kΩ.

P-MOSFET zostanie wykorzystany w roli klucza górnego, czyli jego źródło zostanie podłączone do napięcia zasilającego, a reszta układu (dioda LED i rezystor) zostaną podłączone za drenem tranzystora. Bramka będzie zwierana do masy lub napięcia zasilającego. Podłączenie do 5V powinno wyłączyć tranzystor, a podłączenie do masy powinno skutkować włączeniem tranzystora i zaświeceniem diody LED.



Rys. 10. Obwód testujący komponent *X A*-, który został zidentyfikowany jako P-MOSFET. Po lewej stronie bramka układu została podłączona do 5V, co skutkowało wyłączeniem tranzystora. Po prawej stronie pokazano obwód, gdzie bramka jest podłączona do masy, co załącza P-MOSFET oraz diodę LED

Fig. 10. Circuit used for testing the *X A*- device identified as a P-MOSFET. On the left the device's gate is connected to 5V, which turns off the transistor. On the right the gate is connected to ground, which turns the P-MOSFET and an LED diode on

Komponent *X A*- zachował się zgodnie z przewidywaniami co potwierdza, że jest to P-MOSFET. Na koniec został wykonany dodatkowy pomiar napięcia przewodzenia diody wewnętrznej układu z powodu nieprawidłowego pomiaru na testerze. Zmierzył on spadek napięcia diody w kierunku przewodzenia jako 0V, a w rzeczywistości wynosi ono 0,585V (dokładność UT33A+ dla tego pomiaru to ±0,5% + 2 cyfry, dając nieznaczną niepewność ±4mV), zgodnie ze spadkiem diody krzemowej. Zmierzone parametry wskazują na P-MOSFET do zastosowań logicznych oraz przełączania obciążeń o niskiej mocy. Finalne parametry komponentu przedstawiono w tabeli 4.

Tabela 4. Zmierzone parametry P-MOSFET oznaczonego *X A*-

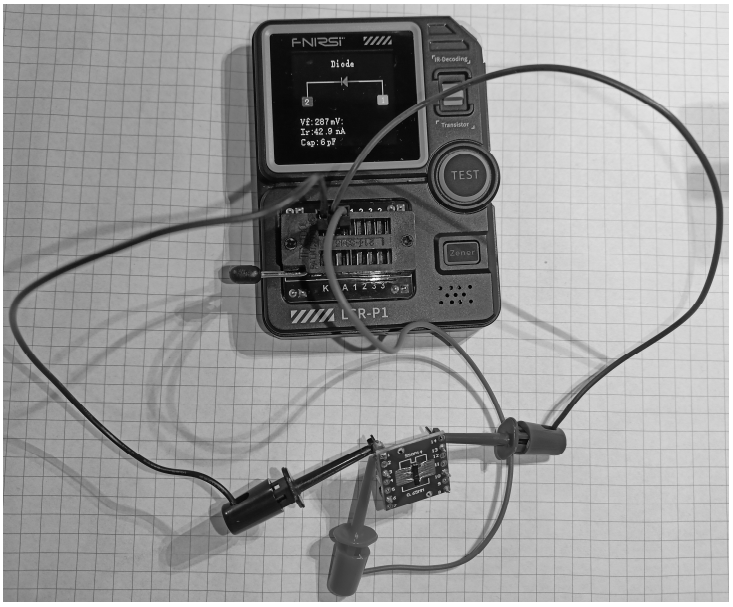
Table 4. Measured parameters of the *X A*- P-MOSFET

Parametr	Wartość	Jednostka
Napięcie włączenia	1,64	V
Rezystancja w stanie włączenia	8,2	Ω
Pojemność bramki	6	pF
Spadek napięcia w kierunku przewodzenia diody wewnętrznej	585	mV

Ze zmierzonych parametrów można wyliczyć ładunek bramki, czyli pojemność bramki pomnożona razy napięcie bramka-źródło, co daje wynik około 9,6pC. Jest to wartość mało prawdopodobna, gdyż jedne z lepszych tranzystorów obecnie dostępnych (np. *DMP22D5UFO-7B*) posiadają znacznie gorsze parametry ładunku bramki (300pC przy 5V). Powodem jest nieznana dokładność pomiarowa testera oraz jego nieznany algorytm pomiaru, co uniemożliwia poprawne porównanie wyników do danych z realnych specyfikacji.

5.2. Układ *L4p 25*

Kolejnym identyfikowanym komponentem będzie układ oznaczony jako *L4p 25* w obudowie SOT-23. Do pomiarów został użyty LCR-P1, który zidentyfikował komponent jako diodę o spadku napięcia 0,287V (pomiar pokazano na rysunku 11.). Wykonano dodatkowe pomiary dokładniejszym multimetrem UNI-T UT33A+, który zmierzył spadek napięcia w kierunku przewodzenia jako 0,208V. Nóżka (według standardowych numeracji SOT-23) numer 1 jest anodą, natomiast nóżka 3 jest katodą, a nóżka numer 2 jest niepodłączona. Spadek napięcia wynoszący około 0,2V wskazuje na diodę schottkiego, a obudowa SOT-23 oznacza zastosowanie do małych mocy, gdyż nie jest ona w stanie odprowadzić dużych ilości ciepła.



Rys. 11. Pomiar komponentu oznaczonego *L4p 25*. Został zidentyfikowany jako dioda o spadku napięcia 0,287V

Fig. 11. Measurement of the component marked as *L4p 25*. The device was identified as a diode with a forward voltage drop of 0,287V

Wyszukanie oznaczenia L4p w katalogu kodów SMD zwróciło diodę schottkyego o oznaczeniu *BAT54*. W tabeli 5. porównano zmierzone parametry z parametrami znajdującymi się w specyfikacji *BAT54* od producenta STMicroelectronics. Z wyjątkiem prądu upływu, dane są bardzo podobne i mierzony element jest najprawdopodobniej właśnie diodą *BAT54*.

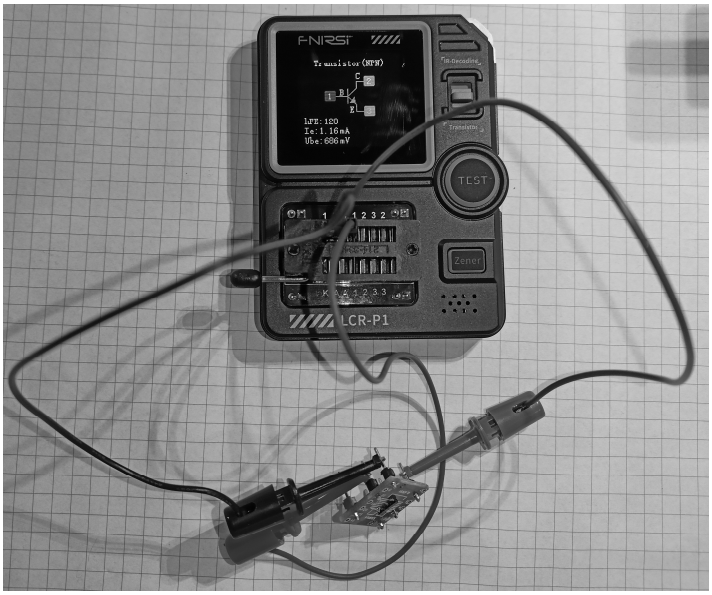
Tabela 5. Porównanie zmierzonych parametrów *L4p* oraz specyfikacji *BAT54*

Table 5. Comparison of measured parameters of *L4p* and *BAT54* specification

Parametr	L4p	BAT54	Jednostka
Napięcie przewodzenia	208	240 @ 100µA	mV
Prąd upływu	0,0429	1 ~ 100	µA
Pojemność diody	6	7	pF

5.3. Układ *R1A 37*

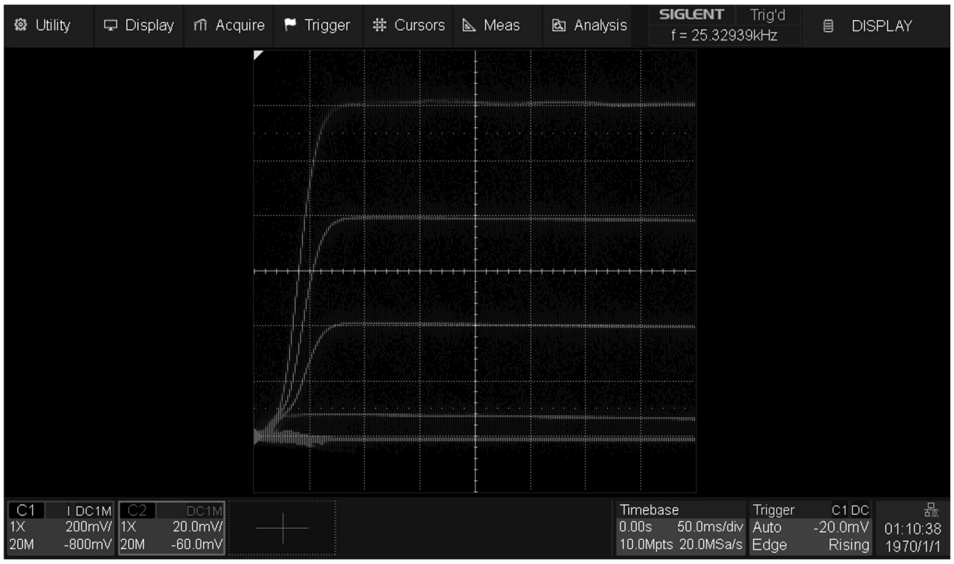
Następnym komponentem który zostanie zidentyfikowany jest układ oznaczony *R1A 37* w obudowie SOT-23. Tester LCR-P1 zidentyfikował go jako tranzystor NPN ze standardowym wyprowadzeniem nóżek (ukazane na rysunku 12.), czyli baza na nóżce 1, emiter na nóżce 2 i kolektor na nóżce 3 (według standardowej numeracji SOT-23, tester był podłączony inaczej przez co wskazania są inne).



Rys. 12. Pomiar układu R1A 37 testerem LCR-P1. Został zidentyfikowany jako tranzystor bipolarny NPN

Fig. 12. Measurement of the R1A 37 marked device using the LCR-P1. It was identified as a bipolar NPN transistor

Zmierzona wartość wzmocnienia prądu stałego *hFE* wynosi 128, a napięcie przewodzenia baza-emiter *Ube* 686mV. Wartości te są typowe dla tranzystorów małej mocy ogólnego zastosowania. Aby móc dokładniej scharakteryzować komponent, został wykonany pomiar krzywej I-V (wynik pokazany na rysunku 13.) za pomocą analizatora krzywych tranzystora.



Rys. 13. Wykres krzywej I-V tranzystora R1A 37 wykonany za pomocą prostego analizatora zbudowanego z dyskretnych komponentów. Charakter analizatora nie pozwala na dokładne określenie prądu bazy bez dodatkowych pomiarów. Prąd kolektora jest mierzony jako napięcie odłożone na rezystorze o wartości 22Ω. Oś X reprezentuje napięcie kolektor-emiter, natomiast oś Y prąd kolektora

Fig. 13. *R1A 37* transistor I-V curve measured using a simple curve tracer built with discrete components. The character of the tracer doesn't allow to accurately define base current without additional measurements. Collector current is measured as voltage drop across a 22Ω resistor. X axis represents collector-emitter voltage, Y axis represents collector current

Na rysunku 13. widoczna jest krzywa I-V tranzystora zmierzona za pomocą prostego analizatora zbudowanego na podstawie schematu dostępnego w internecie [7]. Prosta i niedokładna budowa analizatora nie pozwala na określenie prądu bazy mierzonego tranzystora dla poszczególnych linii bez dodatkowych pomiarów. Analizator jest jednak wystarczający do zaobserwowania kształtu krzywych. Prąd kolektora jest mierzony jako spadek napięcia na rezystorze pomiarowym o wartości 22Ω.

Po wyszukaniu oznaczenia *R1A* w bazie komponentów SMD zostały podane dwa prawdopodobne modele tranzystora; *MMBTH10* oraz *SST3904*, przy czym po sprawdzeniu dokumentacji obu okazało się, że tylko *SST3904* ma oznaczenie *R1A*

(*MMBTH10* jest oznaczany jako *K3Y*). *SST3904* jest tranzystorem ogólnego zastosowania, co zgadza się z przedstawioną klasyfikacją, a zmierzone parametry mieszczą się w specyfikacji *SST3904*, co porównano w tabeli 6.

Tabela 6. Porównanie zmierzonych parametrów *R1A* oraz specyfikacji *SST3904*

Table 6. Comparison of measured parameters of *R1A* and *SST3904 specification*

Parametr	R1A	SST3904	Jednostka
Wzmocnienie DC (h_{FE})	128 @ 1,6mA	>70 @ 1mA	-
		100 ~ 300 @ 10mA	
Napięcie nasycenia (V_{BE})	686	650 ~ 850	mV

5.4. Układ C16 942

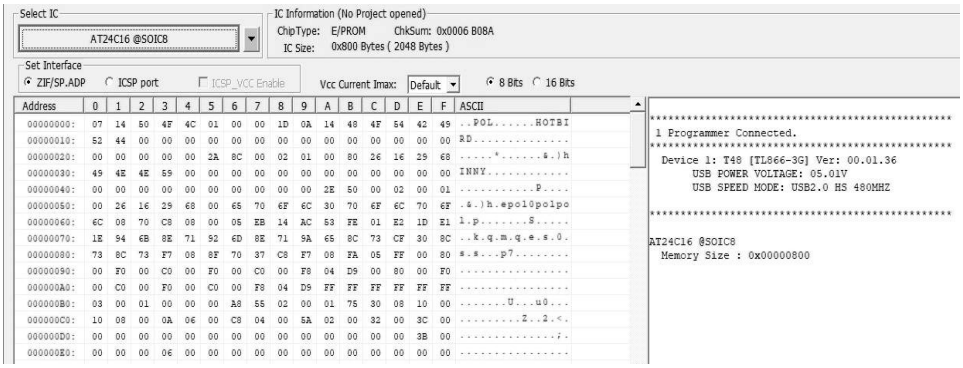
Następnym identyfikowanym komponentem będzie układ w obudowie SO-8 z oznaczeniem *C16 942*. Obudowa SO-8 sugeruje przetwornicę DC/DC, MOSFET, wzmacniacz operacyjny lub pamięć szeregową. Pierwszym testem jest sprawdzenie, czy nóżki są połączone w 3 grupy tak jak w MOSFETach. Po sprawdzeniu multimetrem w trybie testu ciągłości połączenia nóżek, nie znaleziono żadnego połączenia. Kolejnym krokiem jest znalezienie wyprowadzeń zasilających, zazwyczaj pin 4 jest połączeniem ujemnym, a pin 8 dodatnim. W przypadku tego układu nie można zidentyfikować nóżek zasilających na podstawie płytki z której został wymontowany z powodu braku oryginalnej płytki.

Do zweryfikowania połączenia ujemnego można użyć testu diody multimetru. Do nóżki 4 przykładana jest dodatnia sonda, a do pozostałych ujemna. Multimetr wskazuje na każdej nóżce spadek napięcia około 0,6V, co odpowiada spadkowi napięcia diody krzemowej. Sugeruje to odwrotne podłączenie, czyli nóżka 4 jest wyprowadzeniem GND. Układ ma najprawdopodobniej standardowy układ nóżek zasilających (tak jak wspomniano wcześniej). Układ został zasilony napięciem 5V z ograniczeniem prądu ustawionym na 20mA, plus podłączono do pinu 8, a minus do pinu 4. Zasilacz nie wskazywał żadnego poboru prądu po włączeniu zasilania, prawdopodobnie przez ograniczenie dokładności wewnętrznego pomiaru zasilacza.

Pomiary nóżek oscyloskopem wskazały niskie napięcie w okolicy 0V. uwagę zwraca lekka niestabilność napięcia, która wygląda jak sinusoida o częstotliwości 50 Hz. Sugeruje to wysoką impedancję wyprowadzeń układu. Wzmacniacz operacyjny posiada wejścia o wysokiej impedancji, ale ma również wyjścia, które wyróżniały by się niestabilnym napięciem wynikającym z niepodłączenia wejść (*pływanie wejścia*). Jednolite napięcia na wyprowadzeniach sugerują, że badany układ jest cyfrowy i jest prawdopodobnie pamięcią szeregową. W tym przypadku oznaczenie *C16* może oznaczać skrócenie oznaczenia 24Cx lub 25Cx, gdzie x oznacza pojemność układu, a 24 lub 25 oznacza protokół komunikacyjny (SPI lub I^2C). Pojemność układu to zawsze potęga 2, co pasuje do oznaczenia *C16* i oznaczałoby

pojemność 16kb. Aby sprawdzić czy identyfikowany układ jest pamięcią szeregową o pojemności 16kb został on podłączony do uniwersalnego programatora XGecu T48. Użycie programatora w tym przypadku jest zalecane, ponieważ prawdziwa funkcja jest nieznana, a teoria o układzie pamięci jest tylko założeniem i może być nieprawidłowa. Programator posiada wbudowane zabezpieczenia przeciwprądowe i przeciwzwarciowe oraz wykonuje test układu przed każdym odczytem. Dzieki temu w razie prawdopodobnej pomyłki nie dojdzie do uszkodzenia.

W programie obsługującym programator został początkowo wybrany ogólny profil dla pamięci 24C16. Programator z powodzeniem sprawdził układ i odczytał z niego dane (pokazane na rysunku 14.).



Rys. 14. Okno programu XGpro służącego do obsługi programatora XGecu T48 po odczytaniu danych z identyfikowanego układu *C16 942*

Fig. 14. Window of the XGpro software used to contol the XGecu T48 programmer after reading data from the device under test *C16 942*

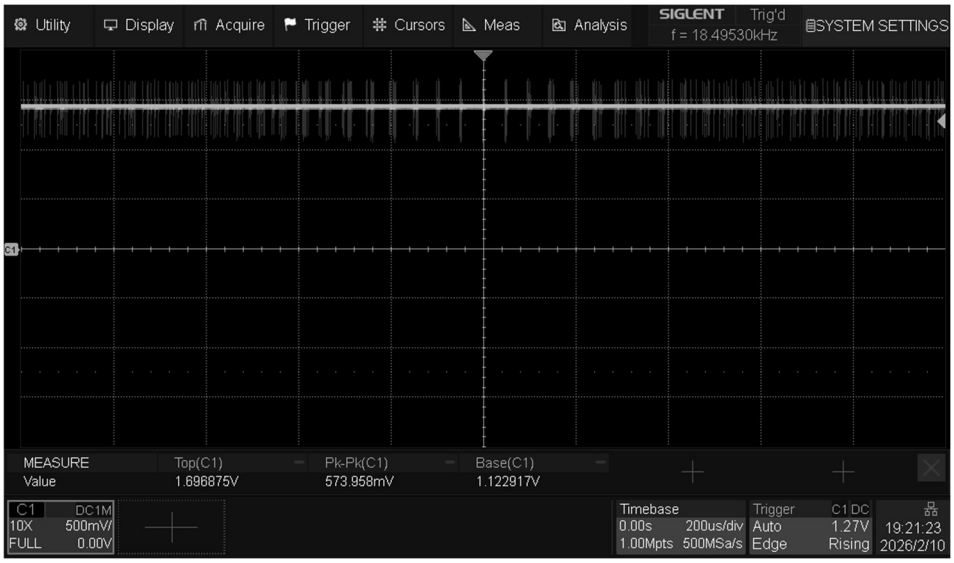
W odczytanych danych można zobaczyć napis *HOTBIRD*. Jest to nazwa serii satelit telekomunikacyjnych, które nadają między innymi sygnały telewizyjne dla Polski. Te dane potwierdzają, że układ jest pamięcią szeregową 24Cx oraz dodatkowo, że pochodzi on z dekodera do telewizji satelitarnej. Aby dodatkowo zweryfikować pojemność został wybrany model pamięci o większej pojemności, a po odczytaniu dane zaczęły się zapętlać, co potwierdza pojemność 16kb.

5.5. Układ 803 812

Kolejnym identyfikowanym układem będzie *803 812*. Ma on obudowę SO-8 i pochodzi z płyty przemysłowego sterownika lasera firmy *Fanuc*. Układ znajduje się między złączami z wieloma stykami a konwerterami analogowo-cyfrowymi i cyfrowo-analogowymi. Jest to najprawdopodobniej sekcja analogowa płytki, gdzie odczytywane są napięcia z różnych czujników. Sugeruje to, że identyfikowany układ jest wzmacniaczem operacyjnym, który działa jako bufor napięcia, lub innym układem analogowym. Teoria o wzmacniaczu operacyjnym jest najbardziej

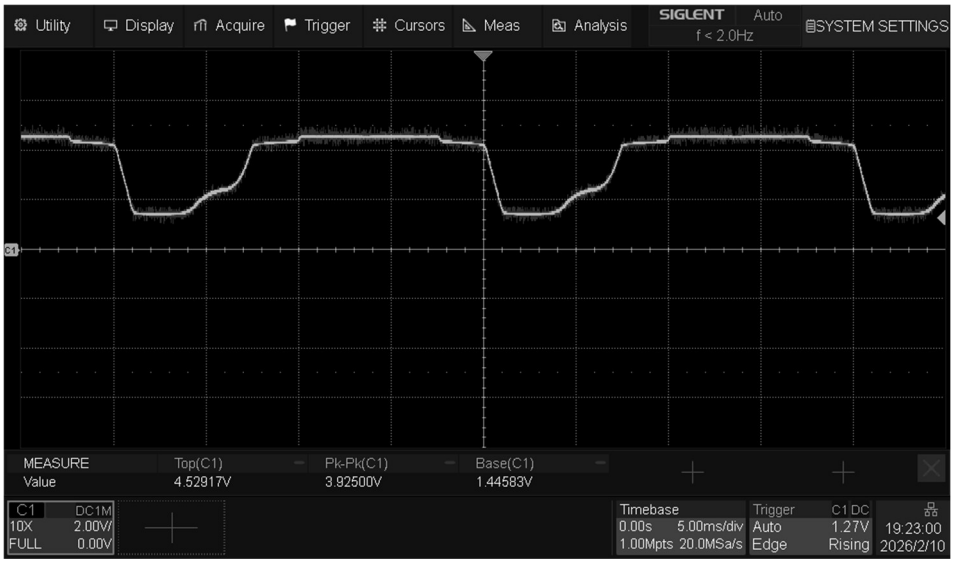
prawdopodobna, lecz należy zachować ostrożność przy testach, aby nie uszkodzić układu.

W pierwszej kolejności wykonano proste pomiary, które nie wprowadzają ryzyka uszkodzeń. Zostało sprawdzone połączenie nóżek układu do linii zasilających obecnych na płyty. Układ był podłączony do zasilania symetrycznego, gdzie -15V było podłączone do wyprowadzenia 4, a +15V do wyprowadzenia 8. Tak szeroki zakres zasilania oraz obecność wielu rezystorów wokół układu praktycznie potwierdza, że badany układ jest wzmacniaczem operacyjnym. Do finalnej weryfikacji został wykonany pomiar nóżek układu po zasileniu go napięciem 5V (GND podłączono do pinu 4, a 5V do pinu 8). Układ nie pobierał nadmiernego zasilania (do 2 mA), a na nóżkach układu pojawiły się przebiegi pokazane na następujących rysunkach (rysunki od 15. do 18.).



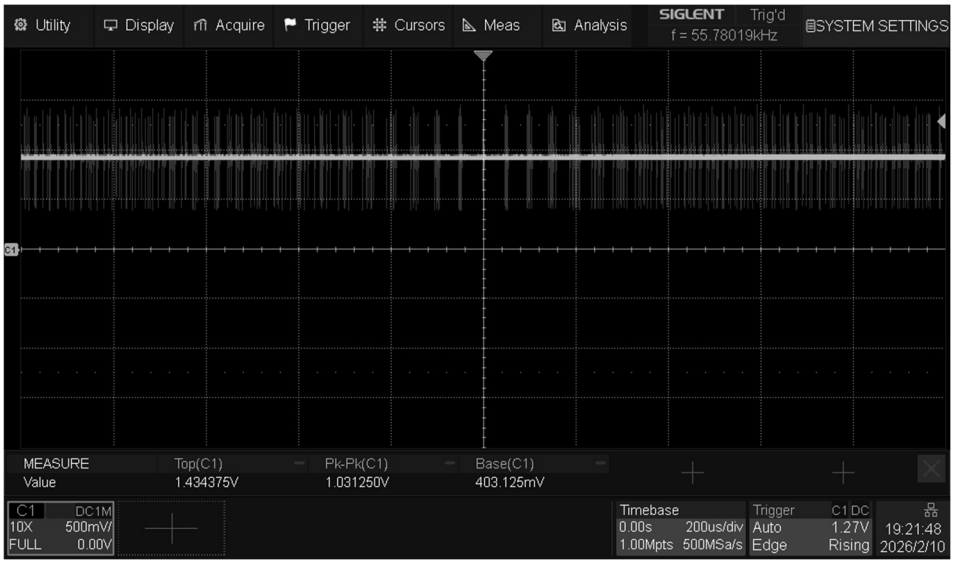
Rys. 15. Przebieg napięcia na nóżce nr. 1 układu

Fig. 15. Voltage waveform on IC pin number 1



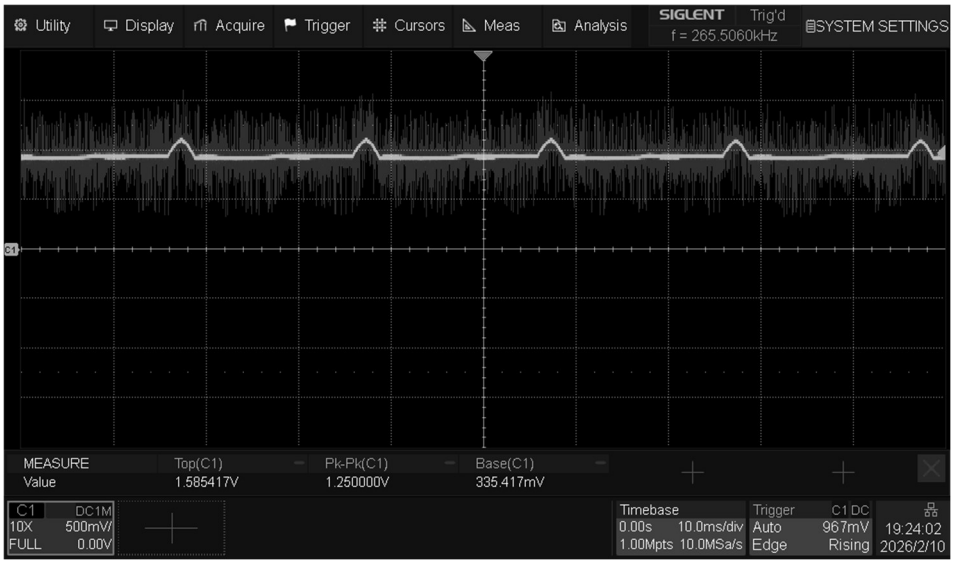
Rys. 16. Przebieg napięcia na nóżce nr. 7 układu

Fig. 16. Voltage waveform on IC pin number 7



Rys. 17. Przebieg napięcia na nóżkach nr. 2 i 3 układu

Fig. 17. Voltage waveform on IC pins number 2 and 3



Rys. 18. Przebieg napięcia na nóżkach nr. 5 i 6 układu

Fig. 18. Voltage waveform on IC pins number 5 and 6

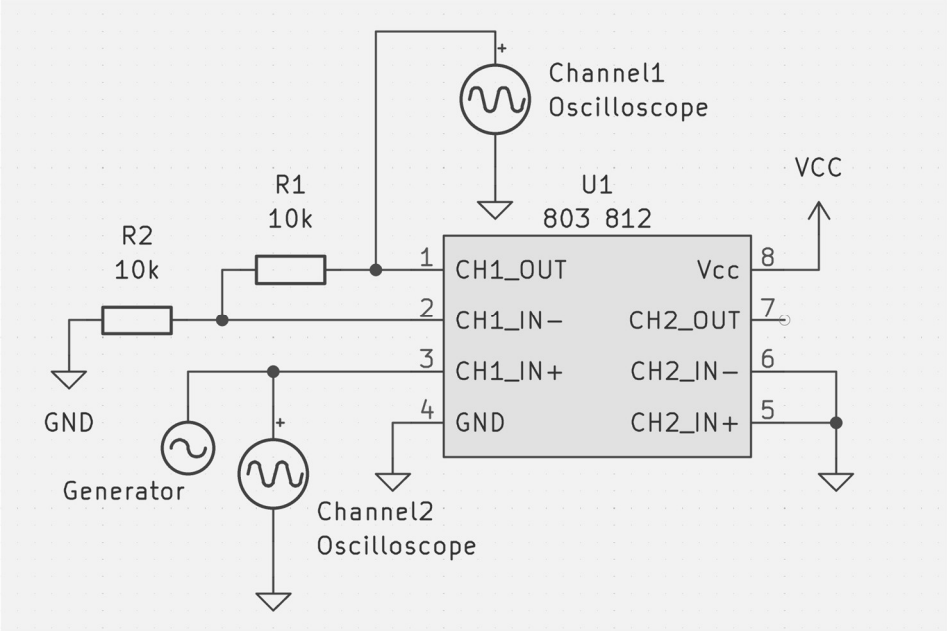
Na powyższych rysunkach wykresów widać, że nóżki 1 i 7 mają unikalne przebiegi, a nóżki 2 i 3 oraz 5 i 6 mają praktycznie takie same przebiegi. Dodatkowo przebiegi na nóżkach 2 i 3 oraz przebieg nóżki 1 są stałe, natomiast przebiegi nóżek 5 i 6 oraz przebieg nóżki 7 zawierają zniekształcenia powtarzające się z tą samą częstotliwością. Ten fakt grupuje nóżki 1, 2 i 3 w jedną grupę oraz 5, 6 i 7 w drugą grupę. Takie pogrupowanie nóżek pokrywa się ze standardowym wyprowadzeniem pinów wzmacniaczy operacyjnych (np. LM358), gdzie nóżki mają funkcje pokazane w tabeli 7.

Tabela 7. Funkcje nóżek w standardowym układzie wzmacniacza operacyjnego

Table 7. Pin functions of a standard pinout operational amplifier

Numer nóżki	**Funkcja**
1	Wyjście kanału 1
2	Wejście odwracające kanału 1
3	Wejście nieodwracające kanału 1
4	GND
5	Wejście nieodwracające kanału 2
6	Wejście odwracające kanału 2
7	Wyjście kanału 2
8	Vcc

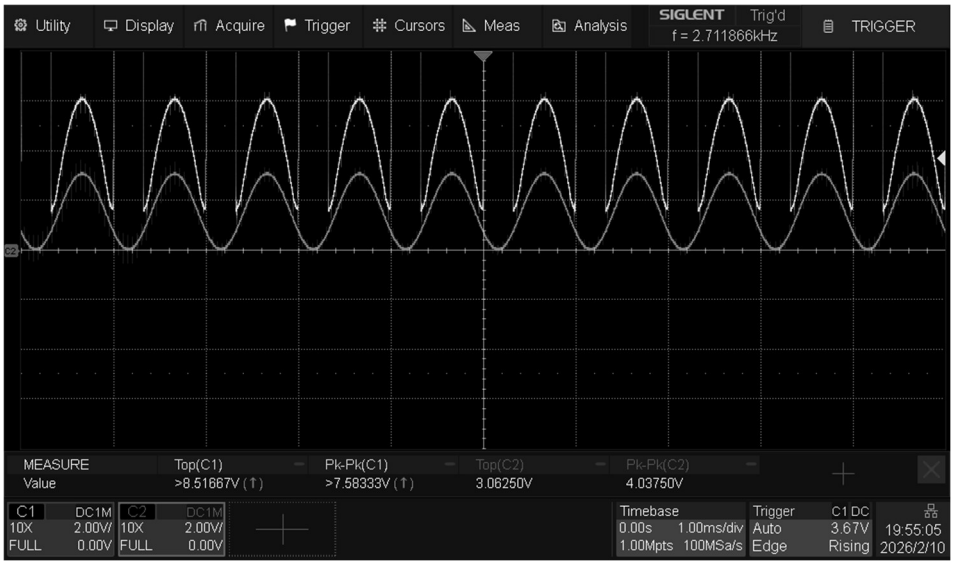
Aby całkowicie zweryfikować funkcję sprawdzanego komponentu *803 812* został zbudowany układ elektroniczny na płytce stykowej konfigurujący identyfikowaną część jako nieodwracający wzmacniacz operacyjny ze wzmocnieniem równym 2. Do ustawienia wzmocnienia użyto rezystorów 10 kΩ. Na rysunku 19. pokazano schemat ideowy układu testowego.



Rys. 19. Schemat ideowy układu testującego identyfikowany komponent
Fig. 19. Test circuit schematic for the device under test

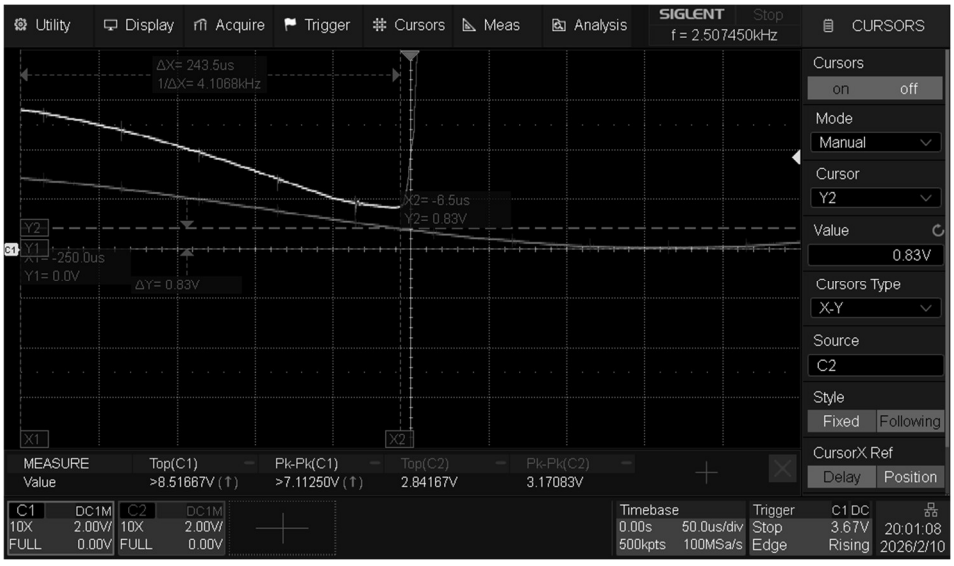
Za pomocą zbudowanego schematu przetestowano działanie komponentu. Napięcie zasilające *VCC* wynosi 10V. Z generatora podawano różne przebiegi mające na celu sprawdzenie działania komponentu jako wzmacniacza operacyjnego. Na rysunku 20. pokazano przebieg testowy (C2) oraz wyjściowy (C1).

Na rysunku 20. wyraźnie widać podwójne wzmocnienie przebiegu wchodzącego co potwierdza w 100%, że badany układ to wzmacniacz operacyjny. Zniekształcenie przebiegu w dolnej części sinusoidy jest wynikiem ograniczonego zakresu napięcia wejściowego. Niektóre wzmacniacze operacyjne (głównie starsze modele) nie pracują prawidłowo przy niskich napięciach wejściowych. Na rysunku 21. pokazano pomiar tego napięcia.



Rys. 20. Przebiegi napięć podczas testu komponentu. Ciemniejszy przebieg (C2) to przebieg wejściowy pochodzący z generatora arbitralnego, a jaśniejszy (C1) to przebieg wyjściowy z testowanego komponentu

Fig. 20. Voltage waveforms during device testing. Darker waveform (C2) represents the input signal coming from the arbitrary waveform generator, lighter waveform (C1) is the output waveform from the device under test

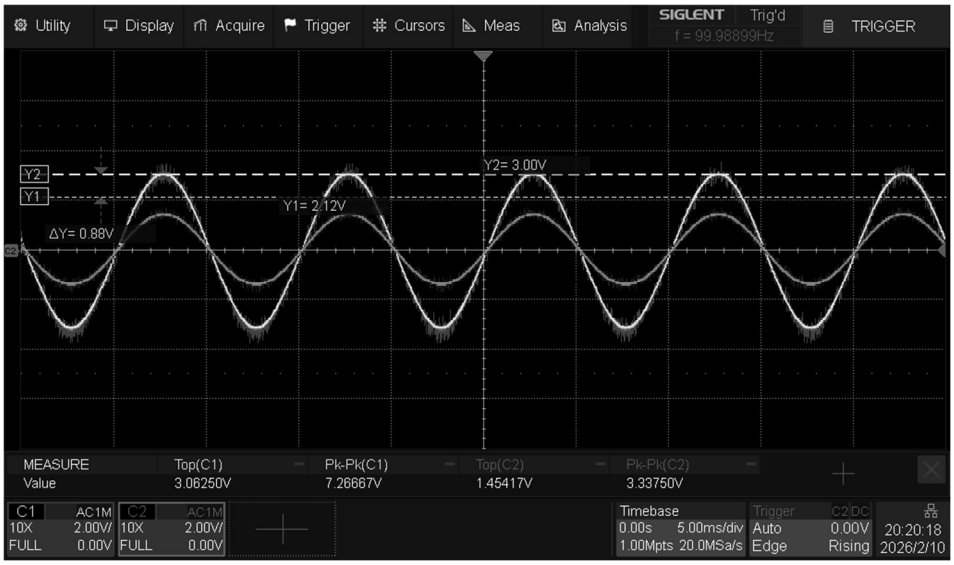


Rys. 21. Pomiar minimalnego napięcia wejściowego wymaganego do pracy wzmacniacza operacyjnego 803 812

Fig. 21. Measured minimal input voltage required for correct 803 812 op-amp operation

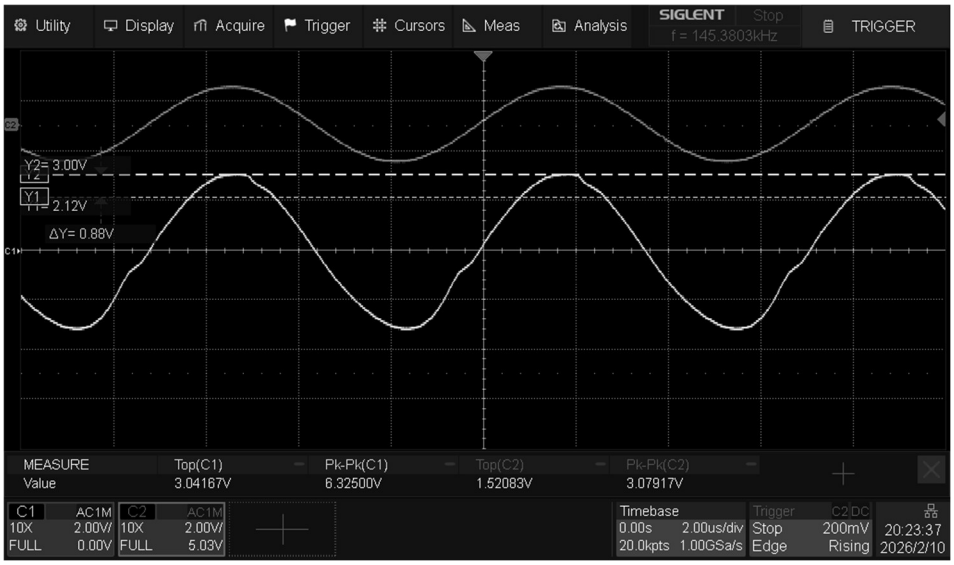
Aby zmierzyć to napięcie ustawiono kursor oscyloskopu w punkcie, gdzie napięcie wyjściowe przycina się. Następnie zmierzono jakie napięcie osiąga sygnał sterujący. Z testu wynika, że wzmacniacz działa poprawnie tylko wtedy, kiedy napięcie wejściowe wynosi przynajmniej 830mV (błąd pomiarowy wynosi ±23mV.

Kolejnym testem będzie sprawdzenie maksymalnej częstotliwości pracy. Test polega na zmierzeniu częstotliwości przy której napięcie wyjściowe wzmacniacza operacyjnego spada do poziomu 70.7% oryginalnego napięcia. Do testu użyto sygnału wejściowego o kształcie sinusoidy, amplitudzie 3Vpp i składowej stałej 3V, która ma zapobiec przycinaniu sygnału wyjściowego. Jaśniejszy przebieg (kanał 1) to sygnał wyjściowy, natomiast ciemniejszy przebieg (kanał 2) to sygnał wejściowy. Ustawione wzmocnienie równe 2 oznacza, że górna połówka sygnału wyjściowego powinna osiągać napięcie 3V (6Vpp). Granica maksymalnej częstotliwości pracy jest wyznaczona spadkiem napięcia wyjściowego do 70,7% teoretycznego napięcia, czyli 2,12V w tym przypadku. Do wykonania pomiaru ustawiono kursory na odpowiednich poziomach napięcia. Na rysunkach 22., 23. i 24. pokazano kolejne kroki przeprowadzania testu maksymalnej częstotliwości



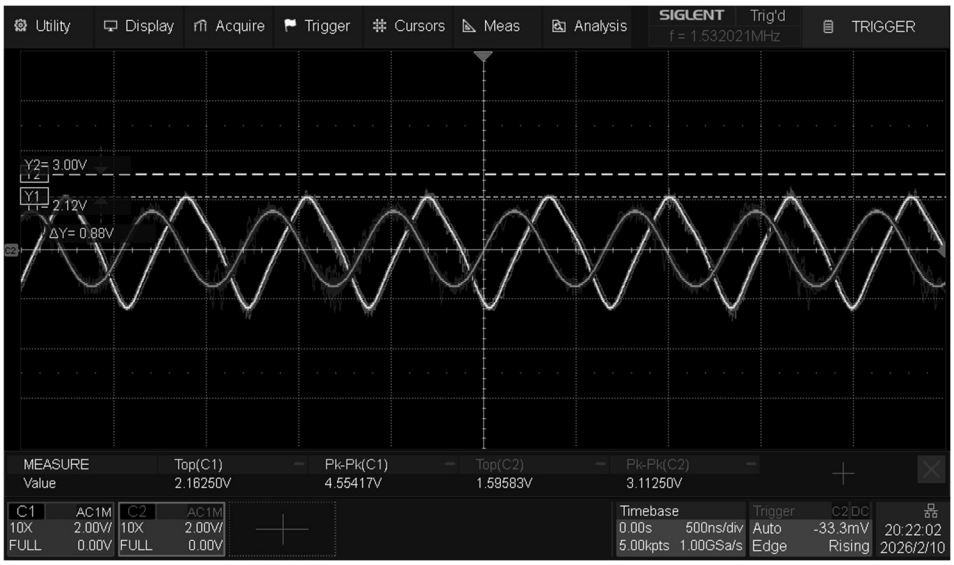
Rys. 22. Początek testu maksymalnej częstotliwości pracy wzmacniacza operacyjnego. Częstotliwość sygnału wynosi 100Hz, jego amplituda to 3Vpp. Do sygnału wejściowego dodano składową stałą wynoszącą 3V, aby uniknąć przycinania sygnału wyjściowego. Sprzężenie oscyloskopu ustawiono w tryb AC aby wyeliminować dodaną składową stałą z pomiarów w celu ich ułatwienia

Fig. 22. Start of the maximum op-amp bandwidth test. Signal frequency is 100Hz with 3Vpp amplitude. A 3V DC offset was added to the input signal to avoid clipping of the output signal. Oscilloscope probe is AC coupled to remove the added offset, which simplifies measurements



Rys. 23. Początek powstawania widocznych zniekształceń w sygnale wyjściowym przy częstotliwości 145,4kHz. Sygnał sterujący przesunięto do góry, aby ułatwić obserwację mierzonego sygnału

Fig. 23. Start of visible output signal distorions near the 145,4kHz frequency. Driving waveform was moved upwards to make the output waveform more visible

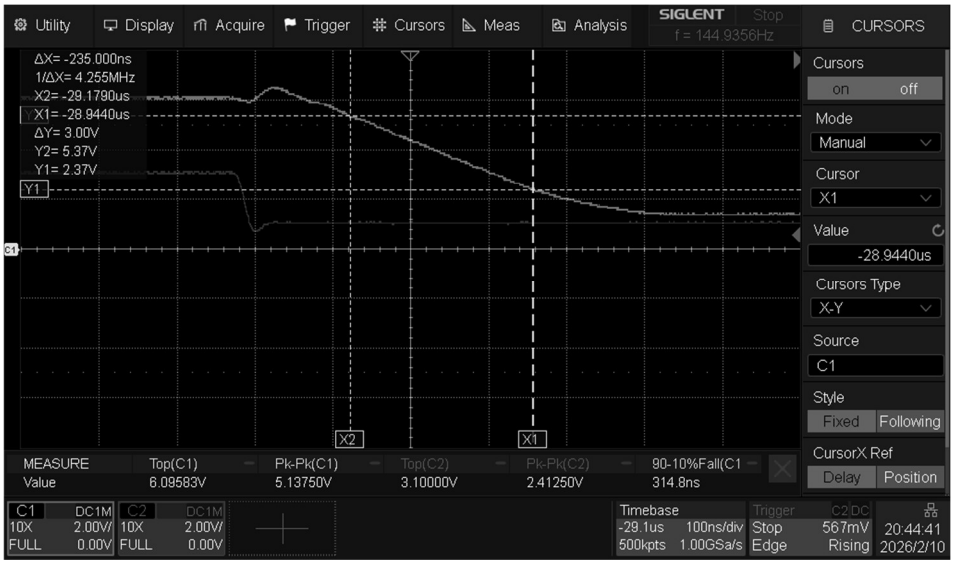


Rys. 24. Napięcie wyjściowe spadło do 70,7% oryginalnego napięcia, co wyznacza maksymalną częstotliwość graniczną wynoszącą 1,53MHz

Fig. 24. Output voltage fell to 70,7% of the original voltage, which marks the maximum bandwidth frequency of 1,53MHz

Testy wykazały, że przy napięciu sterującym wynoszącym 3 Vpp maksymalna częstotliwość pracy wynosi ok. 1,53 MHz. Dodatkowo przy częstotliwości 145,4 kHz pojawiają się zniekształcenia, lecz mogą być one wynikiem nieoptymalnego układu testującego (płytka stykowa, długie przewody połączeniowe).

Ostatnimi testami będzie zmierzenie szybkości narastania napięcia oraz czasu powrotu z przesterowania. Są to jedne z ważniejszych parametrów przy pracy z wysokimi częstotliwościami oraz wyższymi napięciami mogącymi sięgać napięcia zasilającego. Do tego celu użyto przebiegów sterujących z ostrymi zboczami, aby móc wyraźnie zaobserwować zmiany. Jako pierwszy zmierzono czas narastania za pomocą sygnału piło-kształtnego, co ukazano na rysunku 25.

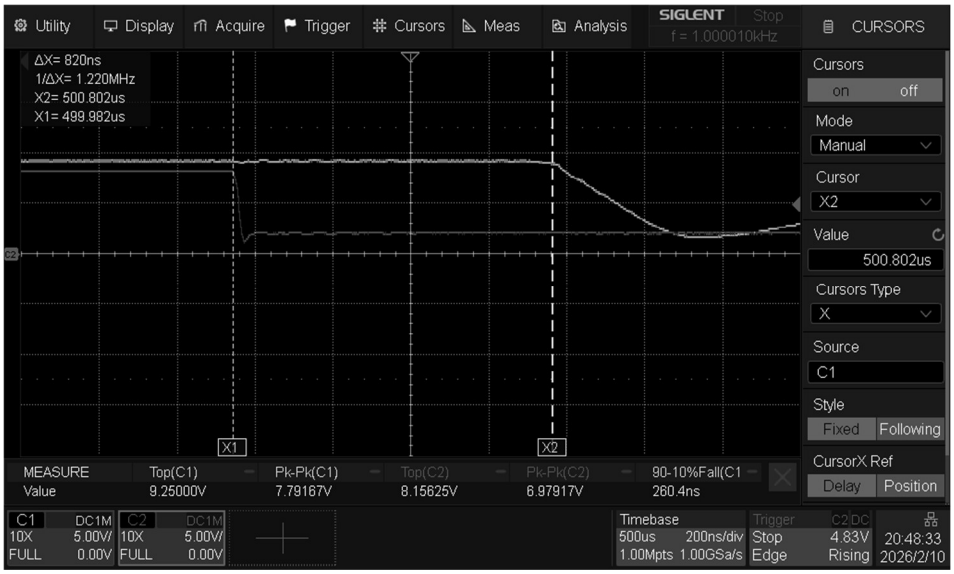


Rys. 25. Pomiar szybkości zmian napięcia na zboczu opadającym. W ciągu 235 ns napięcie spadło o 3 V, co daje wynik ok. 12,7 V/µs

Fig. 25. Slew rate measurement on the falling edge. During 235 ns voltage fell by 3 V, which gives a slew rate of around 12,7 V/µs

Pomiar szybkości zmiany napięcia wykazał, że w czasie 235 ns napięcie spadło o 3V, co po przeliczeniu daje wynik ok. 12,7 V/µs. Jest to dobry wynik w porównaniu do popularnych układów (np. LM358), które osiągają szybkość 0,5 V/µs.

Do pomiaru czasu powrotu z przesterowania użyto sygnału kwadratowego o dużej amplitudzie, co wprowadza wzmacniacz w stan przesterowania (napięcie wyjściowe wyższe od zasilającego, co nie jest możliwe). Pomiar pokazano na rysunku 26.



Rys. 26. Pomiar czasu powrotu z przesterowania. Wynosi on ok. 820 ns

Fig. 26. Oversteer recovery time measurement. Measured time equals around 820 ns

Za pomocą kursorów zmierzono czas powrotu z przesterowania, czyli ile czasu potrzebuje wzmacniacz aby znowu zaczął poprawnie podążać za sygnałem sterującym po wymuszeniu maksymalnego napięcia na wyjściu. Zmierzony czas wynosi ok. 820 ns, co jest dobrym wynikiem w porównaniu do popularnych układów LM358, które nominalnie potrzebują 10 µs na powrót z przesterowania. Dodatkowo na przebiegu można zaobserwować napięcie strat wyjściowych, czyli maksymalne napięcie wyjściowe względem napięcia zasilającego, które wynosi 750 mV. Ten parametr można obliczyć odejmując maksymalne napięcie wyjściowe układu (9,25 V) od napięcia zasilającego (10 V). Oznacza to, że maksymalne napięcie wyjściowe układu jest zawsze o 750 mV niższe od napięcia zasilającego.

W tabeli 8. zestawiono wszystkie zmierzone oraz zauważone parametry identyfikowanego układu (należy pamiętać, że nie są to dokładne wartości; warunki pomiarów nie były standardowe).

Z parametrów można wywnioskować, że jest to dosyć szybki wzmacniacz operacyjny, który dobrze reaguje na zmiany napięcia i może w krótkim czasie powrócić ze stanu przesterowania. Są to cechy pożądane w sterowniku lasera przemysłowego, gdzie potrzebny jest krótki czas reakcji na zmiany parametrów pracy lasera.

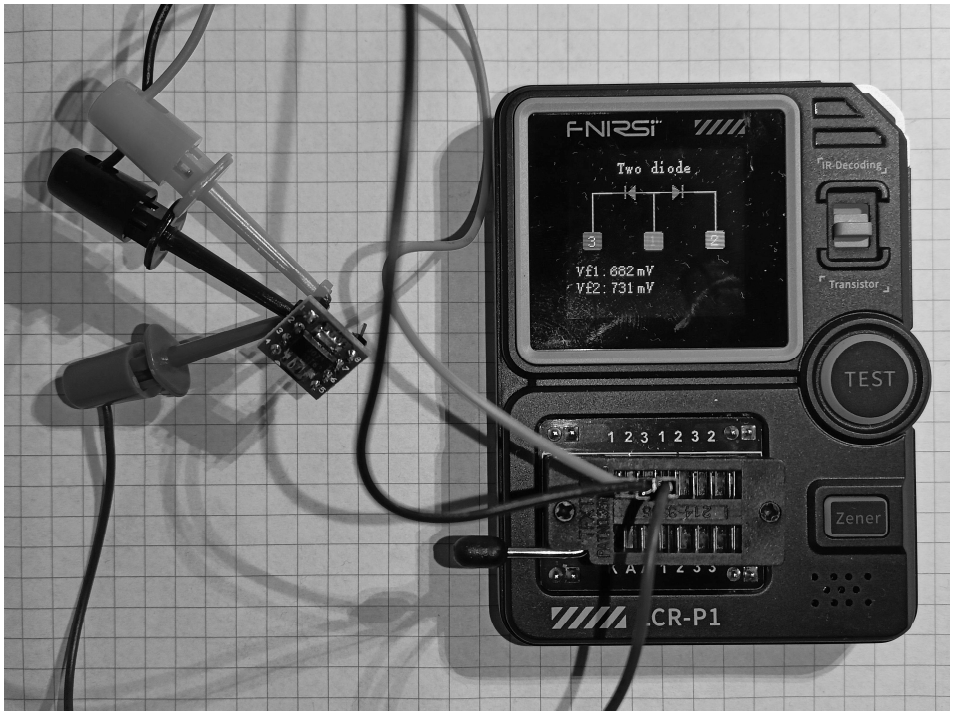
Tabela 8. Zestawienie wyników pomiarów identyfikowanego układu *803 812*

Table 8. *803 812* IC measurement results

Nazwa parametru	Wartość	Jednostka
Napięcie zasilania	30	V
Napięcie strat wyjściowych	750	mV
Minimalne napięcie wejściowe	830	mV
Pasmo przenoszenia (-3dB)	1,5	MHz
Czas powrotu z przesterowania	820	ns
Szybkość narastania napięcia	12,7	V/µs

5.6. Układ J073 903T63

Układ ten posiada obudowę SOT-223, która jest typowa dla tranzystorów typu MOSFET oraz regulatorów napięcia. Do wstępnej identyfikacji zostanie wykorzystany tester komponentów LCR-P1. Wynik testu pokazano na rysunku 27.



Rys. 27. Wynik pomiaru układu J073 za pomocą testera LCR-P1. Układ zidentyfikowano jako dwie diody

Fig. 27. Measurement results of the J073 device using the LCR-P1 component tester. Device was identified as two diodes

Tester wskazuje układ jako dwie diody. Jest to jednak mało prawdopodobne ze względu na relatywnie dużą różnicę napięć przewodzenia, niestandardowy układ wyprowadzeń (punkt wspólny diod powinien być na środkowej nóżce) oraz niepraktyczny kierunek przewodzenia diod (wspólna anoda, czyli plus, nie pozwala na sensowne wykorzystanie diody, bo jedna dioda zacznie przewodzić przed drugą, co uniemożliwi jej wykorzystanie). Mając to na uwadze oraz pamiętając o fakcie, że tester nie potrafi poprawnie rozpoznać regulatorów napięcia bezpiecznym będzie założenie, że właśnie tym jest badany komponent.

Regulator napięcia posiada 3 wyprowadzenia:

- Vin - doprowadzenie zasilania,
- GND - połączenie masowe, czyli minus,
- Vout - wyjście napięcia regulowanego przez układ.

Znając funkcje nóżek można zauważyć, że napięcie doprowadzane jest tylko do jednej z nich, a pozostałe 2 są wyjściami (Vout odprowadza regulowane napięcie, GND przejmuje nadwyżkę napięcia). Zgadza się to ze wskazaniem testera, gdzie jedno wyprowadzenie przewodzi prąd do pozostałych dwóch.

Aby sprawdzić czy układ faktycznie jest regulatorem zostało podłączone do niego zasilanie. Dodatni biegun do nóżki numer 1, a ujemny w pierwszej fazie testu do nóżki numer 2. Napięcie zasilające ustawiono na 5 V oraz zastosowano ograniczenie prądowe do 20 mA na wypadek niepoprawnego podłączenia, co nie pozwoli na zniszczenie układu spowodowane ewentualnym przegrzaniem wywołanym dużym prądem. Po włączeniu zasilania prąd natychmiast osiągnął limit, co jednoznacznie mówi, że jest to nieprawidłowe podłączenie. Wykonano więcej iteracji testów w danej kolejności:

1. Zasilanie do pinu 1, masa do pinu 3. Prąd osiągnął maksymalną wartość, co oznacza nieprawidłowe podłączenie.
2. Zasilanie do pinu 2, masa do pinu 3. Prąd ponownie osiągnął maksymalną wartość.
3. Zasilanie do pinu 2, masa do pinu 1. Tym razem zasilacz nie wskazuje żadnego poboru prądu. Przy pomiarze napięcia między masą a wyprowadzeniem numer 3 multimetr wskazał napięcie 4,5 V. Napięcie podąża za napięciem zasilania, ale jest o 0,5 V niższe. Ustawiając zasilanie na 6 V napięcie na nóżce wzrosło do 5,5 V. Jest to nieprawidłowe zachowanie regulatora
4. Zasilanie do pinu 3, masa do pinu 1. Przy tej próbie napięcie między masą i ostatnią nóżką (numer 2) wyniosło 3,3 V, co jest standardowym napięciem często spotykanym przy regulatorach napięcia. Zwiększanie i zmniejszanie napięcia o 1 V nie miało wpływu na napięcie wyjściowe mierzone na wyprowadzeniu.

Ostatni test ustalił dokładną konfigurację sprawdzanego układu, czyli:

- nóżka 1 - GND,
- nóżka 2 - Vout,
- nóżka 3 - Vin.

Znając konfigurację można ustalić podstawowe parametry regulatora, czyli czy jest to regulator zwykły lub LDO, jaki jest jego minimalny spadek napięcia. oraz jakie jest jego maksymalne napięcie regulowania. Aby to sprawdzić należy zasilić sprawdzany układ i mierząc napięcie wyjściowe zmieniać napięcie zasilające. Zwiększając napięcie z 5 V co 0,1 V udało się ustalić, że maksymalne napięcie wejściowe to 7,9 V. Regulator pracował poprawnie do napięcia 7,5 V. Przy każdym kolejnym kroku napięcie wyjściowe zwiększało się. Przy zasilaniu 7,8 V napięcie wyjściowe wyniosło 3,7 V, co jest wartością niebezpiecznie wysoką dla niewielkich układów z nominalnym zasilaniem 3,3 V. Przy 7,9 V regulator wyłączył się, co jest najprawdopodobniej wbudowanym zabezpieczeniem przeciwko zbyt wysokiemu napięciu wyjściowemu, co ma na celu zabezpieczenie zasilanych komponentów.

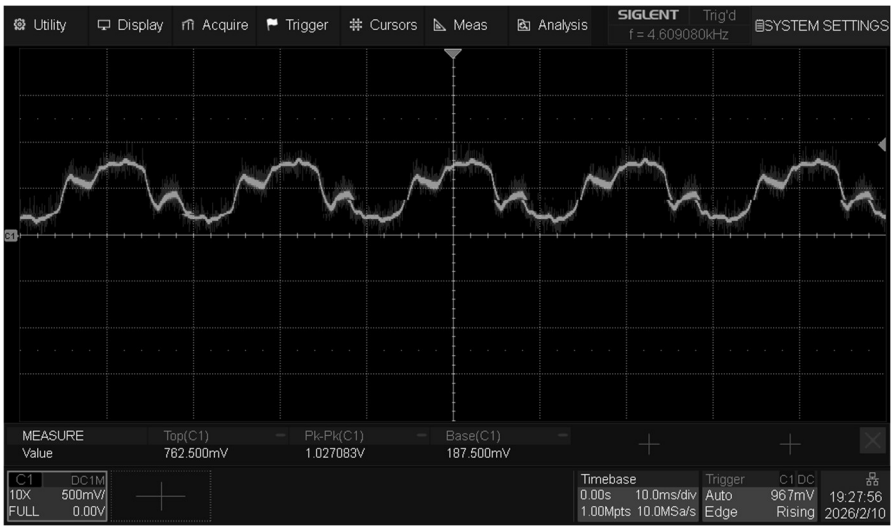
Zmniejszając napięcie regulator utrzymywał stabilne napięcie wyjściowe, nawet kiedy napięcie wejściowe było równe nominalnemu wyjściowemu 3,3 V. Przy zasilaniu napięciem 3 V regulator podawał 3 V na wyjście. Oznacza to, że jest to regulator LDO (Low DropOut), czyli o niskim spadku napięcia. Standardowe regulatory posiadają minimalny spadek napięcia, co oznacza, że napięcie zasilające musi być wyższe od wyjściowego co najmniej o tą wartość. Regulatory LDO nie mają tej wady, bo ich minimalny spadek napięcia jest bardzo niski (w tym przypadku trudno jest w ogóle zmierzyć ten spadek). Ich wadą jest jednak niewielkie maksymalne napięcie zasilania. Standardowe regulatory mogą obniżać napięcie zasilające nawet o 30V, co jest imponującą wartością w porównaniu do maksymalnego napięcia mierzonego układu *J703*, które wynosi 7,9 V.

5.7. Układ ICL 358

Jest to ostatni układ jaki zostanie zidentyfikowany. W tym przypadku jest to znacznie ułatwione przez liczbę 358 w nazwie. Jednym z najpopularniejszych uniwersalnych wzmacniaczy operacyjnych jest LM358, który można spotkać w wielu konstrukcjach. Oznaczenie tego układu to bardzo prawdopodobnie skrócona nazwa producenta oraz skrót oznaczenia LM358. Aby zweryfikować czy część jest faktycznie wzmacniaczem operacyjnym, został wykonany taki sam test jak przy układzie *803 812*, czyli pomiar nóżek zasilonego układu. Zasilanie zostało podłączone w taki sam sposób jak przy teście układu *803 812*, czyli nóżka 4 połączona do masy oraz nóżka 8 do zasilania +5 V. Po podłączeniu układ nie wykazał nadmiernego poboru prądu (<1 mA). Zmierzono wszystkie nóżki poza zasilającymi, co pokazano na rysunkach 28. oraz 29.

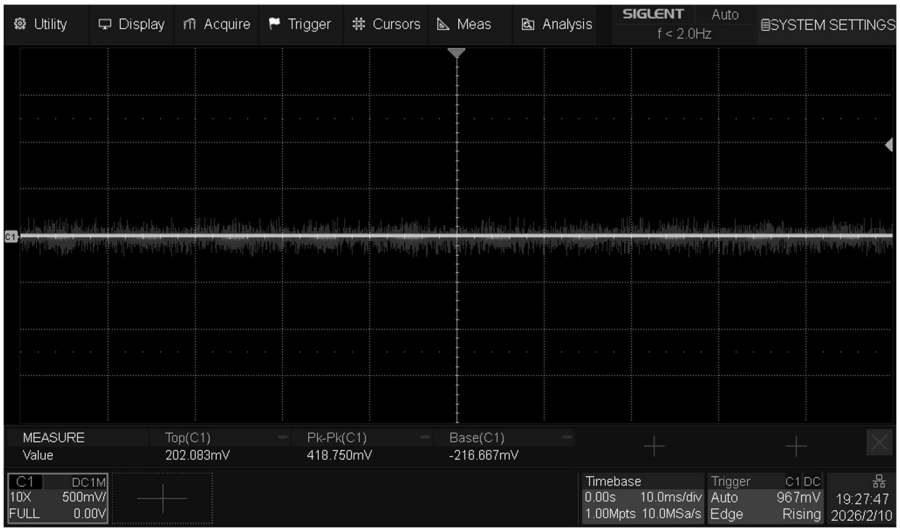
Przebieg widoczny na rysunku 28. można było zaobserwować na nóżkach 2, 3, 5 oraz 6, natomiast przebieg 29 był widoczny na nóżkach 1 oraz 7. Zakłócenia widoczne na rysunku 28. mogą być powodowane wysoką impedancją nóżek, co w połączeniu z faktem, że nóżki te w standardowym układzie wyprowadzeń LM358 są wejściami bardzo mocno sugeruje, że w tym przypadku również są one wejściami. Na pozostałych nóżkach pojawia się napięcie stałe o wartości około

200 mV, co jest zgodne z zachowaniem wyjść układu z niepodłączonymi (pływającymi) wejściami.



Rys. 28. Przebieg napięcia zmierzony na nóżkach 2, 3, 5 i 6. Przebieg przypomina sinusoidę, a jego okres to ok. 20ms, co daje częstotliwość 50Hz. Przechwytywanie napięcia sieciowego świadczy o wysokiej impedancji

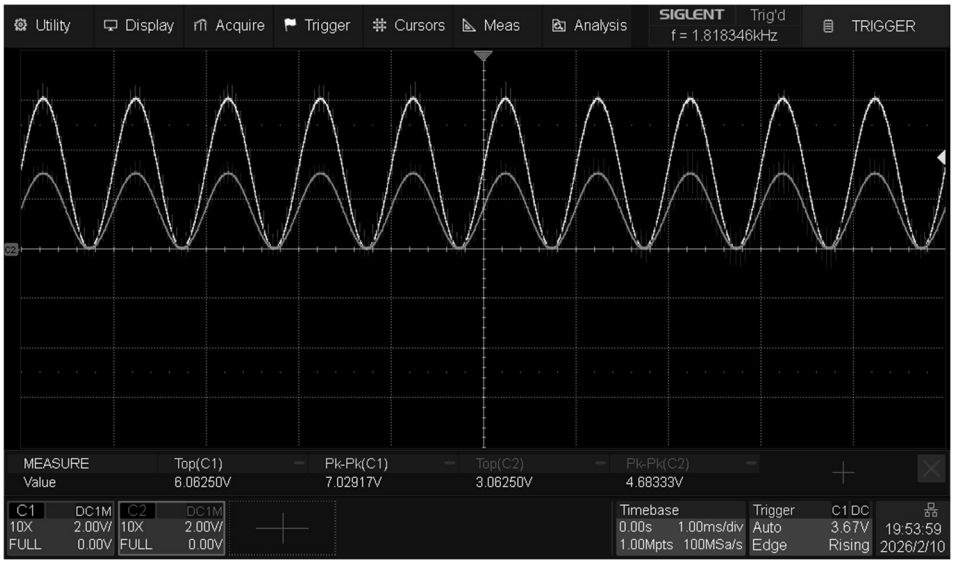
Fig. 28. voltage waveform measured on pins 2, 3, 5 and 6. The waveform resembles a sinusoide, it's period measures around 20ms, which gives a frequency of 50Hz. Interference from the power grid suggests a high impedance



Rys. 29. Pomiar przebiegu napięcia na nóżkach 1 i 7

Fig. 29. Measured voltage waveform on pins 1 and 7

Kolejnym testem na weryfikację układu jest skonfigurowanie go jako wzmacniacz nieodwracający ze wzmocnieniem równym 2 w taki sam sposób, jak poprzedni układ *803 812*, czyli zgodnie ze schematem z rysunku 21. Na rysunku 30. ukazano wynik testu.



Rys. 30. Przebieg napięcia sygnału wejściowego (ciemniejszy) oraz wyjściowego (jaśniejszy) po wzmocnieniu przez wzmacniacz ICL 358

Fig. 30. Voltage waveform of input signal (darker) and output signal (lighter) after amplification by ICL 358 op-amp

Na rysunku 30. możemy zaobserwować sygnał sterujący z generatora (kanał 2, ciemniejszy) oraz zmierzony sygnał wyjściowy (kanał 1, jaśniejszy). Sygnał wejściowy został wzmocniony dwukrotnie zgodnie z założeniami, co potwierdza, że badany układ jest wzmacniaczem operacyjnym. Biorąc pod uwagę, że w oznaczeniu znajduje się numer 358 pewnym jest, że ten układ to LM358.

6. Wnioski

Posiadając odpowiedni sprzęt oraz wiedzę możliwa jest identyfikacja układów bez wyraźnego oznaczenia. Ta umiejętność jest obecnie szczególnie ważna ze względu na częste próby utajnienia modeli i funkcji komponentów przez producentów urządzeń elektronicznych. Praktyka ta ma na celu utrudnienie lub uniemożliwienie naprawy urządzenia, zmuszając użytkownika do zakupu nowego egzemplarza, zamiast naprawy uszkodzonego. Generuje to dodatkowe koszty dla konsumentów oraz znaczne ilości elektrośmieci, które zawierają wiele szkodliwych dla środowiska substancji. Ustalenie funkcji układu elektronicznego

umożliwia jego wymianę na działający, co ratuje całe urządzenie przed wyrzuceniem. Bardzo często zdarza się również, że komponent ulega katastrofalnej awarii, co uniemożliwia odczytanie jego oznaczeń w przypadku eksplozji lub wypalenia. Ustalenie jego funkcji na podstawie otoczenia niejednokrotnie pozwoliło na naprawę urządzenia, co bez tej umiejętności byłoby niemożliwe.

Wykorzystując wskazówki i metodykę z przedstawionego rozdziału możliwe jest rozpoznanie podstawowych układów elektronicznych oraz ich funkcji, co może pozwolić na naprawę uszkodzonego urządzenia lub wykorzystanie wylutowanego układu w innym urządzeniu lub w projekcie. Wtórne użycie zdemontowanych komponentów jest dobrym sposobem na unikanie konieczności zamawiania nowych układów, promowanie ekologii oraz naukę elektroniki, co może otworzyć nowe ścieżki kariery zawodowej, między innymi jako projektant płytek PCB lub serwisant sprzętu elektronicznego. Oba zawody są pożądane przez branżę oraz zapewniają wysokie zarobki.

Przy konieczności identyfikacji komponentu warto jest wykonać opisane w rozdziale czynności:

1. Wyszukać oznaczenie komponentu w internetowych bazach danych oraz zweryfikować, czy takie oznaczenie faktycznie może występować na komponencie (skrócone oznaczenia są zazwyczaj podane w specyfikacji).
2. Sprawdzić sąsiadujące komponenty. Sąsiedztwo cewki sugeruje, że układ ma funkcję zasilającą, wiele rezystorów lub jeden duży rezystor o niskiej wartości sugeruje funkcję pomiarową, a sąsiedztwo dużego układu scalonego oznacza, że komponent może być pamięcią.
3. Należy zmierzyć napięcia (i ich przebiegi jeśli możliwe). Stałe niezmienne napięcie o standardowej wartości (np. 1,8 V, 3,3 V, 5 V) oznaczają zasilanie układu, natomiast ostre szybkie przebiegi przypominające sygnał kwadratowy oznaczają linię danych.

Kroki te pozwolą na ustalenie funkcji układu, lecz znalezienie jego dokładnego modelu (lub pasującego zamiennika) jest trudniejsze i prawdopodobieństwo diametralnie różni się w zależności od funkcji. Układy pamięci są powszechne i różnią się głównie pojemnością oraz protokołem (SPI lub I^2C); znając te 2 parametry można wymienić pamięć, z dużą pewnością, że będzie poprawnie działać. Układy zasilające (w szczególności przetwornice DC/DC) stanowią większe wyzwanie ze względu na różne wyprowadzenia nóżek, częstotliwości przełączania oraz napięcie odniesienia. Wszystkie te parametry różnią się między układami i znalezienie odpowiedniego zamiennika jest bardzo trudne.

Literatura i źródła internetowe

[1] https://www.bunniestudios.com/blog/2014/name-that-ware-september-2014/ (archiwum: https://web.archive.org/web/20250815150408/https://www.bunniestudios.com/blog/2014/name-that-ware-september-2014/) (dostęp: 26.06.2026).

[2] Ćwik P.: Planowane postarzanie produktu generuje góry elektrośmieci. UE chce z tym walczyć. Smog Lab: https://smoglab.pl/planowane-postarzanie-produktu-generuje-gory-elektrosmieci-ue-chce-z-tym-walczyc/ (dostęp: 25.12.2025).

[3] https://www.fnirsi.com/products/lcr-p1 (dostęp: 27.12.2025).

[4] https://ndn.com.pl/zasilacze-laboratoryjne/4867-korad-kkg305p-zasilacz-programowalny-dc-30v5a.html (dostęp: 27.12.2025).

[5] https://www.siglenteu.com/digital-oscilloscopes/sds800x-hd-digital-storage-oscilloscope/ (dostęp: 27.12.2025).

[6] https://instruments.uni-trend.com/products/waveform-generators/UTG900E (dostęp: 27.12.2025).

[7] Bergman E.: Simple but effective Transistor Curve Tracer circuit. https://www.eddybergman.com/2017/06/simple-but-effective-transistor-curve.html (dostęp: 23.02.2026).

IDENTIFYING INTEGRATED CIRCUITS WITHOUT FULL MARKINGS

Summary

This article describes the process of identifying integrated circuits found in random electronic devices that, for various reasons, lack complete markings necessary for quick identification. The functions of these circuits were determined based on measurements, adjacent components, package type, and partial markings. After identification, their parameters were measured, including maximum operating frequency, current consumption, input capacitance, and output voltage rise time. The article outlines the testing methods used and how the function of each circuit was identified. The presented methods and tips can be applied to recognize common components used in electronic devices. Identifying unmarked integrated circuits is essential for repairing devices in case of circuit failure, especially in catastrophic failures where the circuit may be burned beyond recognition. In such cases, only clues from neighboring components and PCB trace layouts can be used. Understanding the circuit's function also allows for its reuse in other devices or the selection of a compatible replacement for the repaired item. The ability to identify electronic components is crucial for repairing electronic equipment, which would otherwise become e-waste, exacerbating the global issue of electronic waste and the climate crisis. As part of this article, seven components removed from discarded electronic devices were examined. All were initially identified, followed by more detailed testing to confirm their functions and measure their parameters.

Keywords: electronics, measurements, electronic repair, ecology, planned obsolescence, reverse engineering.

Gabriela Bober, Dorota Głowacz-Czerwonka
Koło Naukowe Studentów Chemii „Esprit”

WPŁYW GRAFITU EKSPANDOWANEGO NA WŁAŚCIWOŚCI FIZYCZNE I OGNIOWE SZTYWNYCH PIANEK POLIURETANOWYCH ZAWIERAJĄCYCH KWIATY GOŹDZIKA

Streszczenie

Celem pracy było określenie wpływu grafitu ekspandowanego na właściwości fizyczne oraz ogniowe sztywnych pianek poliuretanowych zawierających bio-wypełniacz w postaci suszonych kwiatów goździka. W ramach przeprowadzonych badań otrzymano serię pianek poliuretanowych, obejmującą piankę referencyjną, piankę zawierającą grafit ekspandowany oraz pianki zawierające grafit ekspandowany i różną zawartość suszonych kwiatów goździka. Zastosowanie grafitu ekspandowanego miało na celu poprawę odporności ogniowej materiałów, natomiast wprowadzenie bio-wypełniacza stanowiło próbę modyfikacji struktury oraz właściwości fizycznych pianek poliuretanowych.

Dla otrzymanych materiałów przeprowadzono szereg badań pozwalających na ocenę ich właściwości użytkowych. Oznaczono gęstość pozorną, chłonność wody, stabilność wymiarów oraz współczynnik przewodzenia ciepła. Ponadto przeanalizowano właściwości związane z palnością materiałów, w tym indeks tlenowy oraz przebieg spalania w teście poziomym.

Uzyskane wyniki wskazują, że zastosowanie grafitu ekspandowanego prowadzi do zwiększenia gęstości pianek poliuretanowych oraz wyraźnej poprawy ich odporności ogniowej. Wprowadzenie grafitu wpływa również na przebieg procesu spalania poprzez tworzenie na powierzchni materiału warstwy ochronnej o charakterze zwęglonym, która ogranicza dostęp tlenu oraz zmniejsza szybkość rozprzestrzeniania się płomienia. Obecność kwiatów goździka w strukturze pianek wpływała natomiast na wybrane właściwości fizyczne materiałów, związane przede wszystkim ze zmianą ich struktury komórkowej. Otrzymane wyniki potwierdzają, że grafit ekspandowany może stanowić skuteczny dodatek ograniczający palność pianek poliuretanowych modyfikowanych bio-wypełniaczami roślinnymi.

Słowa kluczowe: kompozyty polimerowe, środki zmniejszające palność, bio-wypełniacze, właściwości ogniowe, materiały izolacyjne.

1. Wprowadzenie

Pianki poliuretanowe stanowią grupę materiałów polimerowych o strukturze komórkowej, powstających w wyniku reakcji polioli z izocyjanianami w obecności katalizatorów, środków powierzchniowo czynnych oraz czynników spieniających. W trakcie reakcji zachodzi jednocześnie proces tworzenia sieci poliuretanowej oraz wydzielania gazu, najczęściej dwutlenku węgla, który odpowiada za powstanie charakterystycznej struktury porowatej materiału [6].

Właściwości pianek poliuretanowych zależą od wielu czynników, takich jak rodzaj zastosowanego poliolu i izocyjanianu, stosunek molowy grup funkcyjnych NCO/OH oraz obecność dodatków modyfikujących strukturę materiału. Istotne znaczenie mają również parametry mikrostruktury piany, w tym wielkość komórek oraz udział fazy zamkniętej i otwartej, które wpływają na właściwości mechaniczne i termoizolacyjne materiału [5]. Jednym z istotnych ograniczeń stosowania pianek poliuretanowych jest jednak ich podatność na działanie wysokiej temperatury oraz stosunkowo wysoka palność. W warunkach oddziaływania płomienia materiały te ulegają degradacji termicznej, czemu może towarzyszyć wydzielanie produktów spalania oraz szybkie rozprzestrzenianie się ognia [7-8]. Z tego względu w ostatnich latach prowadzi się intensywne badania nad metodami ograniczania palności materiałów poliuretanowych.

Jednym z rozwiązań stosowanych w celu poprawy odporności ogniowej pianek poliuretanowych jest wprowadzanie dodatków uniepalniających. Wśród nich szczególne znaczenie ma grafit ekspandowany, który pod wpływem wysokiej temperatury zwiększa swoją objętość, tworząc warstwę ochronną na powierzchni materiału. Powstająca warstwa zwęglona stanowi barierę ograniczającą dostęp tlenu oraz zmniejszającą szybkość rozprzestrzeniania się płomienia [1, 4].

Mechanizm działania grafitu ekspandowanego polega na tworzeniu izolacyjnej warstwy zwęglonej na powierzchni materiału, która ogranicza wymianę ciepła pomiędzy płomieniem a materiałem polimerowym oraz utrudnia dyfuzję gazów palnych [2, 9].

Równocześnie coraz większe znaczenie w badaniach nad materiałami polimerowymi mają dodatki pochodzenia naturalnego. Bio-wypełniacze roślinne mogą wpływać na strukturę komórkową pianek poliuretanowych, a także modyfikować ich właściwości fizyczne oraz mechaniczne [3]. Zastosowanie dodatków roślinnych wpisuje się również w aktualne kierunki rozwoju materiałów kompozytowych o obniżonym wpływie na środowisko. Połączenie dodatków mineralnych o właściwościach uniepalniających z bio-wypełniaczami roślinnymi może prowadzić do otrzymania materiałów kompozytowych o zmodyfikowanej strukturze oraz poprawionych właściwościach użytkowych.

Celem niniejszej pracy była ocena wpływu grafitu ekspandowanego na właściwości fizyczne oraz ogniowe sztywnych pianek poliuretanowych zawierających suszone kwiaty goździka oraz określenie wpływu jednoczesnego zastosowania dodatku mineralnego i roślinnego na charakterystykę badanych materiałów.

2. Część doświadczalna

Pianki poliuretanowe otrzymano poprzez zmieszanie składnika poliolowego z izocyjanianem w obecności odpowiednich dodatków. Grafit ekspandowany (EG290, Sinograf S.A., Toruń, Polska) oraz suszone kwiaty goździka (importer: Ziołowy Dwór, Polska; kraj pochodzenia: Chiny) dodano do składnika poliolowego przed rozpoczęciem reakcji spieniania. Suszone kwiaty goździka przed wprowadzeniem do układu rozdrobniono w moździerzu, a następnie przy użyciu mieszadła mechanicznego. Nie określano wielkości cząstek po rozdrobnieniu, ponieważ ze względu na obecność twardych elementów rośliny nie uzyskano całkowicie jednorodnej frakcji. Bio-wypełniacz stosowano bez dodatkowej obróbki termicznej. Po wymieszaniu składnika poliolowego z dodatkami w temperaturze pokojowej dodawano izocyjanian, inicjując egzotermiczną reakcję prowadzącą do procesu spieniania Po zakończeniu procesu próbki pozostawiono do sezonowania przez 72 godziny w temperaturze pokojowej.

Przygotowano następujące kompozycję:

- Pianka referencyjna RF 151V
- RF 151V + EG290 (60% mas.)
- RF 151V + EG290 (60% mas.) + 5% mas. kwiatów goździka
- RF 151V + EG290 (60% mas.) + 10% mas. kwiatów goździka
- RF 151V + EG290 (60% mas.) + 15% mas. kwiatów goździka

Poniżej przedstawiono jak zmienia się struktura sztywnych pianek poliuretanowych z dodatkiem grafitu ekspandowanego wraz ze wzrostem udziału kwiatów goździka (rys. 1.).



Rys. 1. Zmiana struktury sztywnych pianek poliuretanowych z dodatkiem grafitu ekspandowanego wraz ze wzrostem udziału suszonych kwiatów goździka

Fig. 1. Changes in the Structure of Rigid Polyurethane Foams Containing Expanded Graphite with Increasing Content of Dried Clove Flowers

2.1. Badanie właściwości fizycznych

Właściwości fizyczne otrzymanych pianek poliuretanowych oceniono na podstawie pomiarów gęstości pozornej, chłonności wody, stabilności wymiarów oraz współczynnika przewodzenia ciepła.

Gęstość pozorną określono zgodnie z normą *PN-EN ISO 845:2010* na podstawie stosunku masy do objętości przygotowanych próbek o regularnym kształcie. Wymiary próbek wyznaczono przy użyciu suwmiarki, natomiast masę określono z wykorzystaniem wagi analitycznej.

Chłonność wody oznaczono zgodnie z normą *PN-93/C-89084* poprzez zanurzenie próbek w wodzie destylowanej i określenie zmiany ich objętości po określonym czasie ekspozycji.

Stabilność wymiarów pianek badano według normy *PN-EN 2796:1986*. Próbki poddawano działaniu podwyższonej temperatury w suszarce laboratoryjnej, a następnie określano zmiany ich wymiarów po określonym czasie ekspozycji.

Współczynnik przewodzenia ciepła wyznaczono zgodnie z normą *PN-EN 993-15:2006* przy użyciu sondy pomiarowej wprowadzanej bezpośrednio do badanych materiałów.

2.2. Badanie właściwości ogniowych

Właściwości palne pianek poliuretanowych oceniono na podstawie oznaczenia indeksu tlenowego oraz analizy przebiegu spalania w teście poziomym.

Indeks tlenowy wyznaczono zgodnie z normą *PN-EN ISO 4589-2:2017-06*. Badanie polegało na określeniu minimalnego stężenia tlenu w mieszaninie tlenu i azotu, przy którym możliwe jest podtrzymanie procesu spalania próbki materiału.

Dodatkowo przeprowadzono test palności w układzie poziomym zgodnie z normą *PN-EN ISO 3582:2002*. Podczas badania rejestrowano czas spalania oraz drogę rozprzestrzeniania się płomienia, co umożliwiło wyznaczenie szybkości spalania oraz ubytku masy badanych próbek.

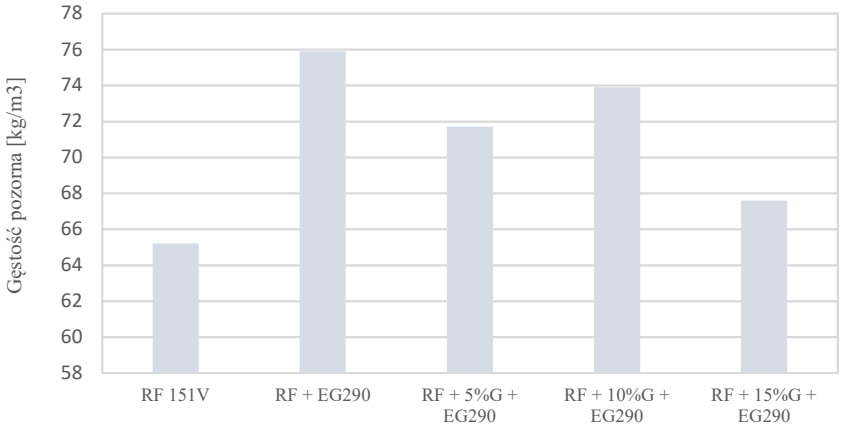
3. Analiza i omówienie wyników

3.1. Gęstość pozorna

Na rysunku 2. przedstawiono wartości gęstości pozornej otrzymanych pianek poliuretanowych. Najniższą wartość tego parametru odnotowano dla pianki referencyjnej RF 151V, natomiast wprowadzenie grafitu ekspandowanego spowodowało zwiększenie gęstości materiału. Zjawisko to można tłumaczyć obecnością cząstek mineralnego dodatku w strukturze komórkowej pianek.

W przypadku pianek zawierających zarówno grafit ekspandowany, jak i kwiaty goździka obserwowano dalsze zmiany wartości gęstości. Dla pianek zawierających kwiaty goździka zauważyć można nieliniowy spadek wartości gęstości pozornej w odniesieniu do pianki zawierającej jedynie grafit ekspandowany,

co może wynikać z porowatej struktury bio-wypełniacza oraz jego wpływu na przebieg procesu spieniania.

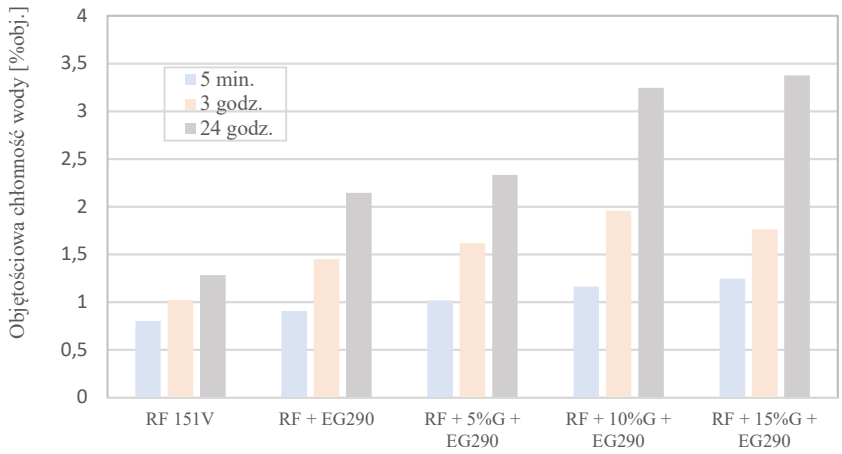


Rys. 2. Gęstość pozorna badanych pianek poliuretanowych

Fig. 2. Apparent density of polyurethane foams

3.2. Chłonność wody

Na rysunku 3. przedstawiono wyniki oznaczenia chłonności wody badanych pianek poliuretanowych. Pianka referencyjna charakteryzowała się najniższą chłonnością wody, co wynika z jednorodnej struktury komórkowej materiału.



Rys. 3. Objętościowa chłonność wody badanych pianek poliuretanowych

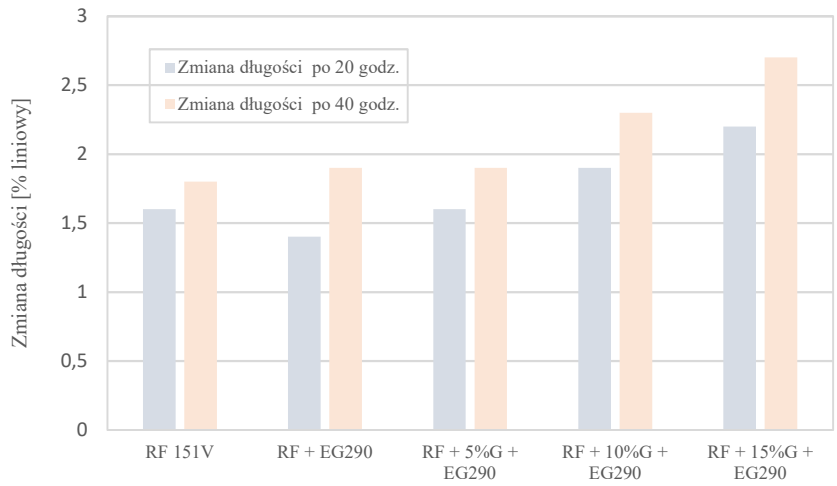
Fig. 3. Volumetric water absorption of the tested polyurethane foams

Dodatek grafitu ekspandowanego powodował niewielki wzrost wartości tego parametru. Może to być związane z obecnością cząstek grafitu w strukturze materiału, które wpływają na zmianę budowy komórkowej pianek.

W przypadku pianek zawierających grafit oraz kwiaty goździka obserwowano dalsze zwiększenie chłonności wody wraz ze wzrostem zawartości bio-wypełniacza. Zjawisko to można tłumaczyć obecnością składników roślinnych o charakterze hydrofilowym, które sprzyjają absorpcji wody oraz zwiększają liczbę otwartych komórek w strukturze materiału.

3.3. Stabilność wymiarów

Wyniki badania stabilności wymiarów przedstawiono na rys. 4. Analizie poddano zmianę długości badanych materiałów. Pianka referencyjna wykazywała niewielkie zmiany długości po ekspozycji w podwyższonej temperaturze, co świadczy o dobrej stabilności struktury materiału.



Rys. 4. Zmiana długości badanych pianek poliuretanowych

Fig. 4. Change in the length of polyurethane foams

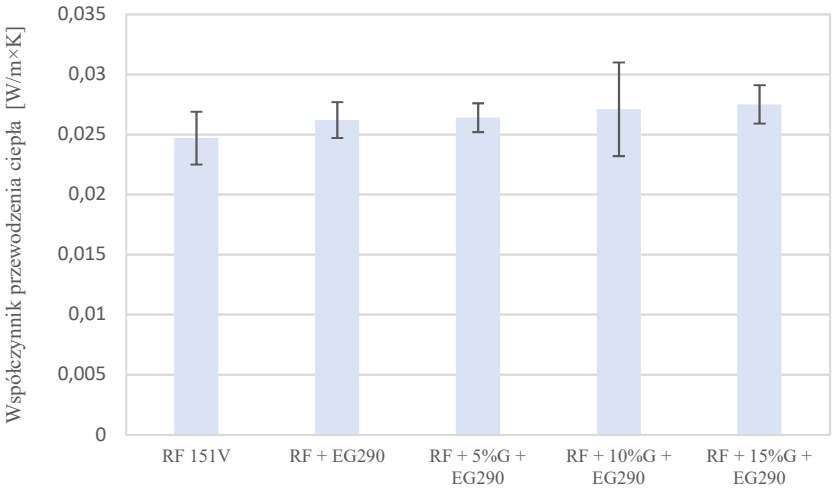
Wprowadzenie grafitu ekspandowanego nie powodowało znaczącego pogorszenia stabilności wymiarów pianek. Obserwowano nawet niewielką poprawę stabilności, co może być związane z obecnością cząstek mineralnego dodatku, które wzmacniają strukturę komórkową materiału.

W przypadku pianek zawierających jednocześnie grafit oraz kwiaty goździka obserwowano większe zmiany długości, szczególnie przy wyższej zawartości dodatku roślinnego. Może to wynikać z mniejszej jednorodności struktury komórkowej oraz obecności cząstek bio-wypełniacza w matrycy poliuretanowej.

3.4. Współczynnik przewodzenia ciepła

Na rysunku 5. przedstawiono wartości współczynnika przewodzenia ciepła dla analizowanych pianek poliuretanowych. Pianka referencyjna charakteryzowała się niską wartością współczynnika przewodzenia ciepła, co potwierdza jej dobre właściwości termoizolacyjne.

Dodatek grafitu ekspandowanego spowodował niewielki wzrost współczynnika przewodzenia ciepła, a przy połączeniu grafitu z kwiatami widać jego stopniowy wzrost, wraz ze wzrostem udziału kwiatów. Wyniki można traktować jako bardzo zbliżone do wartości referencyjnych, zauważalny jest jednak niewielki wpływ dodatków modyfikujących na właściwości izolacyjne badanych pianek poliuretanowych

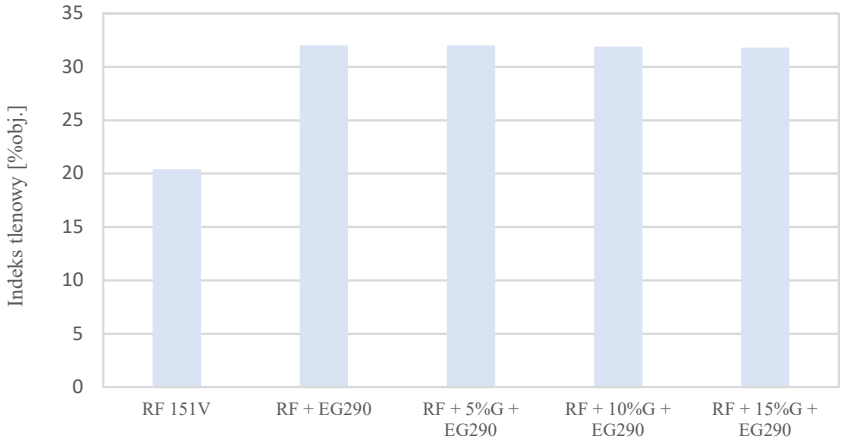


Rys. 5. Współczynnik przewodzenia ciepła badanych pianek poliuretanowych
Fig. 5. Thermal conductivity coefficient of the tested polyurethane foams

3.5. Indeks tlenowy

Wyniki oznaczenia indeksu tlenowego przedstawiono poniżej na rys. 6. Pianka referencyjna charakteryzowała się najniższą wartością tego parametru, co wskazuje na stosunkowo wysoką palność materiału.

Wprowadzenie grafitu ekspandowanego spowodowało wyraźny wzrost wartości indeksu tlenowego. Oznacza to, że do podtrzymania procesu spalania konieczne jest wyższe stężenie tlenu w atmosferze, co świadczy o poprawie odporności materiału na działanie ognia. Zarówno pianka referencyjna z dodatkiem grafitu ekspandowanego, jak i seria pianek o różnym udziale kwiatów goździka uważane są za niepalne.



Rys. 6. Indeks tlenowy badanych pianek poliuretanowych

Fig. 6. Limiting oxygen index of polyurethane foams

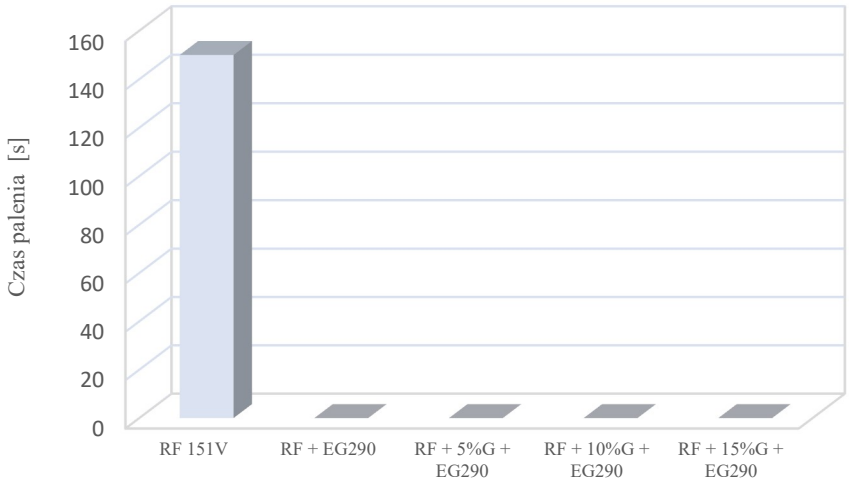
3.6. Test poziomy

Szczegółowa analiza przebiegu spalania badanych materiałów została przedstawiona na rys. 7.-9. Na rysunku 7. zaprezentowano czas spalania próbek, na rys. 8. szybkość rozprzestrzeniania się płomienia, a na rys. 9. ubytek masy materiału po zakończeniu testu. Zestawienie tych parametrów pozwala na kompleksową ocenę wpływu zastosowanych dodatków na zachowanie pianek poliuretanowych w warunkach oddziaływania płomienia.

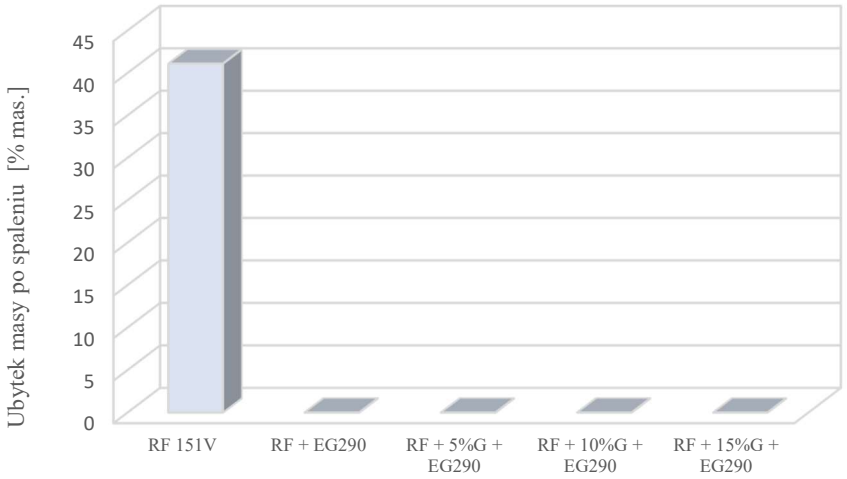
Pianka referencyjna ulegała szybkiemu spalaniu, a płomień rozprzestrzeniał się wzdłuż całej próbki w stosunkowo krótkim czasie (150 sekund). Materiał ten ulegał intensywnej degradacji termicznej, co prowadziło do stosunkowo dużego ubytku masy po zakończeniu badania.

Wprowadzenie grafitu ekspandowanego znacząco zmieniało przebieg procesu spalania. Pod wpływem wysokiej temperatury grafit ulegał ekspansji, tworząc spienioną warstwę ochronną na powierzchni materiału. Warstwa ta ograniczała dostęp tlenu oraz spowalniała rozprzestrzenianie się płomienia.

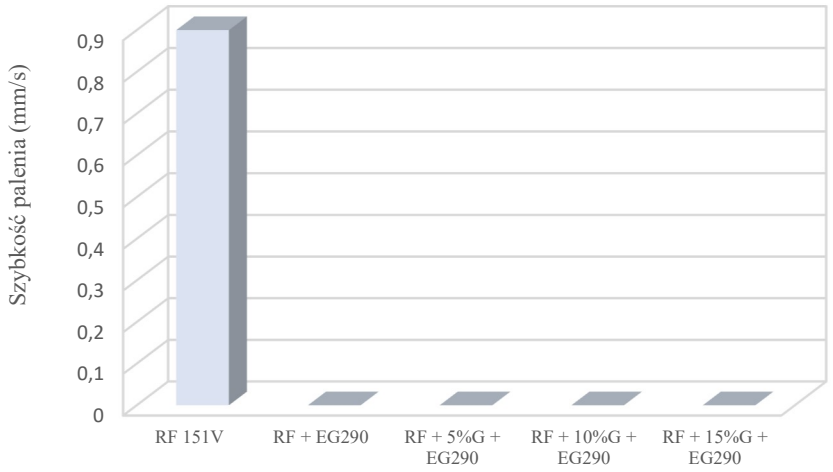
W przypadku pianek zawierających grafit oraz kwiaty goździka obserwowano podobny mechanizm działania. Powstająca warstwa ochronna ograniczała degradację materiału i nie zaobserwowano trwałego podtrzymania procesu spalania. Po usunięciu źródła płomienia próbki ulegały samoczynnemu wygaszeniu, co skutkowało zerową szybkością rozprzestrzeniania się płomienia oraz brakiem istotnego ubytku masy badanych materiałów.



Rys. 7. Czas spalania badanych pianek poliuretanowych w teście poziomym
Fig. 7. Burning time of polyurethane foams in the horizontal flammability test



Rys. 8. Ubytek masy pianek poliuretanowych po teście poziomym
Fig. 8. Mass loss of polyurethane foams after the horizontal burning test



Rys. 9. Szybkość spalania badanych pianek poliuretanowych
Fig. 9. Burning rate of the tested polyurethane foam

4. Podsumowanie

Na podstawie przeprowadzonych badań stwierdzono, że dodatek grafitu ekspandowanego w istotny sposób wpływa na właściwości fizyczne oraz ogniowe sztywnych pianek poliuretanowych. Wprowadzenie grafitu prowadziło do zwiększenia gęstości pozornej materiałów oraz znaczącej poprawy ich odporności na działanie ognia, co potwierdzono wzrostem wartości indeksu tlenowego oraz zmianą charakteru spalania próbek w teście poziomym.

Mechanizm działania grafitu ekspandowanego polega na jego ekspansji pod wpływem wysokiej temperatury i tworzeniu spienionej warstwy ochronnej na powierzchni materiału. Warstwa ta ogranicza dostęp tlenu do materiału oraz spowalnia proces degradacji termicznej pianek poliuretanowych.

Dodatek kwiatów goździka jako bio-wypełniacza wpływał przede wszystkim na strukturę komórkową materiałów, co znajdowało odzwierciedlenie w zmianach takich parametrów jak chłonność wody oraz stabilność wymiarów pianek. Wraz ze wzrostem zawartości dodatku roślinnego obserwowano zwiększenie chłonności wody oraz niewielkie zmiany gęstości materiałów.

Uzyskane wyniki wskazują, że połączenie grafitu ekspandowanego z bio-wypełniaczem roślinnym może stanowić interesujący kierunek modyfikacji pianek poliuretanowych, umożliwiający otrzymanie materiałów o zwiększonej odporności ogniowej oraz zmodyfikowanych właściwościach fizycznych.

Literatura

[1] Bourbigot S., Duquesne S., Fire retardant polymers, *Journal of Materials Chemistry*, 2007, 17, pp. 2283–2300.

[2] Camino G., Costa L., Martinasso G., Intumescent fire-retardant systems, *Polymer Degradation and Stability*, 1989, 23, pp. 359–376.

[3] Członka S., Strąkowska A., Strzelec K., Natural fillers in polyurethane foams – a review, *Journal of Cellular Plastics*, 2019, 55, pp. 1–28.

[4] Hull T., Stec A., *Fire Retardancy of Polymeric Materials*, CRC Press, 2009, pp. 1–356.

[5] Kairytė A., Członka S., Boris R., *Evaluation of the performance of bio-based rigid polyurethane foam with high amounts of sunflower press cake particles*, Materials, 2021, 14(19), s. 5475.

[6] Kurańska M., Prociak A., Michałowski S., Cabulis U., Kirpluks M., *Microcell lose as a natural filler in polyurethane foams based on the biopo yol from rapeseed oil*, Polimery, 2016, 61(9), s. 625–632.

[7] Levchik S., Weil E., Flame retardancy of polyurethane foams, *Polymer International*, 2004, 53, pp. 1585–1610.

[8] Sonnenschein M., *Polyurethanes: Science, Technology, Markets and Trends*, Wiley, 2015, pp. 1–482.

[9] Weil E., Levchik S., *Flame Retardants for Plastics and Textiles*, Hanser Publishers, 2009.

INFLUENCE OF EXPANDED GRAPHITE ON THE PHYSICAL AND FIRE PROPERTIES OF RIGID POLYURETHANE FOAMS CONTAINING CLOVE FLOWERS

Summary

The aim of this study was to determine the influence of expandable graphite on the physical and flammability properties of rigid polyurethane foams containing a bio-filler in the form of dried clove flowers. Within the scope of the research, a series of materials was prepared, including a reference foam, a composition containing expandable graphite, and systems containing both graphite and different amounts of the plant-based additive. The incorporation of expandable graphite was intended to improve the fire resistance of the materials, while the introduction of the bio-filler was aimed at modifying the structure and selected physical properties of polyurethane foams. A number of tests were carried out to evaluate the properties of the obtained materials. Apparent density, water absorption, dimensional stability, and thermal conductivity were determined. In addition, the flammability behavior of the foams was analyzed by determining the limiting oxygen index and performing a horizontal burning test.

The results indicate that the addition of expandable graphite leads to an increase in the density of polyurethane foams and significantly improves their fire resistance. The presence of graphite also affects the combustion process by forming a protective char layer on the material surface, which limits oxygen access and reduces flame propagation. The presence of clove flowers in the foam structure influenced selected physical properties of the materials, mainly due to changes in the cellular structure of the foams. The obtained results confirm that expandable graphite can be an effective additive reducing the flammability of polyurethane foams modified with plant-based bio-fillers.

Keywords: polymer composites, flame retardants, plant fillers, fire properties, thermal insulation materials.

Dorota Głowacz-Czerwonka, Aleksandra Brożyna
Koło Naukowe Studentów Chemii „Esprit”

WPŁYW PŁATKÓW RÓŻY I GRAFITU EKSPANDOWANEGO NA MORFOLOGIĘ I ODPORNOŚĆ OGNIOWĄ SZTYWNYCH PIANEK POLIURETANOWYCH

Streszczenie

Rozdział przedstawia analizę wpływu płatków róży na morfologię sztywnych pianek poliuretanowych. Ponadto zbadano również, w jaki sposób modyfikacja grafitem ekspandowanym uzyskanych materiałów oddziałuje na ich właściwości palne. W celu przeprowadzenia analizy struktury materiału wykonano obserwacjez wykorzystaniem mikroskopu polaryzacyjnego, które umożliwiły określenie wielkości porów w strukturze pianek poliuretanowych. W ramach badań palności materiałów wyznaczono indeks tlenowy oraz przeprowadzono test palności w układzie poziomym, co pozwoliło na ocenę zachowania pianek podczas kontaktu z ogniem. Zastosowanie naturalnego napełniacza wpłynęło na zmianę struktury badanych poliuretanów. Wraz ze wzrostem zawartości biomasy różanej do 5% i 10% zaobserwowano malejącą średnią średnicę komórek. Z kolei wprowadzenie grafitu ekspandowanego przyczyniło się do zmniejszenia średnicy porów w porównaniu do próbki kontrolnej, jednak wartości te pozostawały nadal większe niż w przypadku pianek modyfikowanych wyłącznie dodatkiem biomasy różanej. Analiza wyników testu poziomego wskazała natomiast na brak wyraźnej zależności pomiędzy ilością zastosowanych płatków róży, a czasem spalania tworzyw. Wszystkie warianty modyfikowane wyłącznie biomasą wykazały drogę palenia równą 100%. W badaniu indeksu tlenowego zaobserwowano, iż zastosowanie naturalnego napełniacza nie powoduje istotnych różnic w wartościach tego parametru w porównaniu z piankami kontrolnymi. Uzyskane wyniki wskazują na możliwość modyfikacji pianek grafitem ekspandowanym oraz płatkami róży, otrzymując tym samym materiały o uporządkowanej strukturze oraz podwyższonej odporności termicznej.

Słowa kluczowe: poliuretany, płatki róży, biomasa, uniepalniacz, indeks tlenowy.

1. Wprowadzenie

Pianki poliuretanowe należą do jednej z najszerzej rozwiniętych grup poliuretanów stosowanych w produkcji tworzyw sztucznych. Powstają w wyniku

reakcji poliaddycji izocyjanianu oraz poliolu, której towarzyszy wydzielanie gazu tworzącego porowatą strukturę. Ze względu na budowę komórkową możemy wyróżnić między innymi: pianki elastyczne, sztywne oraz półsztywne [3].

Pianki elastyczne otrzymywane są w wyniku syntezy liniowych poliestrów oraz polieterów. Przedstawiają niskie wartości gęstości oraz otwartokomórkową budowę. Znajdują szerokie zastosowanie ze względu na zdolność do odwracalności odkształceń, a także właściwości tłumiące dźwięk. Zazwyczaj wykorzystywane są w postaci płyt, bloków lub kształtek zarówno jako poduszki czy materace jak i elementy tapicerskie [4].

Struktura sztywnych i półsztywnych pianek opiera się na poliolach o rozgałęzionej budowie. Kluczowym parametrem determinującym ich twardość jest gęstość usieciowania, ściśle powiązana z dużą liczbą grup hydroksylowych (OH) zawartych w stosowanych substratach. Pod względem właściwości fizycznych wyróżniają się znacznie większą gęstością niż pianki elastyczne. Wysoka sztywność oraz odporność na odkształcenia sprawia, że pianki te są powszechnie stosowane jako podstawowe materiały konstrukcyjne. Tworzywa te znajdują zastosowanie nie tylko w budownictwie (np. jako termoizolacja), ale również w sektorach transportu, w tym w konstrukcjach kadłubów statków, poszyciach lotniczych oraz komponentach motoryzacyjnych [3, 5].

Poliuretany pomimo szerokiego zastosowania oraz wszechstronnej implementacji często nie spełniają podstawowych standardów bezpieczeństwa pożarowego. Czynnik ten wynika bezpośrednio ze stosowania substratów charakteryzujących się niską wartością indeksu tlenowego (LOI). Na odporność termiczną istotny wpływ ma również struktura materiału, a w szczególności stopień usieciowania danego polimeru. Wraz z dynamicznie rozwijającym sierynkiem poliuretanów, stale poszukuje się alternatywnych rozwiązań problemu ich wysokiej palności. Pianki poliuretanowe charakteryzują się dużą łatwością modyfikacji składu chemicznego, co pozwala na wprowadzanie do ich struktury skutecznych substancji uniepalniających lub opóźniających proces spalania [1].

Uniepalniacze to związki, których mechanizm opiera się głównie na fizycznym lub chemicznym ograniczaniu dyfuzji tlenu do strefy reakcji. W praktyce powszechnie stosowane są dwa sposoby wprowadzania antypirenów do układu. Za pomocą modyfikacji wewnętrznej lub zewnętrznej. Modyfikacja wewnętrzna polega na wykorzystaniu substancji reaktywnych, które stają się elementem łańcucha polimerowego, natomiast zewnętrzna na wprowadzeniu dodatków niereaktywnych, takich jak napełniacze czy powłoki [2].

Celem przeprowadzonej analizy było zbadanie wpływu dodatku płatków róży na strukturę otrzymanych materiałów, a także określenie w jaki sposób modyfikacja grafitem ekspandowanym wpływa na właściwości palne.

2. Część doświadczalna

2.1. Aparatura laboratoryjna

Sprzęt laboratoryjny wykorzystany do otrzymywania i badań kompozycji piankowych: mieszadło mechaniczne (Meramont WE27T, Polska), suszarka SLW 53 STD (Pol-Eko Aparatura, Polska), aparat do badania indeksu tlenowego LOI (Concept Fire Testing, UK), waga analityczna E42 (Radwag, Polska), waga techniczna PM 600 (Szwajcaria), Suwmiarka elektroniczna (NITEO, Polska), Polska), mikroskop polaryzacyjny, serii Pantera (Motic, Niemcy).

2.2. Otrzymanie pianek poliuretanowych

Przygotowanie kompozycji piankowych rozpoczęto od pianek referencyjnych (bez dodatków oraz z grafitem ekspandowanym EG 290).

Otrzymano dwie serie pianek, pierwsza z nich zawierała płatki róży w ilości 5, 10 oraz 15% masowych natomiast druga zawierała analogiczną ilość kwiatów oraz dodatek uniepalniacza w stałej ilości. W tym celu wykorzystano poliol Rokopol RF 151V (PCC Rokita, Polska), silikon L-6900 (Momentive Performance Materials, Belgia), trietyloaminę (Chempur, Polska), wodę dejonizowaną, płatki roży (Greenok Place, Egipt) oraz w przypadku pianek z grafitem EG290 (Sinograf, Polska). W kolejnym kroku dodawano 4.4’-diizocyjanian difenytometanu Ongronat 2100 (BorsodChem, Węgry).

W pierwszym etapie przygotowywano mieszaninę poliolową poprzez połączenie poliolu z trietyloaminą (15 kropli), silikonem (5 kropli), wodą dejonizowaną (2% mas.) oraz rozdrobnionymi płatkami róży w ilości 5, 10 lub 15% mas. W przypadku drugiej serii materiałów do mieszaniny wprowadzano również grafit ekspandowalny EG290 w ilości 60% mas (12g). Płatki róży wprowadzono do układu bez dodatkowej obróbki termicznej, rozdrabniając je uprzednio ręcznie w moździerzu laboratoryjnym do postaci drobnych, nieregularnych cząstek. Wszystkie składniki fazy poliolowej homogenizowano w temperaturze pokojowej z wykorzystaniem mieszadła mechanicznego do uzyskania możliwie jednorodnego rozproszenia napełniaczy. Następnie do przygotowanej przedmieszki dodawano odpowiednią ilość 4,4'-diizocyjanianu difenylometanu i całość ponownie intensywnie mieszano. Po zakończeniu mieszania mieszaninę reakcyjną niezwłocznie wylewano do formy. Otrzymane materiały pozostawiano następnie w warunkach laboratoryjnych do całkowitego zakończenia procesu sieciowania.

Otrzymano następujące kompozycje:

- Ref. RF 151V (RF) i Ref. RF 151V (RF) + EG290 (rys. 1.).



Rys. 1. Otrzymane pianki poliuretanowe. Od lewej: Pianka referencyjna bez dodatków, pianka referencyjna z dodatkiem grafitu.

Pic. 1. Obtained polyurethane foams. From left: Reference foam without additives, reference foam with graphite added.

oraz: RF+ róża 5%, RF+ róża 10%, RF+ róża 15%, RF+ róża 5% + EG290RF+ róża 10% + EG290, RF+ róża 15% + EG290 (rys. 2.).



Rys. 2. Otrzymane pianki poliuretanowe. Od lewej: pianki z dodatkiem płatków róży, z dodatkiem płatków róży i grafitu ekspandowanego

Pic. 2. Obtained polyurethane foams. From left to right: foams with added rose petals, with added rose petals and expanded graphite

2.3. Badania

2.3.1. Morfologia pianek

Wpływ modyfikacji pianek na strukturę otrzymanych materiałów przeanalizowano z wykorzystaniem mikroskopu polaryzacyjnego z serii Pantera firmy Motic (Niemcy). Z uzyskanych próbek wycięto kształtki w postaci sześcianów, a następnie przeprowadzono pomiary wielkości porów (tab. 1.).

Tabela (tablica) 1. Średnice komórek pianek poliuretanowych

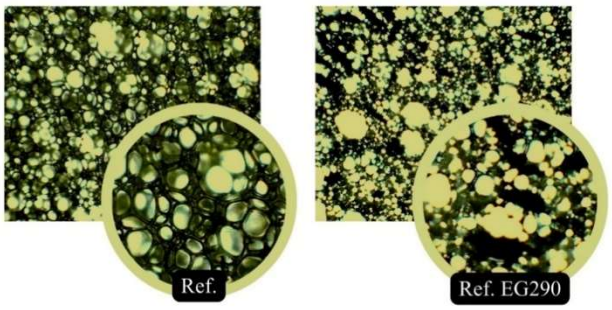
Table 1. Diameters of polyurethane foam cells

Lp.	Nazwa kompozycji	D średnie [µm]
1.	Ref. RF 151V (RF)	160,3
2.	RF+ róża 5%	174,3
3.	RF+ róża 10%	155,3
4.	RF+ róża 15%	221,0
5.	Ref. RF 151V (RF) + EG290	251,5
6.	RF+ róża 5% + EG290	242,3
7.	RF+ róża 10% + EG290	231,5
8.	RF+ róża 15% + EG290	186,5

gdzie: D średnie – średnia średnica wszystkich komórek [µm].

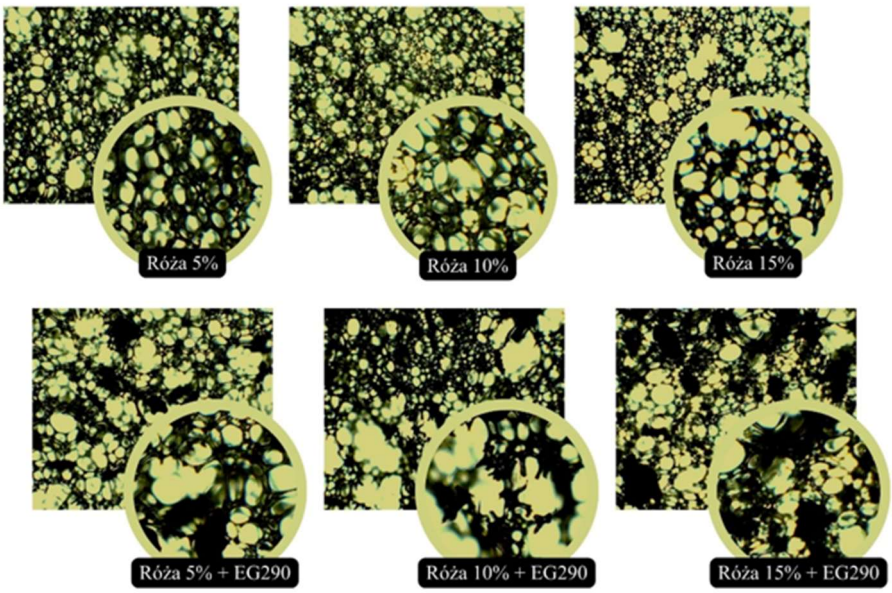
Wprowadzenie do matrycy poliuretanowej biomasy różanej w istotny sposób wpływa na strukturę otrzymanych materiałów. W przypadku pianek z 5% i 10% zawartością róży zaobserwowano malejącą średnią średnicę wraz z rosnącą ilością napełnienia. Taka ilość napełniacza umożliwiła równomierne rozłożenie surowca w matrycy poliuretanowej oraz nie zaburzała wzrostu komórek.

Pianka zawierająca największą ilość napełniacza charakteryzowała się największą średnią średnicą komórek co mogło być spowodowane zbyt dużą ilością fazy stałej w składzie. Skutkiem tego zjawiska była aglomeracja substratu oraz niejednorodny wzrost komórek. Kompozycje modyfikowane grafitem wykazały mniejsze wartości średniej średnicy w porównaniu do pianki referencyjnej. Ponadto średnica komórek malała wraz z rosnącą ilością napełnienia. Świadczy to o tym, iż dodatek grafitu nie wpływa w negatywny sposób na wzrost pianek oraz pozwala na uzyskanie bardziej drobnokomórkowej struktury niż referencja.



Rys. 3. Obraz pianek referencyjnych uzyskany w mikroskopie polaryzacyjnym

Pic. 3. Picture of reference foams obtained in a polarizing microscope



Rys. 4. Obraz pianek poliuretanowych z dodatkiem róży oraz grafitu uzyskany w mikroskopie polaryzacyjnym

Pic. 4. Picture of polyurethane foams with added rose and graphite obtained in a polarizing microscope

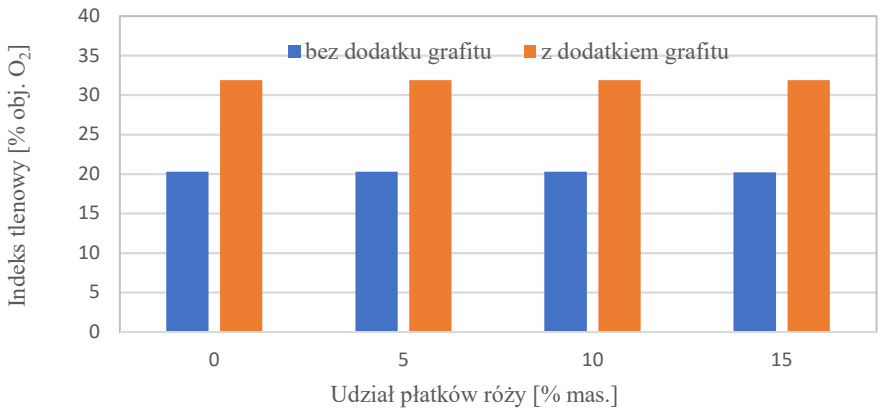
Analizując obraz uzyskany z mikroskopu polaryzacyjnego (rys. 3. oraz rys. 4.) zaobserwowano, iż próbki zawierające dodatek płatków róży charakteryzują się bardziej jednorodną i drobnokomórkową budową niż próbki modyfikowane grafitem. Potwierdzeniem tego są również wyniki zaprezentowane w powyższej tabeli (tab. 1.). W piankach zawierających grafit można zaobserwować ciemniejsze obszary, które wskazują na obecność aglomeratów grafitu.

2.3.2. Indeks tlenowy

Badanie indeksu tlenowego przeprowadzono z wykorzystaniem podłużnych beleczek o wymiarach 10 x 1,2 x 1,2 cm. Przygotowane próbki umieszczano w aparaturze badawczej zapewniającej stały przepływ mieszaniny tlenu i azotu. Wyznaczono minimalną wartość tlenu podtrzymującą spalanie zgodnie z procedurą [PN-EN ISO 4589-2:2017-06].

Analizując otrzymane wyniki (rys. 5.) zaobserwowano, iż wprowadzenie do układu pianek poliuretanowych płatków róży nie wpłynęło w znaczący sposób na wartości indeksu tlenowego. Ponadto zarówno próbka zawierająca 5% i 10% dodatek biomasy wykazuje taką samą wartość jak pianka referencyjna.

W przypadku pianek modyfikowanych grafitem wszystkie przedstawiały wartości równe próbie kontrolnej.



Rys. 5. Wpływ składu kompozycji na wartość indeksu tlenowego

Pic. 5. The effect of composition on the oxygen index value

Ponadto jego dodatek wpłynął na wyraźny wzrost wskaźnika LOI w stosunku do próbek niemodyfikowanych, co bezpośrednio przekłada się na zwiększenie odporności ogniowej uzyskanego materiału.

2.3.3. Test poziomy

Badanie przeprowadzono zgodnie z normą [PN-EN ISO 3582:2002]. W tym celu przygotowano kształtki pianek, które następnie umieszczono w uchwycie badawczym w komorze bez przeciągów. Każdy z badanych materiałów podpalano przez około 30 s, a następnie analizowano zachowanie próbki po odsunięciu źródła ognia, mierząc czas, długość spalania oraz ubytek masy (tab. 2.).

Tabela (tablica) 2. Parametry spalania uzyskane w wyniku badania testu poziomego.

Table 2. Combustion parameters obtained from the horizontal test

Lp.	Nazwa kompozycji	Droga palenia [% liniowy]	Czas palenia [s]	Ubytek masy [% mas.]
1.	Ref. RF 151V (RF)	100	150	41,06
2.	RF+ róża 5%	100	85	48,55
3.	RF+ róża 10%	100	133	43,23
4.	RF+ róża 15%	100	87	54,67
5.	Ref. RF 151V (RF) + EG290	0	0	0,00
6.	RF+ róża 5% + EG290	0	0	0,00
7.	RF+ róża 10% + EG290	0	0	0,00
8.	RF+ róża 15% + EG290	0	0	0,00

W przypadku pianek modyfikowanych biomasą różaną nie zauważono zależności pomiędzy ilością dodanych płatków róży, a czasem spalania. Wszystkie z tych próbek (rys. 6.) charakteryzowały się drogą palenia równą 100%. Próbki wykazały większą szybkość spalania niż próbka referencyjna. Dodatkowo warianty z 5% i 10% dodatkiem róży miały mniejszy ubytek masy w porównaniu do materiału odniesienia. Zastosowanie grafitu ekspandowanego ograniczyło drogę spalania do 0%.



Rys. 6. Próbki przed i po teście poziomym. Od lewej: Próbki zawierające dodatek płatków róży oraz płatków róży i grafitu

Pic. 6. Samples before and after the horizontal test. From left to right: Samples containing rose petals and rose petals and graphite

3. Podsumowanie

Zastosowanie naturalnego napełniacza oraz grafitu ekspandowanego odpowiedzialne jest za zmiany w morfologii otrzymanych tworzyw. W przypadku biomasy różanej zastosowanej w ilości 5% oraz 10% zaobserwowano, iż średnica komórek maleje wraz ze wzrostem napełnienia. Kompozyty modyfikowane grafitem wykazały niższe wartości średniej średnicy komórek w porównaniu do próbki kontrolnej. Uzyskane wyniki dowodzą braku negatywnego wpływu zastosowania grafitu na wzrost pianek, umożliwiając uformowanie struktury bardziej jednorodnej i drobnokomórkowej niż referencja. Należy jednak zaznaczyć, że średnica tych komórek pozostawała nadal większa niż w przypadku próbek zawierających wyłącznie dodatek róży. W przypadku badania indeksu tlenowego nie zaobserwowano, aby badane próbki odbiegały w istotny sposób od pianek referencyjnych,

natomiast dodatek grafitu pozwolił na uzyskanie pianek o charakterze niepalnym. Analiza wartości otrzymanych w wyniku testu poziomego wykazała, iż nie ma zależności pomiędzy ilością zastosowanych płatków róży, a czasem palenia tych tworzyw. Wszystkie warianty modyfikowane wyłącznie biomasą charakteryzowały się całkowitą drogą palenia. Zastosowanie grafitu ekspandowanego pozwoliło na uzyskanie materiałów odpornych na płomień o drodze palenia równej 0%.

Literatura

[1] Czech-Polak J. i in.: Sztywne pianki poliuretanowe o zwiększonej odporności na płomień, Polimery, t. 63, nr 2, 2018, s. 115-124.

[2] McKenna S.T., Hull T.R.: The fire toxicity of polyurethane foams, Fire Science Reviews, vol. 5, no. 3, 2016.

[3] Prociak A., Rokicki G., Ryszkowska J.: Materiały poliuretanowe, Wydawnictwo Naukowe PWN, Warszawa 2016, s. 250-264.

[4] Swinarew B.: Poliuretany – nowoczesne wszechstronne materiały. Część II – pianki poliuretanowe, Przetwórstwo Tworzyw, nr 5, 2015, s. 428-434.

[5] United Nations Environment Programme: Estimated global polymer foam production (Rigid and Flexible PU foams) 2017-2023, Nairobi 2018, s. 50-51.

THE INFLUENCE OF ROSE PETALS AND EXPANDED GRAPHITE ON THE MORPHOLOGY AND THERMAL RESISTANCE OF RIGID POLYURETHANE FOAMS

Summary

The article presents an analysis of the effect of rose petals on the morphology of rigid polyurethane foams. Additionally, examined how the modification of the obtained materials with expanded graphite affects their flammability. Observations were conducted to analyze the material structure using a polarizing microscope, which allowed us to determine the size of the pores in the structure of polyurethane foams. As part of the flammability research, the oxygen index was determined and a flammability test was conducted in a horizontal arrangement, which allowed for the evaluation of the foams' behavior during contact with fire. The use of the natural filler affected the change in the structure of the polyurethanes tested. As the content of rose biomass increased to 5% and 10%, the average cell diameter decreased. The introduction of expanded graphite resulted in a reduction of the pore diameter compared to the control sample, but the values were still higher than in the case of foams modified only with the addition of rose biomass. The analysis of the horizontal test results, however, indicated no clear correlation between the amount of rose petals used and the time combustion of the material. All variants modified only with biomass were characterized by a complete burning path. The test of the oxygen index showed, that the use of a natural filler did not cause significant differences in terms of this parameter compared to the control foams. The results obtained indicate the possibility of modifying foams with expanded graphite and rose petals, which allows for obtaining materials with an ordered structure and increased thermal resistance.

Keywords: polyurethanes, rose petals, biomass, flame retardant, oxygen index.

Andżelika Gronko
Studenckie Koło Naukowe Inżynierii Środowiska

OCENA SATYSFAKCJI KLIENTÓW PRZEDSIĘBIORSTWA WODOCIĄGOWEGO Z UWZGLĘDNIENIEM ASPEKTÓW SPOŁECZNYCH, TECHNICZNYCH I EKONOMICZNYCH

Streszczenie

Badania opinii użytkowników stanowią istotne narzędzie wspomagające doskonalenie funkcjonowania systemów zaopatrzenia w wodę oraz wzmacnianie ich bezpieczeństwa. Celem pracy była ocena jakości i aktualności informacji przekazywanych przez przedsiębiorstwo wodociągowe, poziomu zainteresowania mieszkańców zagadnieniami oszczędzania i jakości wody, a także potrzeby prowadzenia kampanii edukacyjnych oraz gotowości odbiorców do udziału w inicjatywach społecznych. Przeprowadzono badanie ankietowe wśród 100 losowo wybranych mieszkańców korzystających z usług lokalnego wodociągu. Kwestionariusz obejmował 25 pytań dotyczących satysfakcji z usług wodociągowych, w tym oceny jakości wody, komunikacji z odbiorcami, działań edukacyjnych oraz inicjatyw społecznych. Do oceny poziomu satysfakcji zastosowano wskaźnik CSAT (Customer Satisfaction Score). Uzyskane wyniki wskazują na stabilny i stosunkowo wysoki poziom zadowolenia odbiorców, przy czym w większości analizowanych obszarów wartości wskaźnika CSAT mieściły się w przedziale 60–80%. Rezultaty badań podkreślają znaczenie kontynuacji działań informacyjnych i edukacyjnych w celu utrzymania wysokiego poziomu jakości usług wodociągowych.

Słowa kluczowe: badanie ankietowe, system zaopatrzenia w wodę, ocena jakości usług wodociągowych.

1. Wstęp

Zapewnienie niezawodności i bezpieczeństwa dostaw wody stanowi jedno z kluczowych zadań przedsiębiorstw wodociągowych [1]. Współczesne zarządzanie systemami zaopatrzenia w wodę nie ogranicza się jedynie do kwestii technicznych i eksploatacyjnych, lecz obejmuje również aspekty społeczne i ekonomiczne [2], w tym orientację na odbiorcę usług, dla którego istotnym elementem jest regularna ocena poziomu satysfakcji.

Podstawowym narzędziem badań opinii są sondaże [3]. Zgodnie z definicją Wielkiego słownika języka polskiego PAN, sondaż polega na zadawaniu określonej grupie osób tego samego zestawu pytań w celu poznania ich poglądów [4]. W sektorze wodociągowym badania tego typu pozwalają ocenić jakość świadczonych usług, zidentyfikować obszary wymagające usprawnień oraz dostosować działania przedsiębiorstwa do realnych potrzeb odbiorców.

Analiza opinii konsumentów wody ma szczególne znaczenie, ponieważ pozwala na identyfikację problemów związanych z funkcjonowaniem systemu wodociągowego, odnosząc się do jakości wody, ciągłości dostaw wody czy parametrów technicznych funkcjonowania systemu np. wysokości ciśnienia [5]. Satysfakcja klientów nie powinna być rozpatrywana wyłącznie w aspekcie technicznym; równie istotne są czynniki społeczne, takie jak przejrzystość komunikacji, informowanie o przerwach w dostawie, reagowanie na zgłoszenia oraz poziom zaufania społecznego do przedsiębiorstwa.

Badanie satysfakcji pełni zarówno funkcję kontrolną, jak i strategiczną. Umożliwia ocenę skuteczności działań informacyjno-edukacyjnych, poziomu świadomości odbiorców w zakresie racjonalnego gospodarowania wodą oraz ich zaangażowania w inicjatywy prośrodowiskowe. Uwzględnia także wymiar ekonomiczny, obejmujący postrzeganie wysokości opłat, relację ceny do jakości usług oraz akceptację kosztów inwestycji infrastrukturalnych. Kompleksowe uwzględnienie aspektów społecznych, technicznych i ekonomicznych wspiera efektywne zarządzanie przedsiębiorstwem wodociągowym, sprzyjając budowaniu trwałych relacji z odbiorcami opartych na zaufaniu i transparentności.

2. Materiał i metody

Na potrzeby niniejszej pracy przeprowadzono badanie ankietowe wśród mieszkańców korzystających z usług lokalnego wodociągu w mieście wojewódzkim, położonym w południowo-wschodniej Polsce. Ankieta miała charakter hybrydowy – została udostępniona zarówno w formie elektronicznej (Google Forms), jak i papierowej, przekazywanej bezpośrednio respondentom. Kwestionariusze dystrybuowano wśród losowo wybranych mieszkańców badanego obszaru, co umożliwiło pozyskanie zróżnicowanych opinii – zarówno osób świadomie oceniających funkcjonowanie przedsiębiorstwa, jak i tych, które dokonywały takiej oceny po raz pierwszy. Kwestionariusz obejmował 25 pytań zamkniętych opartych na pięciostopniowej skali Likerta. W badaniu wzięło udział 100 osób, co pozwoliło uzyskać wiarygodny obraz opinii odbiorców usług wodociągowych w analizowanym mieście. W tabeli 1. przedstawiono kwestionariusz ankietowy.

Pomiar satysfakcji klientów przedsiębiorstwa wodociągowego oparto na wskaźniku CSAT (Customer Satisfaction Score), będącym jednym z najczęściej wykorzystywanych narzędzi oceny zadowolenia klientów [6]. Metoda ta polega na bezpośrednim określeniu poziomu satysfakcji poprzez odpowiedź na pytanie ogólne dotyczące oceny usługi [7]. Odpowiedzi udzielane są zazwyczaj na pięcio-

lub siedmiostopniowej skali, rozpiętej pomiędzy skrajnymi ocenami negatywnymi i pozytywnymi, z możliwością wskazania odpowiedzi neutralnej. W niniejszym badaniu zastosowano pięciostopniową skalę Likerta, co zapewniło spójność i porównywalność wyników [8].

Tabela 1. Treść pytań z kwestionariusza ankietowego dotyczącego satysfakcji klientów lokalnego wodociągu

Table 1. Customer satisfaction survey questions for the local water utility

Lp.	Pytanie	Skala ocen
1	Jak ocenia Pan/Pani jakość dostarczanej wody pod względem smaku?	bardzo źle – 1 – 5 – bardzo dobrze
2	Jak ocenia Pan/Pani zapach wody z kranu?	bardzo źle – 1 – 5 – bardzo dobrze
3	Jak ocenia Pan/Pani przejrzystość, klarowność wody?	bardzo źle – 1 – 5 - bardzo dobrze
4	W jakim stopniu uważa Pan/Pani wodę wodociągową za bezpieczną do spożycia?	bardzo niskim – 1 – 5 – bardzo wysokim
5	Jak często w ostatnim roku wystąpiły przerwy w dostawie wody?	bardzo często – 1 – 5 – nigdy
6	Jak ocenia Pan/Pani sposób informowania o przerwach w dostawie wody?	bardzo niezadowalający – 1 – 5 – bardzo zadowalający
7	Jak ocenia Pan/Pani łatwość załatwienia sprawy (telefonicznie, mailowo, osobiście)?	bardzo źle – 1 – 5 – bardzo dobrze
8	Czy dostępne kanały kontaktu (telefon, e-mail, eBOK) są dla Pana/Pani wystarczające?	zdecydowanie nie – 1 – 5 – zdecydowanie tak
9	Jak ocenia Pan/Pani przejrzystość rachunków i faktur?	bardzo źle – 1 – 5 – bardzo dobrze
10	Jak ocenia Pan/Pani łatwość odczytania wskazań wodomierza i zgłaszania zużycia?	bardzo źle – 1 – 5 – bardzo dobrze
11	Czy uważa Pan/Pani ceny zaopatrzenia w wodę i odbioru ścieków za adekwatne do jakości usług?	zdecydowanie nie – 1 – 5 – zdecydowanie tak
12	Jak ocenia Pan/Pani komunikację przedsiębiorstwa (ostrzeżenia, komunikaty, informacje o inwestycjach)?	bardzo źle – 1 – 5 – bardzo dobrze
13	Czy informacje publikowane na stronie internetowej lub w mediach społecznościowych są zrozumiałe i aktualne?	zdecydowanie nie – informacje są niejasne i/lub nieaktualne – 1 –5 – zdecydowanie tak – informacje są w pełni zrozumiałe i zawsze aktualne
14	Jak ogólnie ocenia Pan/Pani działanie przedsiębiorstwa wodociągowego?	bardzo źle – 1 – 5 – bardzo dobrze
15	Czy jest Pan/Pani zainteresowany(a) materiałami edukacyjnymi nt. oszczędzania wody lub jakości wody?	zdecydowanie nie – w ogóle nie zainteresowany(a) – 1 – 5 – zdecydowanie tak – bardzo zainteresowany(a)
16	Czy chciał(a)by Pan/Pani, aby wodociągi prowadziły kampanie edukacyjne dotyczące oszczędzania wody i ochrony środowiska?	Zdecydowanie nie – nie uważam, że powinny być prowadzone – 1 –5 – Zdecydowanie tak – bardzo popieram prowadzenie takich kampanii

Tabela 1 (cd.). Treść pytań z kwestionariusza ankietowego dotyczącego satysfakcji klientów lokalnego wodociągu

Table 1 (cont.). Customer satisfaction survey questions for the local water utility

Lp.	Pytanie	Skala ocen
17	Czy jest Pan/Pani zainteresowany(a) materiałami wyjaśniającymi, jak działa oczyszczalnia ścieków i system dostarczania wody?	zdecydowanie nie – w ogóle nie zainteresowany(a) – 1 – 5 – zdecydowanie tak – bardzo zainteresowany(a)
18	Czy uważa Pan/Pani, że wodociągi powinny prowadzić więcej działań informacyjnych i społecznych, np. dni otwarte, spotkania z mieszkańcami, warsztaty?	Zdecydowanie nie – nie popieram prowadzenia takich działań – 1 – 5 – Zdecydowanie tak – zdecydowanie popieram zwiększenie takich działań
19	Czy był(a)by Pan/Pani zainteresowany(a) wizytą edukacyjną / dniem otwartym w przedsiębiorstwie wodociągowym?	zdecydowanie nie – w ogóle nie zainteresowany(a) – 1 – 5 –zdecydowanie tak – bardzo zainteresowany(a)
20	Na ile ważne jest dla Pana/Pani, aby przedsiębiorstwo wodociągowe działało w sposób przyjazny środowisku?	wcale nie ważne – 1 – 5 – bardzo ważne
21	Czy uważa Pan/Pani, że wodociągi dostatecznie informują o swoich działaniach w zakresie ochrony środowiska i oszczędzania wody?	zdecydowanie nie – 1 – 5 – zdecydowanie tak
22	Czy podejmuje Pan/Pani w domu działania wodooszczędne (np. krótsze prysznice, oszczędna spłuczka, zbieranie deszczówki)?	nigdy – 1 – 5 – zawsze
23	Jak dużym zaufaniem darzy Pan/Pani jakość wody z kranu?	bardzo małym – 1 – 5 – bardzo dużym
24	Na ile czuje się Pan/Pani komfortowo pijąc wodę z kranu?	czuję się bardzo niekomfortowo – 1 – 5–czuję się bardzo komfortowo
25	Czy informacje o modernizacjach, badaniach jakości i działaniach wodociągów zwiększają Pana/Pani poczucie bezpieczeństwa?	zdecydowanie nie – 1–5 – zdecydowanie tak

Wartość wskaźnika CSAT może być wyrażona jako odsetek odpowiedzi pozytywnych w ogólnej liczbie wskazań bądź jako średnia arytmetyczna ocen. W przeprowadzonym badaniu przyjęto najczęściej stosowane w praktyce podejście typu Top-2 Box, polegające na uwzględnieniu wyłącznie dwóch najwyższych ocen („4" i „5") jako odpowiedzi pozytywnych. Wskaźnik CSAT obliczono jako odsetek pozytywnych odpowiedzi (oceny 4 i 5 na skali Likerta) zgodnie z formułą [9]:

$$\text{CSAT}\ (\%) = \left(\frac{\text{liczba ocen pozytywnych}}{\text{całkowita liczba odpowiedzi}}\right) \cdot 100\% \qquad (1)$$

gdzie: **CSAT** – wskaźnik poziomu zadowolenia klientów na podstawie wyników ankiety.

Tak obliczona wartość pozwala określić udział respondentów zadowolonych z usług przedsiębiorstwa. Interpretacja wyników została przeprowadzona w oparciu o poziomy satysfakcji przedstawione w tabeli 2.

Tabela 2. Interpretacja wyników CSAT – skala procentowa
(opracowanie własne na podstawie [10])

Table 2. Interpretation of CSAT results – percentage scale

0 – 40%	40 – 60%	60 – 80%	Powyżej 80%
Poziom niski	Poziom umiarkowany	Poziom dobry	Poziom bardzo wysoki

Skuteczność metody zależy od właściwej konstrukcji pytań, ponieważ ich zakres decyduje o tym, czy badanie odnosi się do pojedynczych kontaktów z przedsiębiorstwem, czy do ogólnego doświadczenia klienta. W niniejszej pracy CSAT wykorzystano do oceny całościowego poziomu satysfakcji mieszkańców z usług wodociągowych, obejmujących m.in. jakość wody, komunikację, działania edukacyjne oraz inicjatywy społeczne.

3. Dyskusja otrzymanych wyników

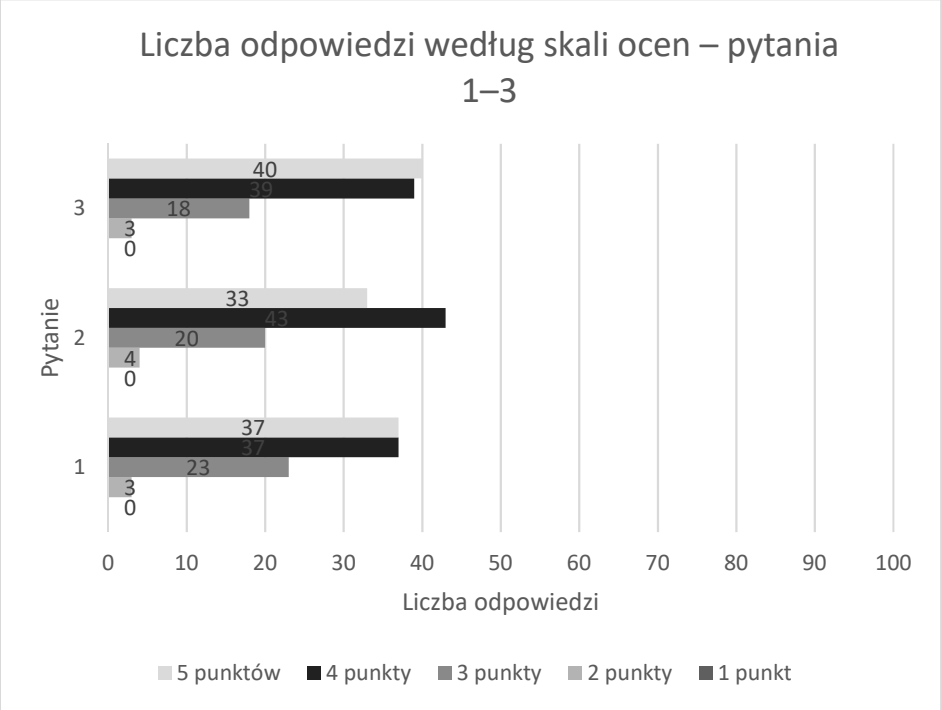
3.1. Jakość i bezpieczeństwo

Pierwsza grupa pytań obejmowała ocenę jakości wody wodociągowej w dwóch komplementarnych wymiarach: cech sensorycznych oraz poczucia bezpieczeństwa jej spożycia.

Na rysunku 1. przedstawiono wyniki ankiety ilustrujące poziom satysfakcji respondentów w zakresie smaku, zapachu i klarowności dostarczanej wody, ukazujące porównanie liczby odpowiedzi w poszczególnych punktach skali.

Analiza pytań dotyczących smaku, zapachu i przejrzystości – które bezpośrednio wpływają na codzienne doświadczenia użytkowników i postrzeganie jakości usługi – wskazuje na dominację ocen wysokich (4 i 5 punktów) we wszystkich trzech aspektach. W przypadku smaku 74% respondentów przyznało oceny pozytywne (po 37% wskazań na 4 i 5 punktów), przy udziale 26% ocen niższych. Zapach oceniono nieco wyżej – 76% badanych wskazało 4 lub 5 punktów (43% i 33%), co sugeruje brak istotnych zastrzeżeń w tym zakresie. Najwyższy poziom akceptacji dotyczył przejrzystości i klarowności wody – 79% ocen wysokich (39% – 4 punkty, 40% – 5 punktów) i 21% niższych, co potwierdza bardzo korzystną ocenę właściwości wizualnych.

Według metodologii CSAT, uzyskane wartości – 74% dla smaku, 76% dla zapachu i 79% dla przejrzystości – mieszczą się w przedziale 60–80%, co oznacza dobry i stabilny poziom satysfakcji klientów. Żaden z parametrów nie przekroczył jednak progu 80%, a najniższy wynik odnotowano w przypadku smaku.



Rys. 1. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytań 1–3

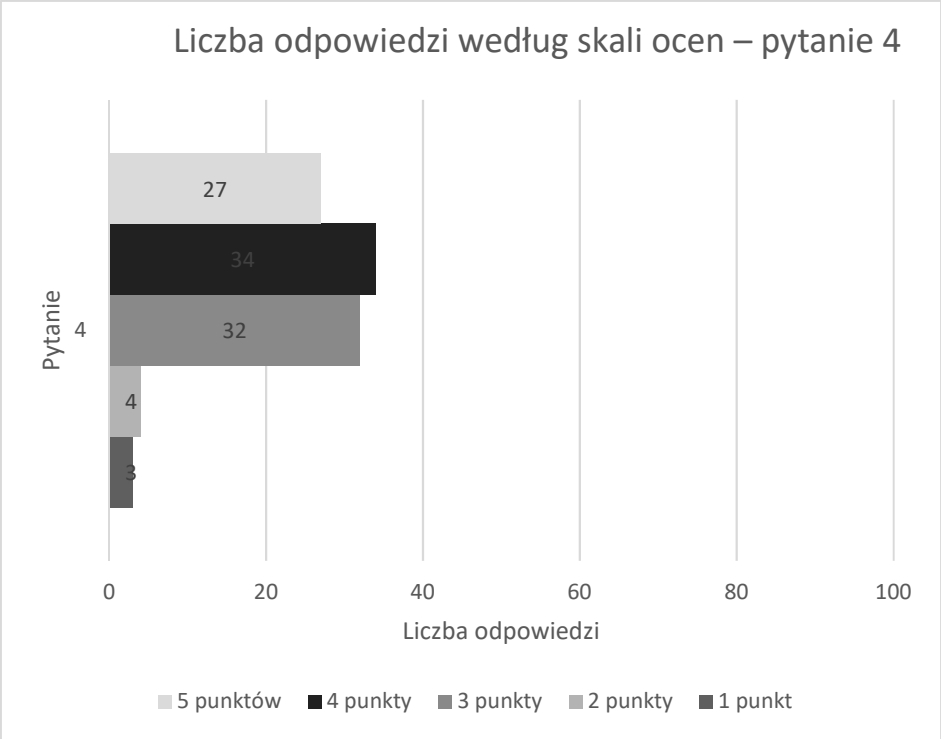
Fig. 1. Chart presenting a comparison of the number of answers in individual points of the scale for questions 1–3

Drugi aspekt dotyczył zagadnienia związanego z poczuciem bezpieczeństwa korzystania z wody wodociągowej, co stanowi jeden z kluczowych elementów zaufania odbiorców do usług przedsiębiorstwa. W badaniu uwzględniono pytanie: „W jakim stopniu uważa Pan/i wodę wodociągową za bezpieczną do spożycia?”.

Na rysunku 2. przedstawiono wyniki ankiety odnoszące się do oceny stopnia, w jakim respondenci uznają wodę wodociągową za bezpieczną do spożycia. Zestawienie to ukazuje rozkład liczby odpowiedzi w poszczególnych punktach zastosowanej skali, odzwierciedlając poziom zaufania mieszkańców do jakości i bezpieczeństwa dostarczanej wody.

Analiza rozkładu odpowiedzi wskazuje na umiarkowanie pozytywną ocenę tego aspektu. Łącznie 39 respondentów przyznało oceny niskie (1–3 punkty), natomiast 34 osoby wskazały ocenę 4 punkty, a 27 osób – 5 punktów. Oznacza to, że oceny wysokie (4 i 5 punktów) stanowiły 61% wszystkich odpowiedzi, przy jednoczesnym udziale 39% ocen niższych. Zgodnie z metodologią CSAT uzyskany wynik 61% mieści się w przedziale 60–80%, co oznacza dobry poziom satysfakcji klientów. Jednocześnie jest to wartość zbliżona do dolnej granicy tego przedziału, co wskazuje na wyraźną rezerwę w zakresie postrzeganego bezpieczeństwa wody

wodociągowej. W porównaniu z ocenami cech sensorycznych wynik ten jest niższy, co sugeruje, że mimo pozytywnych doświadczeń użytkowników związanych z walorami smaku, zapachu i przejrzystości, część respondentów nadal ma wątpliwości dotyczące bezpieczeństwa spożycia wody z sieci.



Rys. 2. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytania 4

Fig. 2. Chart presenting a comparison of the number of answers in individual points of the scale for question 4

Uzyskane wyniki wskazują na wysoki i spójny poziom ocen cech sensorycznych wody wodociągowej, mieszczący się w przedziale dobrym według metodologii CSAT, przy jednoczesnym braku osiągnięcia poziomu bardzo wysokiego. Najkorzystniej oceniono przejrzystość, nieco niżej zapach, natomiast relatywnie najniższe wartości dotyczyły smaku, co identyfikuje ten parametr jako potencjalny obszar dalszego doskonalenia jakości usług.

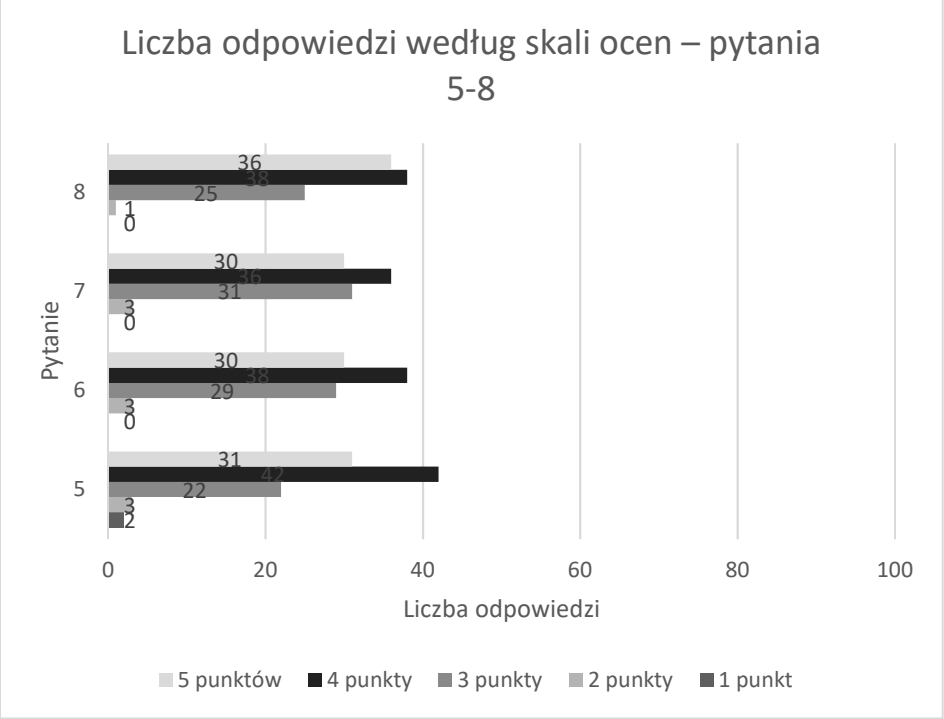
Odmienny charakter ma ocena poczucia bezpieczeństwa spożycia wody, która – mimo utrzymania się w przedziale dobrym – osiąga wartości zbliżone do dolnej granicy skali. Wskazuje to na istniejącą rezerwę w zakresie zaufania odbiorców do bezpieczeństwa wody wodociągowej i sugeruje potrzebę wzmocnienia działań informacyjnych oraz jakościowych. Zestawienie obu wymiarów pokazuje, że

pozytywne doświadczenia sensoryczne nie w pełni przekładają się na przekonanie o bezpieczeństwie wody, co stanowi istotną wskazówkę dla dalszych działań zarządczych i komunikacyjnych przedsiębiorstwa.

3.2. Niezawodność i obsługa klienta

Kolejna kategoria pytań (5–8) dotyczyła obszaru rozliczeń i komunikacji z odbiorcami, obejmującego częstość przerw w dostawie wody, sposób informowania o utrudnieniach oraz dostępność i skuteczność kontaktu z przedsiębiorstwem.

Na rysunku 3. przedstawiono wyniki ankiety odnoszące się do niezawodności dostaw wody oraz jakości obsługi klienta, w tym oceny częstości przerw w ostatnim roku, sposobu przekazywania informacji o planowanych i nieplanowanych utrudnieniach, łatwości załatwiania spraw różnymi kanałami kontaktu oraz wystarczalności dostępnych form komunikacji (telefon, e-mail, eBOK). Zestawienie prezentuje rozkład odpowiedzi w poszczególnych punktach zastosowanej skali, umożliwiając ocenę poziomu satysfakcji respondentów w analizowanych aspektach funkcjonowania przedsiębiorstwa wodociągowego.



Rys. 3. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytań 5-8

Fig. 3. Chart presenting a comparison of the number of answers in individual points of the scale for questions 5-8

W pytaniu dotyczącym liczby przerw w dostawie wody w ciągu roku (gdzie 1 oznaczało „bardzo często", a 5 – „nigdy") 27 respondentów przyznało oceny niskie (1–3 punkty), natomiast 42 osoby wskazały ocenę 4 punkty, a 31 – 5 punktów. Oznacza to, że 73% odpowiedzi miało charakter pozytywny, co przekłada się na wynik CSAT równy 73% i wskazuje na dobry poziom satysfakcji w tym zakresie. Jednocześnie obecność 27% ocen niższych sugeruje, że część odbiorców doświadcza przerw w dostawach wody częściej, niż uznawałaby za akceptowalne.

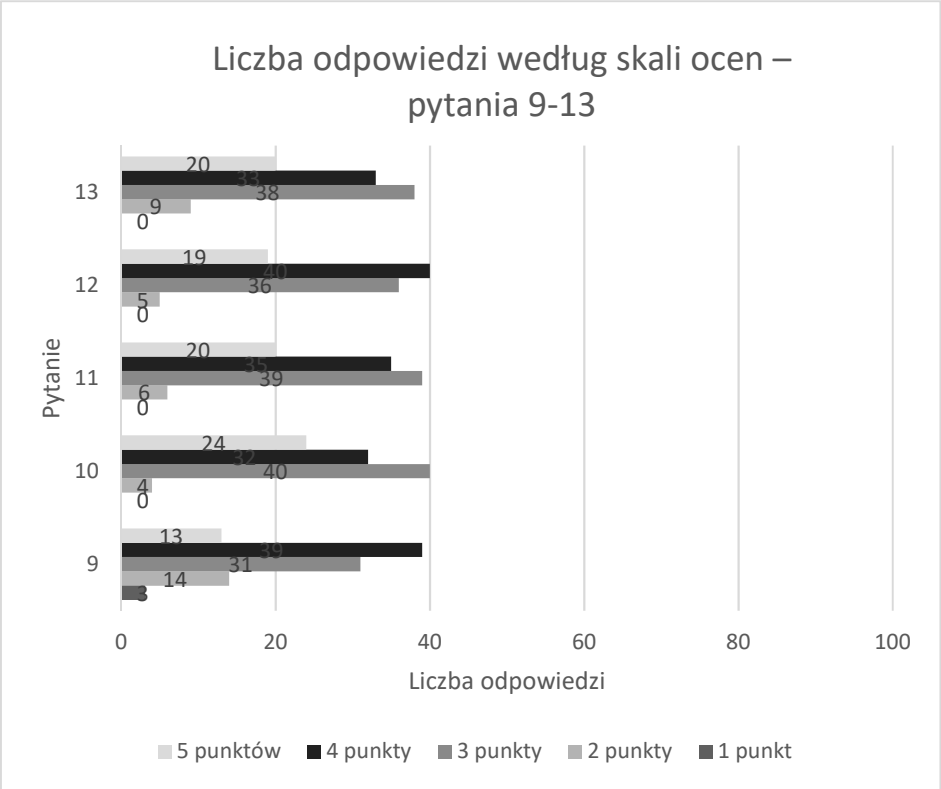
W pytaniu dotyczącym sposobu informowania o przerwach nie odnotowano żadnej oceny na poziomie 1 punktu, natomiast 32 osoby wskazały 2–3 punkty, 38 – 4 punkty, a 30 – 5 punktów (CSAT 68%). Podobną strukturę uzyskano przy ocenie łatwości załatwienia sprawy – brak ocen najniższych, 34 wskazania na 2–3 punkty, 36 na 4 punkty i 30 na 5 punktów – co przełożyło się na CSAT równy 66%. Oznacza to stabilny, lecz umiarkowanie wysoki poziom zadowolenia klientów z procesu obsługi. Najwyższy wynik w tej kategorii uzyskano w pytaniu dotyczącym dostępności kanałów kontaktu z przedsiębiorstwem. Ponownie nie pojawiły się oceny na poziomie 1 punktu, 26 respondentów przyznało 2–3 punkty, 38 – 4 punkty, a 36 – 5 punktów, co dało CSAT równy 74%. Brak skrajnie negatywnych opinii wskazuje na ogólnie dobrą dostępność form kontaktu i poprawną organizację komunikacji z odbiorcami.

Uzyskane wyniki wskazują, że wszystkie analizowane aspekty mieszczą się w przedziale 60–80% według metodologii CSAT, co odpowiada dobremu poziomowi satysfakcji klientów. Brak ocen najniższych w trzech z czterech pytań świadczy o ograniczonej skali silnie negatywnych doświadczeń odbiorców oraz relatywnie stabilnej jakości obsługi. Jednocześnie zauważalny udział ocen średnich (2–3 punkty) wskazuje na potrzebę dalszego doskonalenia procesów komunikacji z klientami, w szczególności w zakresie informowania o przerwach w dostawie wody, usprawniania procedur obsługi oraz rozwijania dostępnych kanałów kontaktu. Realizacja działań w tych obszarach może przyczynić się do wzrostu poziomu zaufania odbiorców oraz poprawy ogólnej oceny usług przedsiębiorstwa.

3.3. Rozliczenia i komunikacja

W kolejnej części badania analizowano ocenę wybranych elementów obsługi informacyjnej i rozliczeniowej przedsiębiorstwa.

Na rysunku 4. zaprezentowano wyniki ankiety odnoszące się do tych zagadnień. Obejmują one opinie respondentów dotyczące przejrzystości dokumentów, łatwości odczytu i zgłaszania zużycia, adekwatności cen względem jakości świadczonych usług oraz skuteczności i zrozumiałości przekazywanych komunikatów. Zestawienie przedstawia rozkład odpowiedzi w poszczególnych punktach skali, umożliwiając ocenę poziomu satysfakcji mieszkańców w analizowanych obszarach.



Rys. 4. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytań 9-13

Fig. 4. Chart presenting a comparison of the number of answers in individual points of the scale for questions 9-13

W pytaniu dotyczącym łatwości odczytania wskazań wodomierza i zgłaszania zużycia nie odnotowano ocen najniższych, jednak aż 40 respondentów przyznało 3 punkty, 4 osoby – 2 punkty, 32 osoby – 4 punkty, a 24 – 5 punktów. Uzyskany wynik CSAT równy 56% lokuje się w przedziale 40–60%, co oznacza umiarkowany poziom satysfakcji i wskazuje, że procedury związane z odczytem oraz raportowaniem zużycia są dla części odbiorców mało intuicyjne.

Podobną tendencję odnotowano w pytaniu o adekwatność cen zaopatrzenia w wodę i odbioru ścieków w stosunku do jakości usług. Przy braku ocen najniższych 6 respondentów wskazało 2 punkty, 39 – 3 punkty, 35 – 4 punkty, a 20 – 5 punktów, co przełożyło się na CSAT równy 55%. Wynik ten również mieści się w przedziale umiarkowanym, sugerując, że część klientów nie jest w pełni przekonana o relacji ceny do jakości świadczonych usług.

W pytaniu dotyczącym ogólnej komunikacji przedsiębiorstwa (ostrzeżenia, komunikaty, informacje o inwestycjach) nie odnotowano ocen skrajnie negatywnych; 5 respondentów przyznało 2 punkty, 39 – 3 punkty, 40 – 4 punkty, a 19 – 5 punktów, co przełożyło się na CSAT równy 59%. Wartość ta znajduje się blisko górnej granicy przedziału umiarkowanego (40–60%), wskazując na relatywnie lepszą ocenę komunikacji w porównaniu z innymi elementami tej kategorii. Odmiennie kształtowały się wyniki w pytaniu o zrozumiałość i aktualność informacji publikowanych na stronie internetowej i w mediach społecznościowych: 0 ocen najniższych, 9 wskazań na 2 punkty, 38 na 3 punkty, 33 na 4 punkty i 20 na 5 punktów, co dało CSAT równy 53%. Wynik ten pozostaje wyraźnie niższy i świadczy o większym udziale ocen średnich, co sugeruje potrzebę poprawy przejrzystości oraz aktualności publikowanych treści.

Wyniki tej kategorii wskazują, że mimo braku ocen skrajnie negatywnych, dominacja ocen średnich (3 punkty) obniża ogólny poziom satysfakcji i lokuje wszystkie analizowane aspekty w przedziale umiarkowanym według skali CSAT (40–60%). Oznacza to potrzebę usprawnienia procesów informacyjnych i rozliczeniowych, zwłaszcza w zakresie czytelności wskazań wodomierza, przejrzystości procedur zgłaszania zużycia oraz jakości komunikacji cenowej i informacyjnej. Poprawa tych elementów może istotnie wpłynąć na postrzeganie transparentności przedsiębiorstwa i zwiększyć zaufanie odbiorców do świadczonych usług.

3.4. Ocena ogólna

Kategoria obejmowała jedno pytanie odnoszące się do ogólnej oceny działalności przedsiębiorstwa wodociągowego.

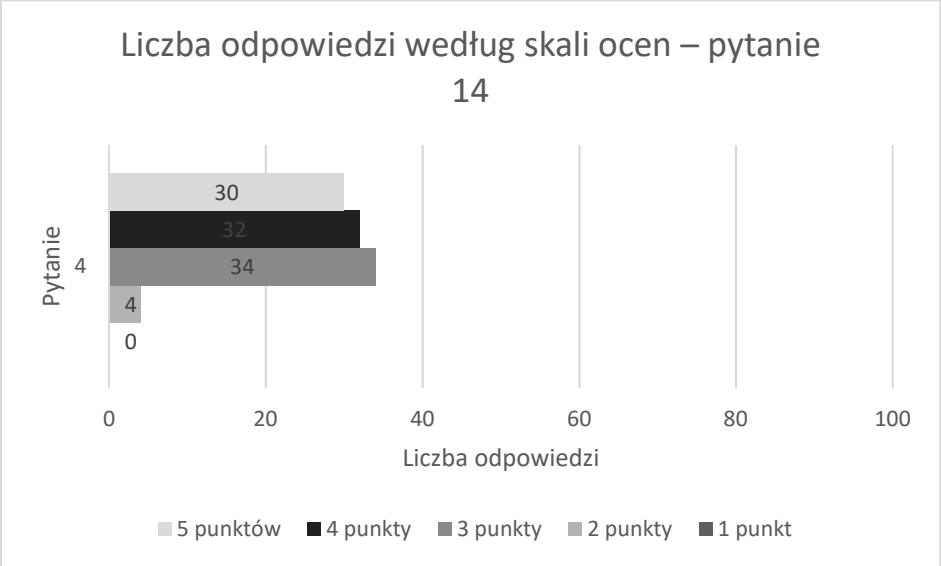
Na rysunku 5. zaprezentowano uzyskane wyniki. Zestawienie przedstawia rozkład odpowiedzi w poszczególnych punktach zastosowanej skali, odzwierciedlając ogólny poziom satysfakcji mieszkańców z jakości świadczonych usług oraz funkcjonowania przedsiębiorstwa.

Rozkład odpowiedzi wskazuje na brak ocen skrajnie negatywnych, przy jednoczesnej dominacji ocen średnich i wysokich: 4 wskazania na 2 punkty, 34 na 3 punkty, 32 na 4 punkty oraz 30 na 5 punktów. Taki rozkład świadczy o generalnie pozytywnej, choć umiarkowanej ocenie działalności przedsiębiorstwa przez respondentów.

Uzyskany wynik CSAT na poziomie 62% lokuje się – zgodnie z przyjętą w badaniu skalą interpretacyjną – w przedziale 60–80%, oznaczającym dobry poziom satysfakcji klientów. Jednocześnie wartość ta znajduje się blisko dolnej granicy tego przedziału, co wskazuje, że choć ogólna percepcja funkcjonowania przedsiębiorstwa jest korzystna, istnieje przestrzeń do dalszego doskonalenia jakości usług i komunikacji z odbiorcami.

Wynik ogólnej oceny pozostaje spójny z rezultatami uzyskanymi w poszczególnych obszarach badania – wysokimi ocenami jakości sensorycznej wody i działań edukacyjnych oraz umiarkowanymi ocenami elementów informacyjnych

i rozliczeniowych – co potwierdza, że całościowa satysfakcja klientów jest wypadkową doświadczeń w różnych aspektach kontaktu z przedsiębiorstwem.



Rys. 5. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytania 14

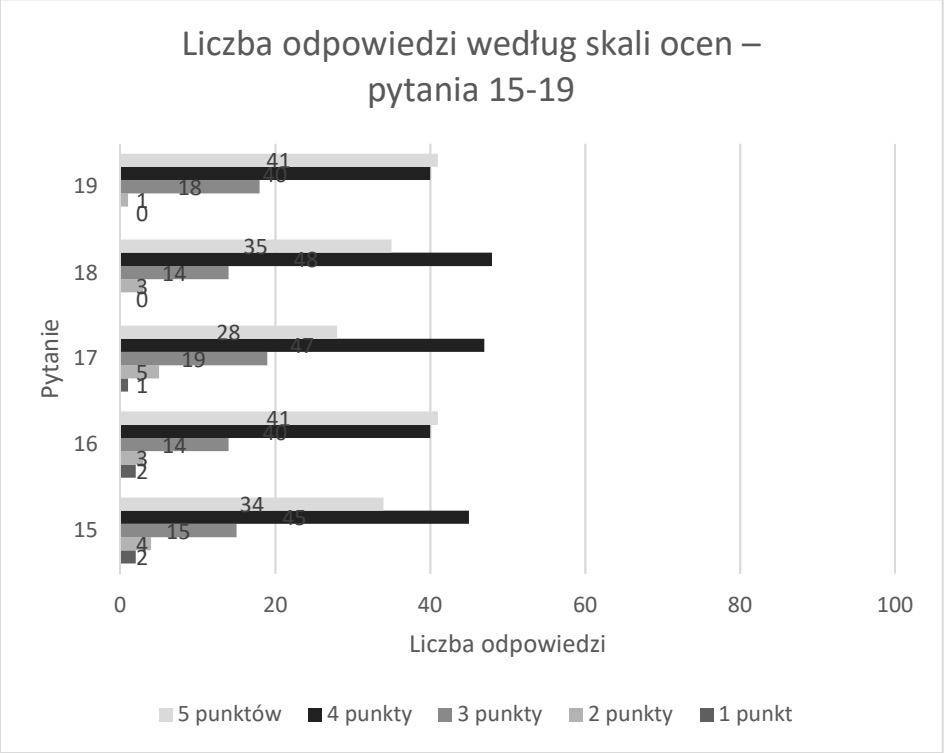
Fig. 5. Chart presenting a comparison of the number of answers in individual points of the scale for question 14

3.5. Edukacja i działania społeczne

Kolejna analizowana kategoria dotyczyła edukacji oraz działań społecznych przedsiębiorstwa wodociągowego. Uzyskane wyniki wskazują na bardzo wysoki poziom zainteresowania respondentów tematyką edukacyjną oraz wyraźne oczekiwanie aktywnej roli przedsiębiorstwa w obszarze inicjatyw prośrodowiskowych i informacyjnych.

Na rysunku 6. zaprezentowano szczegółowy rozkład odpowiedzi odnoszących się do tych zagadnień. Obejmują one opinie dotyczące potrzeby udostępniania materiałów edukacyjnych, organizowania kampanii proekologicznych, wyjaśniania zasad funkcjonowania systemu wodociągowo-kanalizacyjnego oraz podejmowania inicjatyw o charakterze społecznym i otwartym dla mieszkańców.

Zestawienie przedstawia rozkład odpowiedzi w poszczególnych punktach skali, umożliwiając ocenę stopnia zainteresowania odbiorców usług proponowanymi formami aktywności.



Rys. 6. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytań 15-19

Fig. 6. Chart presenting a comparison of the number of answers in individual points of the scale for questions 15-19

W pytaniu dotyczącym zainteresowania materiałami na temat oszczędzania wody i jej jakości dominowały oceny wysokie: 45 wskazań na 4 punkty oraz 34 na 5 punktów, przy marginalnym udziale ocen niskich. Przełożyło się to na wynik CSAT na poziomie 79%, co oznacza wysoki poziom satysfakcji i bardzo pozytywną postawę wobec tego rodzaju inicjatyw. Jeszcze wyższy rezultat odnotowano w pytaniu o zasadność prowadzenia kampanii edukacyjnych dotyczących oszczędzania wody i ochrony środowiska – 40 ocen na poziomie 4 punktów oraz 41 na poziomie 5 punktów dało CSAT równy 81%, co kwalifikuje ten obszar jako bardzo wysoki i potwierdza silne poparcie społeczne dla tego typu działań.

W przypadku zainteresowania materiałami wyjaśniającymi funkcjonowanie oczyszczalni ścieków i systemu dostarczania wody również przeważały oceny wysokie (47 wskazań na 4 punkty i 28 na 5 punktów), co przełożyło się na CSAT na poziomie 75%, a więc wciąż wyraźnie dobrym. Jeszcze wyraźniejsze poparcie odnotowano w pytaniu o potrzebę prowadzenia szerszych działań informacyjnych i społecznych (np. dni otwarte, warsztaty, spotkania z mieszkańcami) – brak ocen

najniższych oraz 83% CSAT wskazują na bardzo wysoki poziom akceptacji i oczekiwań w tym zakresie. Równie istotny jest wynik dotyczący zainteresowania respondentów udziałem w wizycie edukacyjnej lub dniu otwartym w przedsiębiorstwie wodociągowym – 80% CSAT, przy dominacji ocen 4 i 5 punktów, wskazuje na wysoki poziom gotowości odbiorców do bezpośredniego poznania działalności przedsiębiorstwa. Rezultat ten potwierdza, że działania o charakterze edukacyjnym i otwartym cieszą się dużym zainteresowaniem i mogą stanowić skuteczne narzędzie budowania świadomości społecznej, wzmacniania zaufania klientów oraz kształtowania bardziej partnerskich relacji między przedsiębiorstwem a odbiorcami usług.

Zestawienie wyników (79%, 81%, 75%, 83%, 80%) wskazuje na wyjątkowo wysoki poziom aprobaty społecznej dla działań edukacyjnych i prośrodowiskowych. W przeciwieństwie do wcześniejszych kategorii, większość analizowanych wskaźników osiąga poziom bardzo wysoki lub znajduje się na granicy tego przedziału według metodologii CSAT. Oznacza to, że społeczeństwo wykazuje realną gotowość do pogłębiania wiedzy oraz aktywnego uczestnictwa w inicjatywach edukacyjnych. Stanowi to istotny potencjał rozwojowy i jednoznaczną przesłankę do dalszego intensyfikowania działań w obszarze edukacji ekologicznej, transparentności funkcjonowania systemu wodociągowego oraz budowania partnerskich relacji z mieszkańcami.

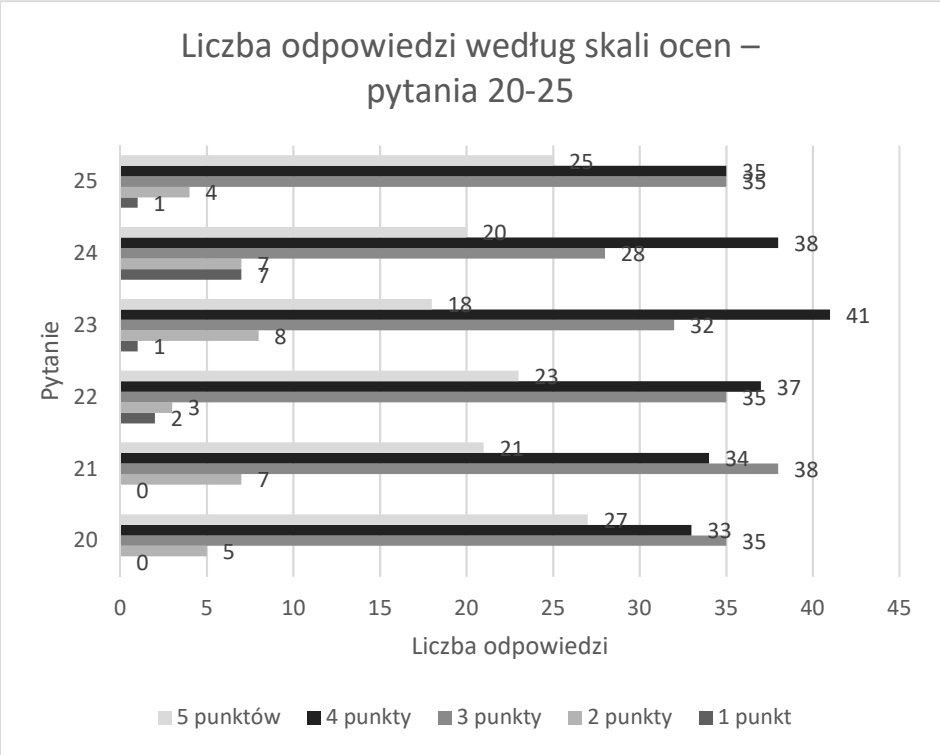
3.6. Ekologia, zaufanie i postawy

Ostatni segment ankiety dotyczył zagadnień związanych z ekologią, zaufaniem oraz postawami odbiorców wobec usług wodociągowych.

Na rysunku 7. zaprezentowano wyniki odnoszące się do postaw mieszkańców wobec kwestii środowiskowych oraz ich zaufania do jakości dostarczanej wody. Zestawienie obejmuje opinie dotyczące znaczenia proekologicznego funkcjonowania przedsiębiorstwa, oceny sposobu informowania o działaniach w zakresie ochrony środowiska i oszczędzania zasobów, a także deklarowanych praktyk wodooszczędnych w gospodarstwach domowych.

Uwzględniono również poziom zaufania do wody z kranu, komfort jej spożywania oraz wpływ informacji o modernizacjach i badaniach jakości na poczucie bezpieczeństwa. Rozkład odpowiedzi w poszczególnych punktach skali pozwala na ocenę świadomości ekologicznej respondentów oraz ich nastawienia do jakości i bezpieczeństwa świadczonych usług.

W pytaniu odnoszącym się do znaczenia prowadzenia działalności przedsiębiorstwa w sposób przyjazny dla środowiska dominowały oceny wysokie (40 wskazań na 4 punkty i 41 na 5 punktów), przy niemal całkowitym braku ocen niskich. Mimo bardzo korzystnego rozkładu odpowiedzi uzyskany wskaźnik CSAT na poziomie 60% lokuje się na dolnej granicy przedziału dobrego (60–80%), co sugeruje pozytywne nastawienie respondentów, lecz bez jednoznacznie bardzo wysokiej oceny.



Rys. 7. Wykres przedstawiający porównanie liczby odpowiedzi w poszczególnych punktach skali dla pytań 20-25

Fig. 7. Chart presenting a comparison of the number of answers in individual points of the scale for questions 20-25

Podobny rezultat odnotowano w pytaniu dotyczącym oceny informowania przez przedsiębiorstwo o działaniach w zakresie ochrony środowiska i oszczędzania wody (CSAT = 60%). Choć dominowały oceny średnie i wysokie, wskazuje to na potrzebę dalszego rozwijania komunikacji ekologicznej i zwiększania widoczności podejmowanych inicjatyw. W pytaniu o indywidualne działania wodooszczędne w gospodarstwach domowych uzyskano niższy wynik – CSAT równy 55% – co wskazuje, że postawy proekologiczne respondentów są umiarkowane i mogą wymagać dalszego wzmacniania poprzez edukację i działania informacyjne.

W obszarze zaufania wobec jakości wody z kranu uzyskano CSAT równy 60%, przy dominacji ocen 3–4 punkty, co potwierdza umiarkowanie pozytywną ocenę, lecz wskazuje na istniejące wątpliwości części odbiorców. Podobnie kształtowały się odpowiedzi dotyczące komfortu picia wody z kranu (CSAT = 58%), gdzie większy udział ocen średnich i niskich sugeruje barierę psychologiczną mimo ogólnie dobrej oceny jakości wody. Na zbliżonym poziomie (CSAT = 60%) oceniono wpływ informacji o modernizacjach, badaniach jakości i działaniach

przedsiębiorstwa na poczucie bezpieczeństwa respondentów, co potwierdza znaczenie transparentnej komunikacji w budowaniu zaufania społecznego.

Łącznie wyniki tego segmentu wskazują na umiarkowany poziom satysfakcji w obszarze ekologii i zaufania, mieszczący się głównie w dolnej części przedziału dobrego według przyjętej skali CSAT. Oznacza to, że respondenci dostrzegają znaczenie działań prośrodowiskowych oraz rolę przedsiębiorstwa w ich realizacji, jednak poziom zaufania i komfortu korzystania z wody z kranu pozostaje niższy niż w przypadku ocen jakości sensorycznej czy działań edukacyjnych. Rezultaty te wskazują na potrzebę dalszego wzmacniania komunikacji ekologicznej, działań informacyjnych oraz inicjatyw budujących świadomość i zaufanie odbiorców.

4. Podsumowanie i wnioski

Przeprowadzona analiza wyników badania ankietowego wskazuje, że ogólny poziom zadowolenia odbiorców usług wodociągowych można uznać za stabilny i relatywnie wysoki. W większości analizowanych badanych obszarów wartości wskaźnika CSAT mieściły się w przedziale 60–80%, co odpowiada dobremu poziomowi satysfakcji klientów. Najlepiej oceniono cechy sensoryczne wody (smak, zapach oraz klarowność) oraz podejmowanie przez przedsiębiorstwo inicjatywy o charakterze edukacyjnym. Potwierdza to fakt, że codzienne doświadczenia użytkowników związane z jakością wody są w przeważającej mierze pozytywne, a działania informacyjne przedsiębiorstwa spotykają się z aprobatą odbiorców. Wyniki te wskazują, że jakość dostarczanego produktu oraz otwartość na dialog społeczny stanowią ważny fundament budowania zaufania klientów przedsiębiorstwa wodociągów.

Jednocześnie uzyskane rezultaty ujawniają obszary wymagające dalszej poprawy. Niższe oceny pojawiły się w odniesieniu do postrzeganego bezpieczeństwa wody, przejrzystości procesów informacyjnych oraz niektórych elementów rozliczeń i obsługi klienta. Stosunkowo duży udział ocen średnich oraz wyniki CSAT mieszczące się w przedziale umiarkowanym (40–60%) wskazują na potrzebę uproszczenia procedur związanych z odczytem wodomierza i zgłaszaniem zużycia, poprawy komunikacji dotyczącej cen usług oraz dalszego usprawnienia systemu informowania o przerwach w dostawie wody. Sugeruje to, że pozytywne odczucia związane z właściwościami sensorycznymi wody nie zawsze w pełni przekładają się na poczucie bezpieczeństwa i wysoką ocenę komunikacji przedsiębiorstwa.

Zestawienie wyników w poszczególnych kategoriach pokazuje, że ogólna ocena działalności przedsiębiorstwa jest efektem doświadczeń klientów w wielu wymiarach – od jakości wody, poprzez obsługę i kontakt z klientem, aż po transparentność działań organizacyjnych. Otrzymany obraz wskazuje na dobrą, choć zróżnicowaną ocenę funkcjonowania przedsiębiorstwa oraz na istotny potencjał dalszego rozwoju, zwłaszcza w obszarze komunikacji, edukacji i budowania zaufania społecznego. Wdrażanie działań usprawniających w tych zakresach może

przyczynić się do wzrostu satysfakcji odbiorców, lepszego postrzegania jakości usług oraz umacniania długofalowych relacji między przedsiębiorstwem wodociągowym a społecznością lokalną.

Literatura

[1] Rak J., Tchórzewska-Cieślak B., Studziński A., Pietrucha-Urbanik K., Boryczko K., Niezawodność i bezpieczeństwo systemów zbiorowego zaopatrzenia w wodę, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2012.

[2] Rak J., Zasady badań marketingowych związanych z zaopatrzeniem w wodę, Wodo ciągi – Kanalizacja, 5(14)2005, Wydawnictwo Abrys, Poznań, s. 13-16.

[3] Churchill G. A., Badania marketingowe. Podstawy metodologiczne, Wydawnictwo Naukowe PWN, Warszawa 2002.

[4] https://wsjp.pl/haslo/podglad/21774/sondaz/4599417/zbieranie-opinii (dostęp: 15.02.2026).

[5] Rak J., Pietrucha-Urbanik K., Badanie sondażowe związane z brakiem dostawy wody wodociągowej w sytuacjach kryzysowych, Instal 2/2016, Warszawa, s. 54-58.

[6] Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. Journal of Marketing.

[7] Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. Journal of Marketing Research.

[8] Malhotra, N. K. (2010). Marketing Research: An Applied Orientation. Pearson Education.

[9] Hill, N., Brierley, J., MacDougall, R. (2003). How to Measure Customer Satisfaction. Gower Publishing, Ltd.

[10] https://www.marcomrobot.com/blog/what-is-a-good-csat-score (dostęp: 16.02.2026).

CUSTOMER SATISFACTION ASSESSMENT OF A WATER UTILITY COMPANY INCLUDING SOCIAL, TECHNICAL, AND ECONOMIC ASPECTS

Summary

User opinion surveys constitute an essential tool for enhancing the performance of water supply systems and strengthening their security. The objective of this study was to evaluate the quality and timeliness of information provided by a water utility company, the level of residents' interest in water conservation and water quality issues, the necessity for educational campaigns, and the readiness of consumers to participate in social initiatives. A survey was conducted among 100 randomly selected residents utilising the services of a local water supplier. The questionnaire comprised 25 questions regarding satisfaction with water services, including assessments of water quality, communication with consumers, educational activities, and social initiatives. The Customer Satisfaction Score (CSAT) was employed to evaluate satisfaction levels. The results indicate a stable and relatively high level of consumer satisfaction, with CSAT values for most analysed areas ranging between 60% and 80%. The findings underscore the importance of continuing information and education efforts to maintain a high standard of water service quality.

Keywords: survey research, water supply system, quality assessment of water supply services.

Norbert Mokrzycki, Filip Naróg, Patryk Przybyło, Adrian Rak, Kacper Saletnik, Krystian Sołtys, Andrzej Sowiński, Colin Sudół, Eryk Twardowski, Katarzyna Tylutka
Koło Naukowe Interakcji Człowiek - Komputer GEST

dr hab. inż. Tomasz Kapuściński, prof. PRz, dr inż. Dawid Kalandyk
opiekunowie Koła Naukowego

ROBOTYCZNE RAMIĘ SORTUJĄCE

Streszczenie

Celem projektu opisanego w rozdziale jest zaprojektowanie i wykonanie robota służącego do nauki sztucznej inteligencji poprzez praktyczne programowanie ramienia robotycznego. Ramię wyposażone w chwytak, przenosi obiekty o różnych kształtach, takich jak stożki, sześciany czy walce. Części robota są wykonane w technologii druku 3D, co zapewnia dużą elastyczność projektową. Robot wykorzystuje również standardowe komponenty mechaniczne stosowane w tego typu konstrukcjach, takie jak serwomechanizmy, silniki krokowe, silniki DC, siłowniki linowe oraz popularny mikrokomputer Raspberry Pi. Dla robota zostanie stworzone własne API na Raspberry Pi, ograniczając maksymalne ruchy robota, aby uniemożliwić użytkownikowi przypadkowe uszkodzenie urządzenia poprzez nadmierny obrót którejkolwiek z części. Robot będzie korzystał ze sztucznej inteligencji, aby lokalizować obiekty na planszy (również wydrukowanej przy użyciu technologii 3D). Na podstawie obliczonej pozycji z kamer, zamontowanych na robocie, będzie on przenosił obiekt z planszy do odpowiedniego pojemnika. Robot obliczy trasę, po której będzie przemieszczał chwytak, aby złapać obiekt, następnie wraz ze złapanym obiektem, przesunie go nad wyznaczone pudełko i upuści go. Zarówno plansza, jak i pojemniki będą elementami stacjonarnymi. Dzięki robotowi studenci będą mogli w przystępny i praktyczny sposób zdobyć doświadczenie w zakresie obsługi systemów sztucznej inteligencji. Projekt łączy zagadnienia z zakresu mechaniki, elektroniki, programowania, projektowania systemów wbudowanych oraz przetwarzania obrazu, stanowiąc interdyscyplinarne środowisko dydaktyczne.

Słowa kluczowe: Sztuczna inteligencja, Raspberry Pi, modelowanie CAD, kinematyka odwrotna.

1. Wprowadzenie

W ostatnich latach dynamiczny rozwój sztucznej inteligencji sprawił, że przestała ona być tylko tematem badań naukowych i futurystycznych wizji. Stała się

integralną częścią współczesnej edukacji, przemysłu i życia codziennego. W raporcie OECD „AI adoption in the education system" podkreślono, że rośnie zainteresowanie integracją sztucznej inteligencji (AI, ang. Artifical Intelligence) w procesach nauczania i uczenia się oraz w kształtowaniu kompetencji przyszłości, co skłania instytucje edukacyjne do modyfikowania programów nauczania i wdrażania narzędzi AI wspierających praktyki dydaktyczne [1]. Dodatkowo aż 92% uczniów na świecie korzysta z narzędzi opartych na sztucznej inteligencji [2].

W świetle tych przesłanek projekt opisany w niniejszym rozdziale stanowi odpowiedź na rosnące zapotrzebowanie na praktyczne narzędzia edukacyjne. Celem jest opracowanie robota edukacyjnego, który stanie się praktyczną platformą do nauki programowania sztucznej inteligencji. Robot umożliwi studentom bezpośrednią interakcję z fizycznym ramieniem, które sortuje różne kształty przy użyciu chwytaka. Dzięki temu możliwe będzie zgłębienie teoretycznych aspektów algorytmów AI, ale przede wszystkim przełożenie ich na rzeczywiste działanie mechaniczne. Poza tym robot integruje elementy sztucznej inteligencji, które pozwalają mu rozpoznawać i lokalizować obiekty na planszy. Taka konstrukcja stwarza wyjątkową przestrzeń edukacyjną, łącząc mechanikę, elektronikę i programowanie AI w jednym, praktycznym narzędziu naukowym.

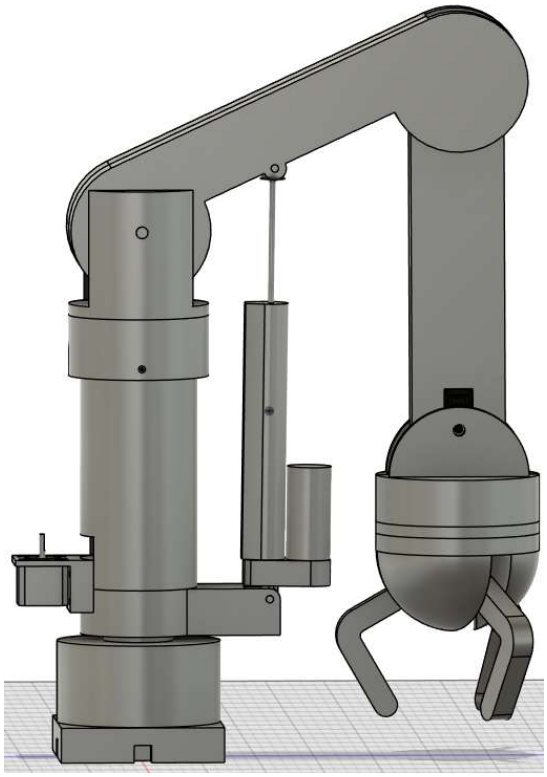
2. Konstrukcja ramienia robotycznego

Do stworzenia robota wykorzystano serwomechanizmy, siłownik liniowy, silnik krokowy, silnik prądu stałego (DC) wraz z czujnikiem prądu INA226 oraz mostkiem H. W projekcie zastosowano Raspberry Pi, który umożliwia uruchamianie bardziej złożonych algorytmów niż typowe mikrokontrolery. Raspberry Pi wyposażono w porty USB, port Ethernet oraz moduł komunikacji Wi-Fi, co umożliwia zdalną kontrolę poprzez sieć. Mikrokomputer posiada również wyprowadzenia GPIO, które pozwalają na integrację z elementami wykonawczymi i czujnikami, co czyni go odpowiednim rozwiązaniem do budowy systemów robotycznych.

Jeden serwomechanizm w podstawowej konstrukcji robota służy do poruszania drugim ramieniem, natomiast dwa kolejne realizują obrót chwytaka. Pierwsze ramię jest podnoszone poprzez siłownik liniowy, co zwiększa jego zdolność do przenoszenia obciążeń, niezbędną do unoszenia chwytaka oraz pozostałych elementów ramienia. W celu zapewnienia stabilności konstrukcji kolumnę robota zaprojektowano jako walec osadzony na dwóch łożyskach. Obrót kolumnowy jest realizowany przez silnik krokowy. W celu ograniczenia zakresu ruchu zastosowano dwa styczniki krańcowe, umieszczone w dolnej części konstrukcji, pod obracającą się kolumną. Napęd przenoszony jest z silnika na kolumnę za pomocą przekładni pasowej. Silnik jest mechanicznie połączony z elementem obrotowym, co umożliwia swobodny i kontrolowany obrót całej kolumny w zadanym zakresie. Dzięki zastosowaniu serwomechanizmów o momencie obrotowym 30 kg·cm [4] możliwe jest przenoszenie obciążeń wynikających z masy elementów połączonych

z ramieniem. Chwytak składa się z silnika krokowego oraz przekładni ślimakowej, która poprzez koła zębate o prostych zębach napędza poszczególne elementy chwytaka. Stan chwytaka (otwarty/zamknięty) będzie określany na podstawie wartości prądu pobieranego przez silnik prądu stałego.

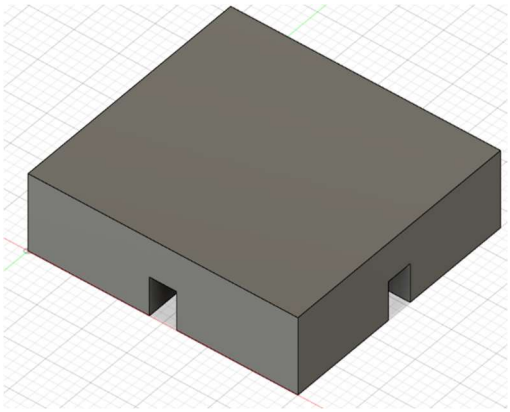
Możliwość obrotu chwytaka w dwóch osiach zwiększa elastyczność manipulacyjną systemu, umożliwiając przekazywanie obiektów pomiędzy dwoma robotami oraz optymalne dopasowanie orientacji chwytaka do podnoszonego elementu. Modularna budowa będzie umożliwiała łatwą wymianę chwytaka. Zastosowanie pustych wewnętrznie elementów ramienia pozwala na zmniejszenie jego masy oraz ułatwia prowadzenie przewodów do podzespołów elektronicznych. Kolumna wewnętrzna, na której osadzone są łożyska, również ma pustą strukturę, co redukuje naprężenia skrętne podczas obrotu ramienia i tym samym zmniejsza ryzyko ich uszkodzenia. Na rysunku (rys. 1.) przedstawiono pierwszy prototyp ramienia. W programie Fusion 360, możliwe jest wstawianie gotowych połączeń między elementami co pozwala to na analizę ruchu ramienia przed jedo wydrukiem i montażem [5].



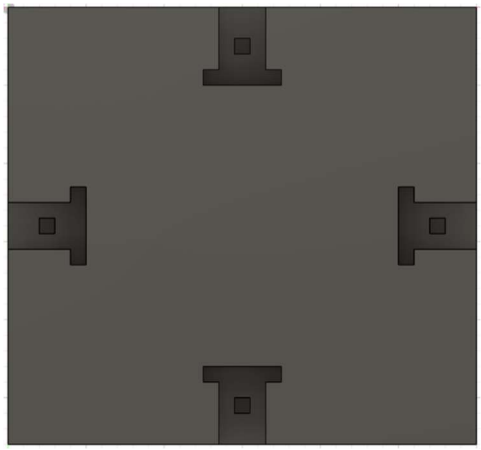
Rys. 1. Prototyp konstrukcji ramienia robotycznego wykonany w Fusion 360
Fig. 1. A prototype of the robotic arm design made in Fusion 360
Źródło: Opracowanie własne.

3. Konstrukcja platformy robota

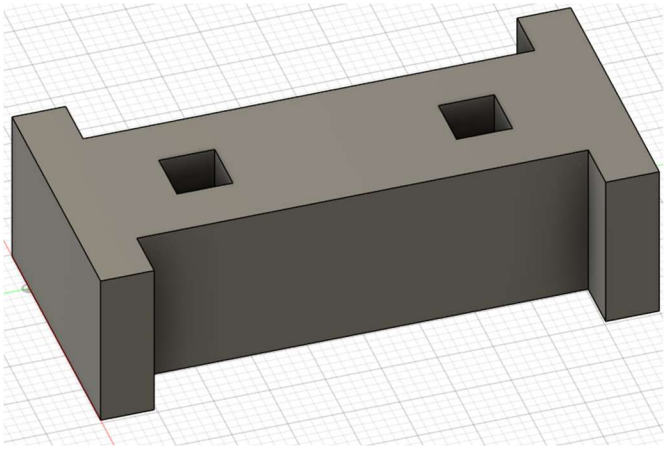
Główna platforma zrobiona jest z bloków (rys. 2. i 3.), które są ze sobą połączone za pomocą łączników (rys. 4.). Dla zapewnienia silnego połączenia oprócz samego kształtu, który całkowicie blokuje ruch w poziomie, zastosowane są magnesy neodymowe 5x5x5 mm. Elementy te zostały drukowane przy użyciu technologii 3D [6].



Rys. 2. Pojedynczy blok platformy (widok z boku)
Fig. 2. Single platform block (side view)
Źródło: Opracowanie własne.

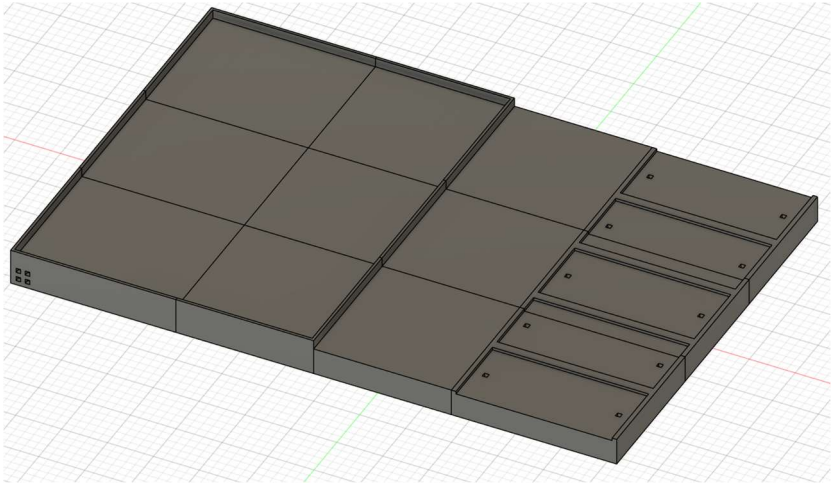


Rys. 3. Pojedynczy blok platformy formy (widok z dołu)
Fig. 3. Single platform block (down view)
Źródło: Opracowanie własne.

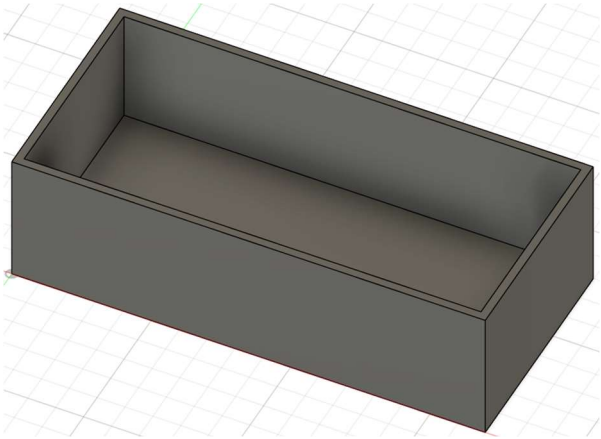


Rys. 4. Pojedynczy łącznik (widok z boku)
Fig. 4. Single connector (side view)
Źródło: Opracowanie własne.

Bloki po lewej stronie, widziane od przodu (rys. 5.) mają dodatkowe ściany zapobiegające spadaniu sortowanych elementów. W rogach znajdują się otwory na magnesy neodymowe, które umożliwiają montaż uchwytów na kamery. Bloki z prawej strony mają gniazda montażowe na pojemniki, do których odkładane są posortowane elementy. Pojemniki (rys. 6.) mają wymiary 8x17x5 cm i są montowane za pomocą magnesów neodymowych. Cała platforma po połączeniu wszystkich części ma wymiary 44x66 cm.



Rys. 5. Cała platforma
Fig. 5. The entire platform
Źródło: Opracowanie własne.



Rys. 6. Pojemnik (widok z boku)
Fig. 6. Container (side view)
Źródło: Opracowanie własne.

Taka konstrukcja znacznie ułatwia wydrukowanie poszczególnych części oraz przenoszenie całej platformy, dzięki możliwości jej szybkiego demontażu i ponownego montażu. Zapewnia to łatwą wymianę części w przypadku ich uszkodzenia lub modyfikacji związanych z dalszym rozwojem projektu.

4. API robota

Kluczowym elementem systemu będzie własne API zaimplementowane na Raspberry Pi, które będzie pełnić rolę warstwy abstrakcji sprzętowej, umożliwiając sterowanie robotem bez konieczności bezpośredniego operowania sygnałami PWM czy sterownikami silników. Głównym celem jest zapewnienie bezpieczeństwa mechanicznego konstrukcji, uproszczenie interfejsu dla użytkownika uczącego się programowania oraz separacji logiki wysokiego poziomu od sterowania niskopoziomowego.

API będzie udostępniać funkcję do poruszanie każdą częścią osobno o zadany kąt, pod warunkiem, że nie przekroczy on ustalonego limitu z wyjątkiem elementu napędzanego przez siłownik liniowy, dla której przesunięcie określane będzie w milimetrach. Można będzie też użyć funkcji, która przemieści całe ramię na podstawie podanych współrzędnych, o ile ramię może zostać ustawione w tej pozycji bez przekraczania swoich limitów. System zostanie wyposażony w funkcję do zamykania i otwierania chwytaka oraz funkcję ustawiania ramienia w pozycji domyślniej co gwarantuje, że po jej wywołaniu całe ramię zostanie ustawione w ściśle określonej pozycji, definiując układ odniesienia dla dalszych operacji.

Funkcje ruchu będą realizować na podstawie równań kinematycznym robota. Dla znanych współrzędnych x, y, z pozycji końcowej chwytaka obliczane będą

kąty obrotu poszczególnych części. W pierwszym kroku wykonywany jest rzut na płaszczyznę XY w celu wyznaczenia pierwszego kąta.

$$r = \sqrt{(x^2 + y^2)} \tag{1}$$

$$\theta_1 = \text{atan2}(\text{y}, \text{x}) \tag{2}$$

gdzie: r – odległość rzutu punktu końcowego chwytaka w układzie bazowym robota,
x, y – znane współrzędne chwytaka,
θ_1 – kąt obrotu robota u podstawy w poziomie.

Reszta kątów będzie liczona w płaszczyźnie pionowej (r, Z)

$$z_1 = z - h \tag{3}$$

gdzie: z_1 – wysokość względem osi pierwszego przegubu,
h – wysokość osi pierwszego przegubu nad podstawą,
z – znany koordynat chwytaka.

Z twierdzenia cosinusów

$$D = \frac{r^2 + z_1^2 - L_1^2 - L_2^2}{2L_1L_2} \tag{4}$$

gdzie: D – pomocnicza zmienna równa $\cos(\theta_3)$,
r – odległość rzutu punktu końcowego chwytaka w układzie bazowym,
L_1, L_2 – długość pierwszego i drugiego ramienia.

Jeśli

$$-1 \leq D \leq 1 \tag{5}$$

To istnieje rozwiązanie

$$\theta_3 = atan2(\pm\sqrt{1 - D^2}, D) \tag{6}$$

gdzie: θ_3 - kąt obrotu robota w drugim ramieniu.

Istnieją dwa rozwiązania, odpowiadające dwóm różnym konfiguracjom, co uzasadnia obecność symbolu ± w równaniu (6). Wybór zależy od ograniczeń mechanicznych oraz aktualnego położenia ramienia.

$$\theta_2 = atan2(z_1, r) - atan2(L_2 sin\theta_3,\ L_1 + L_2 cos\theta_3) \tag{7}$$

gdzie: θ_2 – kąt obrotu pierwszego ramienia.

Szczegółowe wyprowadzenie wzoru (7) można znaleźć w [7].

$$\theta_4 = \phi - (\theta_2 + \theta_3) \tag{8}$$

gdzie: θ_4 – kąt obrotu chwytaka,
ϕ – żądana orientacja chwytaka, dla pozycji pionowej równe 0.

Są to wszystkie kąty wpływające na ruch ramienia, jednak za ruch pierwszego ramienia odpowiada siłownik liniowy, dlatego konieczne jest wyprowadzenie zależności łączącej długość siłownika z odpowiadającym jej katem obrotu. Z twierdzenia cosinusów:

$$l = \sqrt{(a^2 + b^2 - 2ab\cos(\theta_2 + \alpha))} \tag{9}$$

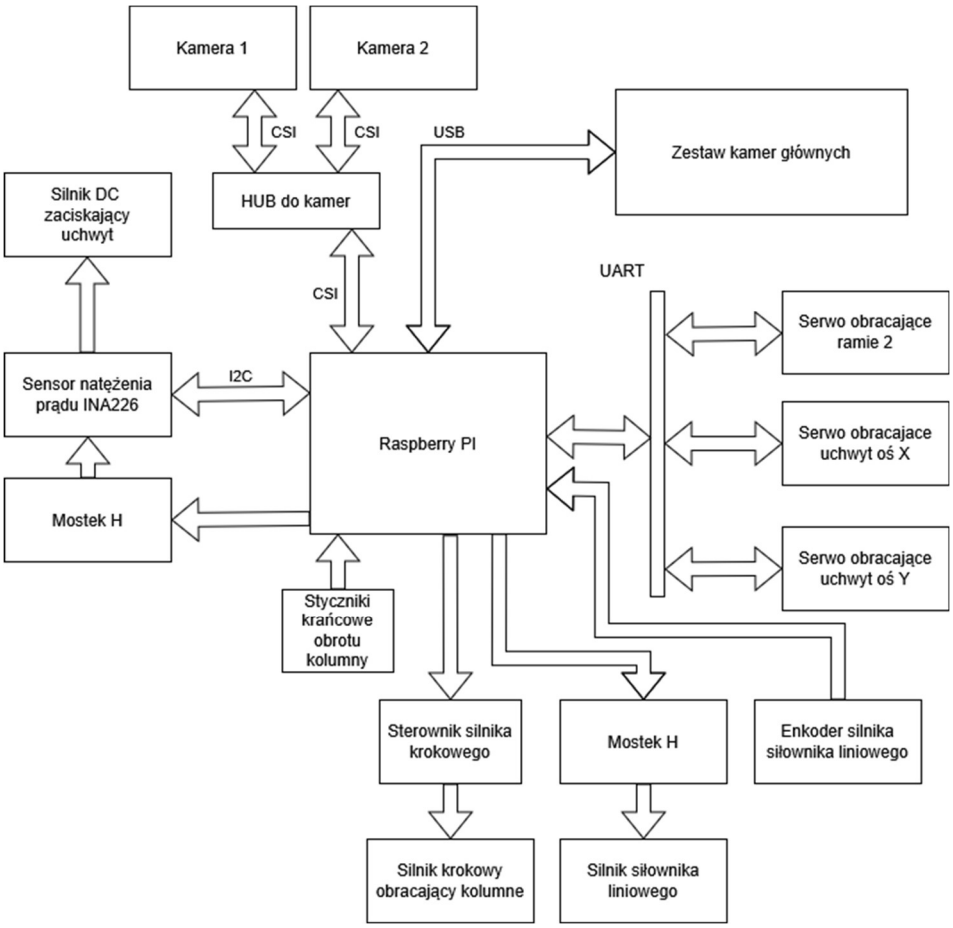
gdzie: l – długość siłownika,
a – odległość punktu mocowania podstawy,
b – odległość punktu mocowania na ramieniu,
α – kąt nachylenia siłownika.

Powyższe równania umożliwiają sterownie ramieniem na podstawie współrzędnych x, y, z określonych w układzie odniesienia podstawy ramienia.

5. Architektura systemu

Architektura została zaprojektowana w sposób zapewniający bezpośrednią komunikację mikrokomputera z elementami wykonawczymi oraz czujnikami. Schemat połączeń elektrycznych przedstawiono na rys. 7. System jest zarządzany przez mikrokomputer Raspberry Pi. Każdy komponent w systemie jest z nim połączony bez pośrednich mikrokontrolerów. Dzięki zastosowaniu trzech kamer technologia AI będzie mogła efektywnie lokalizować przedmioty w przestrzeni roboczej robota. Mikrokomputer jest połączony z serwomechanizmami poprzez porty szeregowe UART, wykorzystując złącza umożliwiające połączenie szeregowe. Silnik krokowy jest sterowany za pośrednictwem dedykowanego sterownik, kontrolowanego za pomocą sygnałów STEP i DIR. Aby wyznaczyć pozycję referencyjną obrotu kolumny oraz zapobiec obracaniu 360°, w konstrukcji kolumny zastosowane są dwa styczniki krańcowe, połączone bezpośrednio przez GPIO do Raspberry Pi. Zacisk robota jest napędzany silnikiem DC, połączonym z mostkiem H poprzez sensor INA226 [8]. Zadaniem czujnika prądu jest monitorowanie poboru, co pozwala określić, czy uchwyt jest otwarty czy zamknięty, nawet jeśli w nim znajduje się dowolny obiekt do podniesienia przez ramię. Mostek H jest połączony bezpośrednio z Raspberry Pi, kontrolowany przez sygnał PWM oraz sygnały FORWARDS i BACKWARDS. Siłownik liniowy podobnie jak silnik prądu stałego w chwytaku jest połączony z mikrokomputerem przez mostek H oraz

posiada enkoder obrotowy, za pomocą którego mikrokomputer jest w stanie precyzyjnie określić aktualną długość ramienia. Aby skalibrować silnik liniowy, wykorzystane zostaną wewnętrzne styczniki krańcowe, które odłączą go od zasilania po osiągnięciu skrajnych pozycji. Informację tę można odczytać z enkodera obrotowego.



Rys. 7. Schemat połączeń robota
Fig. 7. Robot connection diagram
Źródło: Opracowanie własne.

6. Zastosowanie w nauce

Dzięki możliwościom ramienia i modularnej budowie robota będzie można zastosować przy nauce podstaw mechatroniki oraz sztucznej inteligencji. Zastosowanie kilku kamer umożliwia implementację algorytmów sztucznej inteligencji

odpowiedzialnych za detekcję oraz klasyfikację obiektów znajdujących się w przestrzeni roboczej. Odpowiednio zaprogramowany system może identyfikować przedmioty, a następnie przenosić je do wyznaczonych pojemników. Możliwe jest również realizowanie scenariuszy sortowania, w których w danym pojemniku znajduje się określona liczba obiektów wybranego typu, koloru lub kształtu. Dzięki modularnej budowie, będzie można w prosty sposób budować nowe chwytaki oraz zamieniać elementy w ramieniu. Dzięki ograniczeniom wbudowanym we własne API, robot nie będzie w stanie przekroczyć określonych limitów, co zapobiegnie uszkodzeniom części podczas nauki.

Ponadto platforma może być wykorzystywana jako środowisko do realizacji projektów zespołowych oraz badań eksperymentalnych z zakresu sterowania i optymalizacji procesów manipulacyjnych. Studenci mogą analizować dokładność pozycjonowania, czas realizacji zadań czy skuteczność algorytmów detekcji w różnych warunkach oświetlenia. Robot może służyć również do porównywania różnych podejść programistycznych od klasycznych metod przetwarzania obrazu, przez algorytmy uczenia maszynowego, aż po elementy uczenia ze wzmocnieniem co umożliwi ocenę ich efektywności w rzeczywistym układzie fizycznym. Takie podejście sprzyja rozwijaniu umiejętności analitycznych, pracy projektowej oraz dokumentowania wyników eksperymentów co stanowi istotny element kształcenia inżynierskiego.

7. Podsumowanie

Projekt ten pokazuje, że możliwe jest stworzenie stosunkowo nisko kosztowej, modularnej i bezpiecznej platformy dydaktycznej. Zastosowanie technologii druku 3D umożliwiło szybkie prototypowanie oraz elastyczne dostosowywanie konstrukcji do zmieniających się potrzeb edukacyjnych. Wykorzystanie popularnych i łatwo dostępnych komponentów, takich jak Raspberry Pi, serwomechanizmy oraz silniki krokowe pozwalają na jego replikację w innych jednostkach dydaktycznych.

Własne API ograniczające ruchy robota minimalizują ryzyko uszkodzenia mechaniki podczas nauki. Takie rozwiązanie sprzyja bezpiecznemu środowisku eksperymentalnemu, szczególnie dla osób rozpoczynających pracę z robotyką. Robot może stanowić skuteczne narzędzie wspierające kształcenie w obszarze nowoczesnych technologii. Modularność konstrukcji oraz możliwość modyfikacji oprogramowania czynią system otwartą platformą rozwojową, którą można dostosowywać do różnych poziomów zaawansowania użytkowników. Ponadto projekt posiada potencjał skalowania i adaptacji, co czyni go perspektywicznym rozwiązaniem zarówno w dydaktyce jak i w szkoleniach technicznych.

Literatura i źródła internetowe

[1] Borgonovi, F. et al. (2025), AI adoption in the education system: International insights and policy considerations for Italy, OECD Artificial Intelligence Papers, No. 52, OECD Publishing, Paris/Fondazione Agnelli, Turin, https://doi.org/10.1787/69bd0a4a-en

[2] Freeman, J. Student generative ai survey 2025. Higher Education Policy Institute: London, UK, 2025.

[3] Gareth Halfacree, The Official Raspberry Pi Beginner's Guide, 6th ed. Cambridge, U.K: Raspberry Pi Press, 2026.

[4] https://www.waveshare.com/st3215-servo.htm

[5] S. Tickoo, Autodesk Fusion 360: A tutorial Approach, 6th Edition Schererville, IN, USA: CADCIM Technologies, 2024

[6] I Gibson, D. W. Rosen, B Stucker, and Mahyar Khorasani, Additive manufacturing technologies. Cham, Switzerland: Springer, 2021.

[7] J. J. Craig, Introduction to Robotics: Mechanics and Control, 3rd ed. Upper Saddle River, NJ, USA: Pearson Education, 2004, pp. 117-121

[8] https://www.adafruit.com/product/5832

ROBOTIC SORTING ARM

Summary

The objective of the project described in this article is to design and build a robot intended for learning artificial intelligence through practical programming of a robotic arm. The arm, equipped with a gripper, transfers objects of various shapes, such as cones, cubes, and cylinders. The robot's components are manufactured using 3D printing technology, which provides high design flexibility. The robot also uses standard mechanical components commonly applied in such constructions, including servomechanisms, stepper motors, DC motors, cable actuators, and the popular Raspberry Pi microcomputer. Dedicated API was created for the robot on the Raspberry Pi, limiting the robot's maximum movements to prevent the user from accidentally damaging the device by excessively rotating any of its parts. The robot will use artificial intelligence to locate objects (also 3D-printed) on the board. Based on the calculated position from cameras mounted on the robot, it will move the object from the board to the appropriate container. The robot will compute the path along which the gripper will travel to pick up the object and then move it above the designated box and release it. Both the board and the containers will be stationary elements. Thanks to the robot, students will be able to gain experience in operating artificial intelligence systems in an accessible and practical way. The project combines topics from mechanics, electronics, programming, embedded system design, and image processing, creating an interdisciplinary educational environment.

Keywords: Artificial Intelligence, Raspberry Pi, CAD modeling, inverse kinematics.

Sandra Kudła
ForResearchin@proton.me Wincentego Pola 2, 35-021 Rzeszów
Koło Naukowe Elektroniki i Technologii Informacyjnych

dr inż. Bartosz Pawłowicz
opiekun Koła Naukowego

AUTOMATYZACJA ODCZYTU KSIĄG METRYKALNYCH URODZEŃ Z WYKORZYSTANIEM *GOOGLE VISION API*

Streszczenie

Celem rozdziału jest przedstawienie aplikacji internetowej wspomagającej odczyt dawnych ksiąg metrykalnych urodzeń z wykorzystaniem Google Vision API. System umożliwia wgranie skanu strony, automatyczne rozpoznanie tekstu, zapis pełnego wyniku OCR oraz ekstrakcję podstawowych danych osobowych, takich jak imię, nazwisko, data urodzenia i imiona rodziców. Rozwiązanie łączy gotową usługę OCR z własnym modułem analizy układu tabeli i prostymi regułami dopasowanymi do specyfiki łacińsko-polskich zapisów metrykalnych. Dane po przetworzeniu są zapisywane w bazie PostgreSQL i wstępnie walidowane z użyciem biblioteki Pydantic. W rozdziale omówiono architekturę systemu, wybrane elementy implementacji oraz przykład działania na rzeczywistych skanach ksiąg parafialnych. Zwrócono również uwagę na ograniczenia rozwiązania, w szczególności na błędy OCR i konieczność ręcznej weryfikacji części wyników. Zaproponowane narzędzie może stanowić praktyczne wsparcie dla badań genealogicznych i projektów humanistyki cyfrowej

Słowa kluczowe: OCR, Google Vision API, PostgreSQL, Python, księgi metrykalne, cyfryzacja danych, genealogia, archiwistyka, rozpoznawanie pisma ręcznego.

1. Wprowadzenie

Księgi metrykalne urodzeń są jednym z najważniejszych źródeł wiedzy dla genealogów, badaczy historii ludności i lokalnych społeczności. Dotyczy to zwłaszcza okresów sprzed wprowadzenia współczesnych, państwowych rejestrów stanu cywilnego. Choć wiele ksiąg zostało już zeskanowanych, wciąż najczęściej

są one dostępne jako obrazy, czyli zdjęcia stron, z których tekst trzeba przepisywać ręcznie, co jest pracochłonne i podatne na pomyłki [[4], [5]].

Celem rozdziału jest przedstawienie aplikacji internetowej, która wspomaga automatyczne odczytywanie wpisów z ksiąg metrykalnych urodzeń. System wykorzystuje Google Vision API do rozpoznawania tekstu, a następnie przetwarza wynik OCR tak, aby odtworzyć strukturę tabeli i zapisać dane w uporządkowanej postaci. Szczególny nacisk położono na połączenie gotowej usługi OCR z prostymi regułami ekstrakcji dostosowanymi do układu dawnych ksiąg parafialnych. Dzięki temu możliwe jest częściowe zautomatyzowanie pracy, która w praktyce nadal często wykonywana jest ręcznie.

2. Podstawy teoretyczne i przegląd rozwiązań

W niniejszym rozdziale przedstawiono podstawy teoretyczne, na których opiera się projekt oraz najważniejsze istniejące podejścia do przetwarzania dokumentów historycznych z wykorzystaniem narzędzi OCR i HTR. Szczególną uwagę poświęcono charakterystyce ksiąg metrykalnych jako źródła danych, a następnie metodom automatycznego odczytu i analizy ich treści [[1], [3], [6]].

Księgi metrykalne jako źródło danych

Parafialne księgi chrztów i urodzeń z XVI–XVII wieku i późniejszych stuleci pozwalają odtworzyć dzieje rodzin i parafii. Zawierają w zwartej formie informacje o datach, miejscach, stopniu pokrewieństwa czy migracjach. Zapis ma zwykle częściowo tabelaryczną postać (kolumny: numer, data, imię, rodzice), ale bywa uzupełniony luźnym tekstem, skrótami po łacinie oraz lokalnymi wariantami językowymi. To wszystko utrudnia całkowicie automatyczne przetwarzanie takich źródeł [[4], [5]].

Z perspektywy informatyka księga metrykalna to zbiór stron o podobnym, ale nie identycznym układzie. Kluczowe jest prawidłowe odtworzenie wierszy i kolumn, tak aby imiona, nazwiska, daty i informacje o rodzicach trafiły do właściwych „pól" w rekordzie osoby. Wymaga to połączenia wyników z OCR z analizą położenia słów na stronie – w uproszczeniu można to porównać do rysowania na zeskanowanej stronie niewidzialnej siatki w kratkę i przypisywania wyrazów do odpowiednich kratek. Przykładowe słowniki słów łacińskich wykorzystywane do filtrowania szumu OCR oraz identyfikacji nagłówków tabel przedstawiono w **listingu 1**.

```
# Lista słów łacińskich i skrótów, które NIE są imionami ani nazwiskami
# (np. NUMERUS = numer wpisu, PATER = ojciec)
```python
LATIN_BLACKLIST = {
  'NUMERUS', 'DOMUS', 'NOMEN', 'COGNOMEN', 'PATRINI', 'CONDITIO',  # słowa księgi
  'PATER', 'MATER', 'FILIUS', 'FILIA', 'NATUS', 'BAPTISATUS',  # relacje rodzinne
  'JANUARIUS', 'FEBRUARIUS', 'MARTIUS', 'APRILIS', 'MAIUS',  # miesiące po łacinie
  'I', 'II', 'III', 'IV', 'V', 'VI', 'VII', 'VIII', 'IX', 'X'  # numery
}
#  Słowa typowe dla nagłówków tabel
HEADER_KEYWORDS = {
  'NUMERUS', 'NOMEN', 'NATUS', 'PATER', 'MATER', 'SERIES', 'ANNUS'
}
```

Listing 1. Słowniki łacińskie do filtrowania szumu i identyfikacji nagłówków
Listing 1. Latin dictionaries for noise filtering and header identification

OCR i HTR w dokumentach historycznych

Tradycyjne OCR sprawdza się dobrze w przypadku druków, lecz w dokumentach historycznych pojawia się szereg wielu dodatkowych wyzwań: specyficzne kroje pisma, degradacja papieru po czasie, użycie wielu języków, nieregularny układ oraz wiele użytych skrótów. Platformy HTR, takie jak *Transkribus*, pokazują, że trenowane na rękopisach modele głębokich sieci neuronowych pozwalają osiągać o wiele niższy wskaźnik błędów, nawet w trudnych przypadkach.

W praktyce humanistyki cyfrowej rzadko polega się na jednym narzędziu. Często wykorzystuje się OCR do wstępnego rozpoznania tekstu, a potem poprawia wynik przy pomocy modeli HTR lub ręcznej korekty. Coraz większe znaczenie mają też kolejne etapy przetwarzania: podział strony na logiczne fragmenty, odtworzenie tabel, rozpoznawanie osób, miejsc i innych jednostek. To właśnie te dodatkowe kroki decydują o tym, czy z rozpoznanego tekstu da się później wygodnie korzystać w badaniach [[1], [7], [8]].

Usługi chmurowe do OCR – *Google Vision API*

Google Vision API to usługa, która na podstawie obrazu potrafi zwrócić tekst oraz jego szczegółowy opis. Dla gęsto zapisanych dokumentów, takich jak księgi metrykalne, używa się trybu *document_text_detection*. W tym trybie wynik ma postać hierarchii: strony, bloki, akapity, słowa, a dodatkowo każde słowo ma zapisane współrzędne na obrazie (tzw. *bounding box*, czyli prostokąt obejmujący dane
```

słowo) [[1], [2]]. Przykładowy kod wyciągający z odpowiedzi Google Vision API tekst oraz współrzędne słów (bounding boxes) przedstawiono w **listingu 2**. W opisywanym systemie *Google Vision* działa jako „silnik rozpoznawania znaków" – coś w rodzaju specjalisty od odczytu liter. Wynik z tej usługi nie jest jednak pozostawiony w surowej formie. Własne algorytmy systemu przetwarzają go dalej: wyciągają z niego strukturę tabeli, szukają kolumn z datą, imieniem dziecka czy informacją o rodzicach i dopasowują je do z gór zdefiniowanego modelu danych. Dzięki temu można korzystać z gotowej, ogólnodostępnej usługi w chmurze, a jednocześnie mieć pełną kontrolę nad tym, jak konkretny typ dokumentu (tutaj: księgi metrykalne) jest interpretowany [[2]].

```
# Wywołanie Google Vision API z podpowiedziami odnośnie języka (łacina, polski)
response = vision_client.document_text_detection(
   image=gcp_image,
   image_context=vision.ImageContext(language_hints=['la', 'pl']) #  la = łacina,
pl = polski
)

# Wyciągnięcie listy słów z ich położeniem na stronie (bounding box)
words = []
for page in response.full_text_annotation.pages:
   for block in page.blocks:
      for para in block.paragraphs:
         for word in para.words:
            text = ''.join(symbol.text for symbol in word.symbols)
            box = word.bounding_box.vertices
            words.append({
               "text": text,
               "x": box[0].x,
               "y": box[0].y
            })
```

Listing 2. Przekształcanie odpowiedzi Google Vision API z „bounding boxes"
Listing 2. Transforming Google Vision API response with bounding boxes

Panel administracyjny Google Cloud Vision API z metrykami użycia (requesty, błędy, opóźnienia) przedstawiono na rys. 1.



Rys. 1. Konsola Google Cloud z metrykami Vision API

Fig. 1. Google Cloud console with Vision API metrics

2. Założenia projektowe oraz architektura systemu

Na podstawie ogólnych rozważań z rozdziału poprzedniego sformułowano kilka prostych wymagań, które zdecydowały o ostatecznej postaci rozwiązania. W tej części opisano, co dokładnie ma robić aplikacja, z jakich głównych klocków się składa oraz jak wyglądają dane, którymi się posługuje.

Cel funkcjonalny i wymagania

Podstawowym celem systemu jest:

1. przyjmowanie skanów stron ksiąg metrykalnych urodzeń,
2. uruchamianie procesu OCR z wykorzystaniem *Google Vision API*,
3. przekształcanie wyniku OCR w uporządkowane rekordy osób,

4. udostępnianie tych rekordów przez serwis sieciowy (API) i prosty interfejs w przeglądarce, tak aby można je było dalej analizować pod kątem genealogii i demografii.

Aplikacja przechowuje zarówno pełen tekst rozpoznany na stronie, jak i jego tłumaczenie na język polski. Dzięki temu można szukać informacji zarówno po oryginalnych słowach łacińskich, jak i po ich polskich odpowiednikach, co jest ważne dla użytkowników, którzy łaciną się nie posługują.

Wymagania poza funkcjonalne obejmują m.in. współpracę z popularną bazą *PostgreSQL*, możliwość uruchomienia wszystkich elementów w tzw. kontenerach (czyli „pudełkach" z gotową konfiguracją, które da się łatwo przenosić między serwerami), asynchroniczne przetwarzanie plików przez osobny moduł roboczy OCR (pliki są wysyłane do osobnego programu, który rozpoznaje tekst na skanach i przetwarza je w tle, tak żeby użytkownik mógł dalej normalnie pracować i nie musiał czekać, aż to się skończy) oraz prosty w utrzymaniu interfejs sieciowy opisany w formacie *OpenAPI*. Założono również, że aplikacja powinna radzić sobie z plikami o różnej jakości (różne formaty, tryby kolorów) oraz dawać się w przyszłości rozbudować o obsługę innych typów ksiąg, na przykład małżeństw i zgonów.

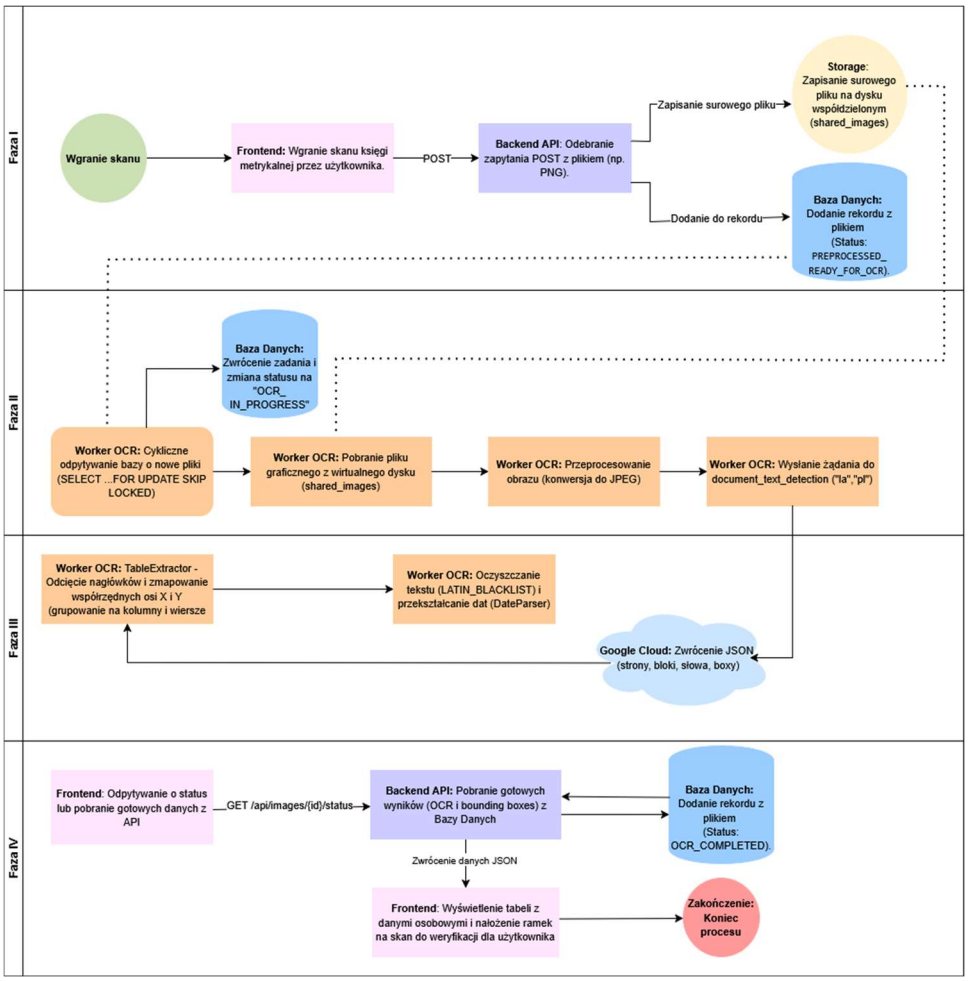
Przegląd architektury

System składa się z czterech głównych elementów: bazy danych PostgreSQL, serwisu API, modułu roboczego OCR oraz interfejsu użytkownika dostępnego w przeglądarce. Użytkownik przesyła skan strony, po czym moduł OCR przetwarza obraz z wykorzystaniem Google Vision API i zapisuje wynik w bazie danych. Następnie warstwa ekstrakcji przekształca surowy wynik OCR w uporządkowane rekordy osób. Interfejs użytkownika umożliwia podgląd skanu, tekstu rozpoznanego przez OCR oraz danych strukturalnych, które mogą być dalej poprawiane i eksportowane.

Moduł roboczy OCR działa w sposób ciągły: co chwilę sprawdza, czy w bazie pojawiły się nowe pliki do przetworzenia, wybiera jeden z nich, zmienia jego status, wysyła obraz do usługi Vision, a po otrzymaniu odpowiedzi zapisuje dane i oznacza zadanie jako ukończone. Aby łatwo śledzić, na jakim etapie jest dany plik, stosuje się prosty zestaw stanów, na przykład: „gotowy do OCR", „w trakcie przetwarzania", „zakończony", „błąd". Ogólny schemat architektury systemu, obejmujący frontend, API, worker OCR, bazę PostgreSQL i katalog plików, przedstawiono na rys. 2.

Model danych

Po stronie aplikacji napisanej w Python dane reprezentowane są za pomocą klas z biblioteki Pydantic. Można je traktować jak „szablony" określające, jakie pola powinien mieć poprawny rekord. Dla osoby są to na przykład: imiona, nazwisko, data urodzenia, imię ojca, imię matki oraz liczbowy wskaźnik wskazujący, na ile system „jest pewien", że rozpoznał te dane poprawnie.



Rys. 2. Architektura systemu i przepływ danych w procesie przetwarzania ksiąg metrykalnych
Fig. 2. System architecture and data flow for processing parish register books

Dodatkowo modele zawierają proste zasady sprawdzające sensowność danych, na przykład walidator, który odrzuca datę urodzenia, jeśli znajduje się ona w przyszłości względem roku bieżącego. Model Pydantic reprezentujący jedną osobę z księgi metrykalnej, wraz z walidacją daty urodzenia, przedstawiono w listingu 3. W innych polach przewidziano miejsce na informacje o parafii, roku zapisu i numerze wpisu, co pozwala później filtrować dane według miejsca i czasu oraz łączyć je z innymi źródłami, takimi jak gotowe indeksy czy rejestry sądowe [[7]]. Model Pydantic reprezentujący jedną osobę z księgi metrykalnej, wraz z walidacją daty urodzenia, przedstawiono w **listingu 3**. Natomiast rysunku 3. ilustruje strukturę tabeli w bazie PostgreSQL (structured_data) z kolumnami image_id, given_names, family_name itp.

```
from pydantic import BaseModel, validator, Field  #  biblioteka do tworzenia modeli
z walidacją
from typing import Optional  #  typ “opcjonalny”
from datetime import date  #  typ daty
class Person(BaseModel):  #  szablon dla jednej osoby z księgi metrykalnej
    given_names: str = Field(..., max_length=255)  #  Imiona dziecka
    family_name: str = Field(..., max_length=255)  #  Nazwisko rodzinne
    birth_date: Optional[date] = None  #  Data urodzenia
    father_name: str = Field(default="", max_length=255)  #  Imię ojca
    mother_name: str = Field(default="", max_length=255)  #  Imię matki
    confidence: float = Field(ge=0.0, le=1.0, default=0.95)  #  Pewność rozpoznania (%)
    @validator('birth_date')  #  automatyczna walidacja pola daty
    def valid_date(cls, v):
        if v and v.year > 2026:  #  Sprawdzanie czy data nie jest z przyszłości
            raise ValueError('Data urodzenia nie może być z przyszłości')
        return v
```

Listing 3. Model Pydantic, który weryfikuje rekord osoby z ograniczeniami daty

Listing 3. Pydantic model validating person record with date constraints

```
parish_archive=# \d structured_data;
                                      Table "public.structured_data"
   Column    |            Type             | Collation | Nullable |                  Default
-------------+-----------------------------+-----------+----------+-------------------------------------------
 id          | integer                     |           | not null | nextval('structured_data_id_seq'::regclass)
 image_id    | integer                     |           |          |
 given_names | character varying(255)      |           |          |
 family_name | character varying(255)      |           |          |
 birth_date  | date                        |           |          |
 father_name | character varying(255)      |           |          |
 mother_name | character varying(255)      |           |          |
 confidence  | numeric(3,2)                |           |          |
 created_at  | timestamp without time zone |           |          | CURRENT_TIMESTAMP
 updated_at  | timestamp without time zone |           |          | CURRENT_TIMESTAMP
Indexes:
    "structured_data_pkey" PRIMARY KEY, btree (id)
    "idx_structured_image" btree (image_id)
    "idx_structured_name" btree (family_name)
Foreign-key constraints:
    "structured_data_image_id_fkey" FOREIGN KEY (image_id) REFERENCES files(id)
parish_archive=# SELECT image_id, given_names, family_name FROM structured_data ORDER BY id DESC LIMIT 5;
 image_id | given_names |                                  family_name
----------+-------------+-------------------------------------------------------------------------
       21 | Legitimn    | Andres Josephi Sovcaya Francised Legitoina Dehman Sagan Mariamma Forsingum
       21 | Isthma      | Cherepakna Mari Laws Mariama
       21 | Sohotelvic  | Adalbertus Ishra Catharina Joan Victoria Lawst
       21 | Antoninos   | Mypatha Valentinit Krzypink.com
       21 | Andres      | Madej Galeakset
(5 rows)
```

Rys. 3. Zawartość tabeli structured_data w bazie PostgreSQL z zapisanymi rekordami osób

Fig. 3. Contents of the structured_data table in PostgreSQL containing saved person records

3. Implementacja wybranych komponentów

W tym rozdziale pokazano, jak omówione wcześniej założenia zostały przełożone na konkretne rozwiązania w kodzie programu oraz w konfiguracji środowiska uruchomieniowego. Opis ma prowadzić czytelnika krok po kroku: od modułu roboczego OCR, przez wycinanie tabeli z obrazu, aż po serwis API i sposób uruchamiania wszystkiego w kontenerach.

Worker OCR i przetwarzanie wsadowe

Moduł roboczy OCR jest osobnym programem w *Pythonie*, który po uruchomieniu łączy się z bazą *PostgreSQL*, a następnie w nieskończonej pętli szuka w niej zadań oznaczonych jako „gotowe do OCR". Gdy takie zadanie znajdzie, blokuje je dla siebie, dzięki czemu dwa procesy naraz nie próbują przetworzyć tego samego pliku. Potem wczytuje obraz, w razie potrzeby zamienia go na format JPEG, który najlepiej współpracuje z *Google Vision*, i wysyła zawartość do usługi rozpoznawania tekstu.

Po otrzymaniu wyniku *worker* wyciąga z odpowiedzi pełny tekst oraz dane o położeniu słów na stronie, zapisuje je w tabeli wyników w bazie i zmienia stan pliku na „zakończony". Jeżeli po drodze pojawi się błąd (na przykład usługa Google nie odpowie), *worker* oznacza dane zadanie odpowiednim statusem błędu, zapisuje informacje diagnostyczne w logach i przechodzi do kolejnego pliku. Dzięki temu przetwarzanie wielu stron może być realizowane „hurtowo", bez ręcznej interwencji [[2]]. Główną pętlę modułu roboczego OCR, realizującą wyszukiwanie zadań, blokowanie rekordu (SKIP LOCKED) oraz obsługę błędów, pokazano w **listingu 4**.

```python
def process_pending_files():
    with get_db_connection() as conn:  #  połączenie z bazą PostgreSQL
        with conn.cursor(cursor_factory=RealDictCursor) as cur:
            #  Pobiera jeden plik do OCR (blokada SKIP LOCKED zapobiega konfliktom)
            cur.execute("""  #  Znajdź pierwszy plik gotowy do przerobienia przez OCR
(bez blokad)
                SELECT id, filepath FROM files
                WHERE status = 'PREPROCESSED_READY_FOR_OCR'
                ORDER BY created_at ASC LIMIT 1 FOR UPDATE SKIP LOCKED
            """)
            row = cur.fetchone()
            if not row:
                return False

            file_id, filepath = row['id'], row['filepath']
```

```
        logging.info(f"OCR dla pliku {file_id}: {filepath}")
        cur.execute("UPDATE files SET status = 'OCR_IN_PROGRESS' WHERE id
= %s", (file_id,))  #  Zamiana statusu na “w trakcie”
        conn.commit()  #  zapis zmian w bazie
        try:
          # Konwersja PNG → JPEG dla Vision API (Vision lepiej sobie radzi
z JPEG)
          image_content = convert_png_to_readable_jpeg(filepath)
          gcp_image = vision.Image(content=image_content)

          # OCR z łaciną i polskim
          response = vision_client.document_text_detection(
            image=gcp_image,
            image_context=vision.ImageContext(language_hints=['la', 'pl'])
          )
          # Wyciągnięcie tekstu i bounding boxes
          extracted_data = extract_ocr_data(response)
          cur.execute("""
            INSERT INTO ocr_results (file_id, fulltext, bounding_boxes, confidence)
            VALUES (%s, %s, %s, %s)
          """, (file_id, extracted_data['fulltext'],
              json.dumps(extracted_data['bounding_boxes']),
              extracted_data['confidence']))
          cur.execute("UPDATE files SET status = 'OCR_COMPLETED' WHERE
id = %s", (file_id,))  #  Zakończenie sukcesem
          conn.commit()
          logging.info(f"OCR OK {file_id}")

        except Exception as e:
          logging.error(f"Błąd OCR {file_id}: {e}")
          cur.execute("UPDATE files SET status = 'OCR_FAILED' WHERE
id = %s", (file_id,))
          conn.commit()
      return True  #  kontynuuj pętle
```

Listing 4. Fragment workera OCR – pętla przetwarzania z obsługą błędów i bounding boxes
Listing 4. OCR worker fragment – processing loop with error handling and bounding boxes

Ekstrakcja struktury zapisu metrykalnego

Kolejnym ważnym elementem jest moduł, który z surowej odpowiedzi *Vision* tworzy coś, co dla człowieka wygląda jak prosta tabela. Najpierw zbiera on wszystkie słowa wraz z ich współrzędnymi na stronie. Następnie dzieli stronę na

wiersze (na podstawie podobnej wysokości wyrazów) oraz na cztery główne kolumny: numer i informacje pomocnicze, datę urodzenia, imię i nazwisko dziecka oraz opis rodziców. Można to porównać do nakładania na skan linijek poziomych i pionowych i odczytywania, co znalazło się w każdym „okienku".

Aby nie pomylić nagłówków kolumn z właściwymi danymi, moduł korzysta ze słowników zawierających typowe łacińskie słowa wynikające z formularza, a także prostych reguł typu „imiona zaczynają się wielką literą, mają pewną długość i nie występują na liście słów-nagłówków". W przypadku kolumny opisującej rodziców system szuka słów-znaczników, takich jak PATER / OJCIEC i MATER / MATKA, i na tej podstawie dzieli tekst na listę imion ojca i listę imion matki.

Z obu list budowane są następnie pola *fathername* i *mothername* w rekordzie osoby. Funkcje walidacji imion oraz podziału kolumny rodziców na ojca i matkę na podstawie znaczników (PATER/MATER) przedstawiono w **listingu 5**.

```
def is_valid_name(word):
    if not word or len(word) < 2:
        return False
    if TableExtractor.is_header(word) or TableExtractor.is_keyword(word):
        return False
    return word[0].isupper()

for w in raw_parents:
    token = re.sub(r'[,.;]', '', w.text).upper()
    if 'PATER' in token or 'OJCIEC' in token:
        target = fathers
    elif 'MATER' in token or 'MATKA' in token:
        target = mothers
    if is_valid_name(w.text):
        target.append(w.text)
```

Listing 5. Fragment ekstrakcji struktury – walidacja imion i podział rodziców

Listing 5. Structure extraction fragment – name validation and parent separation

API i aplikacja webowa

Warstwa API systemu została opisana w specyfikacji OpenAPI 3.0.3. Definiuje ona operacje związane z przesyłaniem skanów, uruchamianiem OCR, pobieraniem wyników rozpoznawania, kontrolą statusu przetwarzania oraz zapisem poprawionych danych strukturalnych. Na tej podstawie przygotowano aplikację webową w bibliotece React, która umożliwia wgrywanie obrazów, śledzenie postępu przetwarzania oraz podgląd wyników OCR i danych tabelarycznych. Użytkownik widzi równocześnie skan strony, tekst rozpoznany przez OCR oraz tabelę wyekstrahowanych rekordów. Dodatkowo interfejs umożliwia powiązanie fragmentów

tekstu z ich położeniem na obrazie za pomocą bounding boxes, co ułatwia weryfikację poprawności ekstrakcji. Uproszczony fragment specyfikacji OpenAPI dla punktu końcowego służącego do zapisu poprawionych rekordów osobowych przedstawiono w **listingu 6**.

```
paths:
  /api/images/{imageId}/structured-data:
    post:
      summary: Save structured person records
      parameters:
        - name: imageId
          in: path
          required: true
          schema:
            type: integer
      requestBody:
        content:
          application/json:
            schema:
              type: object
              properties:
                people:
                  type: array
```

Listing 6. Specyfikacja OpenAPI dla zapisu rekordów osobowych

Listing 6. OpenAPI specification for saving person records

Przykładową tabelę strukturalnych rekordów osobowych wyekstrahowanych z księgi metrykalnej, gotową do eksportu CSV, jak i w formacie JSON przedstawiono na rys. 4. oraz rys. 5.

	A	B	C	D	E
1	Imię	Nazwisko	Data	Ojciec	Matka
2	Thoma	Andrea Catherine Ravbars Stee Navionet Pantai	1884	Sharbare Wola Vicolor Stic Marima	
3	Michiel	Lyp Ama Typ Theresse Jacobi Karbar Klosiert Rawol Ponte		Joannes Dilia Valentine Sophie Dudek Dorothea Udalbert Saganagrieche	
4	Indovica		1884	Sophia Harbare Kruse Thecla	
5	Jects	Parton		Catavina Madej	
6	Cavoloes	Hawbars Francisca Karbave Joan- Magdalena Chyla Partei Machiël Banach Sophia	1886	Mazur Jacobus So- Mazur Ihrawiec MMa Texmeid Person Plavia Torcayed Tusiage Simonciscus	
7	Laurenti	Banach Agnetis Stee Maviorai		Zueiega Banachet Fron Bis Antoni Bis	
8	Josephus	Machay Trade Mayarithw Machaj So- Falcas Istles Typ	1944-09-14	Francisco Waush Taleak Joanne's Turin Francisca Thonce Polcare Josepha	
9	Parties				
10	Thomas	Salcak		Burcayez	
11	Auto	Sophie Taloak Mar- Mazer Myszka Andrea	1903	Marus Michiel Ben Evie Peter Mar Suvia Martini	
12	Portin	Ich		Sithora	
13	Fetus	Palcas Pauli Amma Galeat Marianne Prince Jilice Matosha Peteriet Catha Antoni Agnetis		Soveryed Legeli Parturi Marianno For- Machaj Thoma Catharina Ma	
14	Margaritha			Lo	
15	James	Harbary Marianna		Lyple	
16	Tho		1954-08-18	The Joarmes Wola-	
17	Agnetis	Galeak Jacobs Lyn Catharina		Machai Cotborita Rublov Madei	

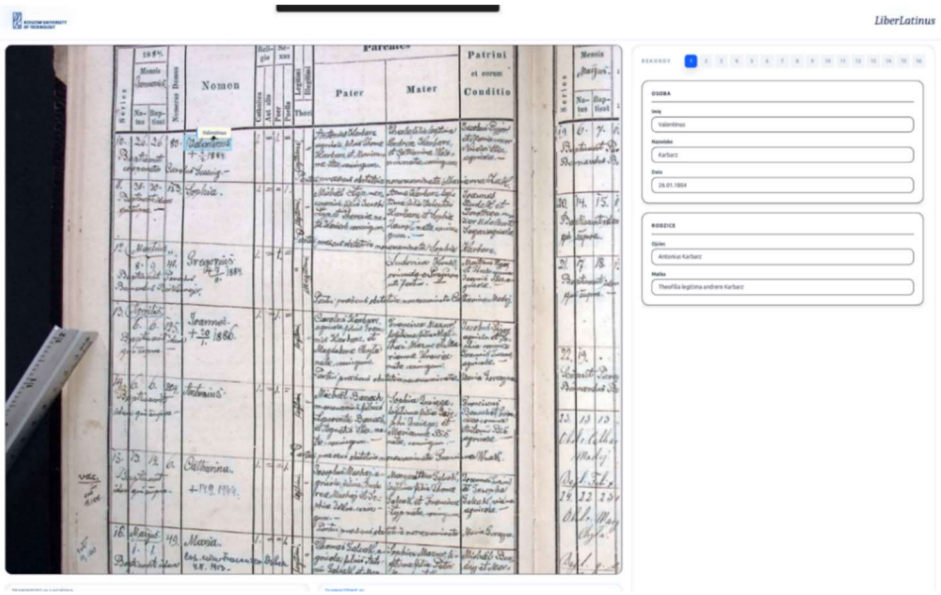
Rys. 4. Tabela strukturalnych rekordów osobowych wywyekstrahowanych z księgi – dane gotowe do eksportu CSV

Fig. 4. Table of structured person records extracted from the book – data ready for CSV export



Rys. 5. Strukturalny zapis wyników OCR i tłumaczenia w formacie JSON
(pełny tekst i wyodrębnione osoby)

Fig. 5. Structured representation of OCR and translation results in JSON format
(full text and extracted persons)



Rys. 6. Interfejs uwidaczniający bounding boxes oraz pełny tekst, bloki, słowa

Fig. 6.Interface displaying bounding boxes, full text, blocks, and words

Interfejs użytkownika umożliwia równoczesny podgląd skanu strony, tekstu rozpoznanego przez OCR oraz tabeli wyekstrahowanych rekordów. Dodatkowo interfejs umożliwia powiązanie fragmentów tekstu z ich położeniem na obrazie za pomocą bounding boxes, co ułatwia weryfikację poprawności ekstrakcji. Widok interfejsu z podświetlonymi bounding boxes przedstawiono na rys. 6.

Implementację punktu końcowego FastAPI do zapisywania listy rekordów osób z walidacją daty urodzenia pokazano w **listingu 7**.

```
@app.post("/api/images/{image_id}/structured-data")  # adres: /api/images/(np.21)/
structured-data
async def save_structured_data(image_id: int, request: Request):
    data = await request.json()
    people_list = data.get('people', [])  # wyciągnięcie pola "people"

    conn = get_db()  # połączenie z bazą PostgreSQL
    cur = conn.cursor()
    for person in people_list:  # pętla po każdej osobie
        # Parsowanie daty: "1884" → "1884-01-01" (format PostgreSQL)
        birth_date = None
        raw_date = person.get('birth_date', '')
        if len(raw_date) == 4 and raw_date.isdigit():
            birth_date = f"{raw_date}-01-01"

        # INSERT do tabeli structured_data
        cur.execute("""
            INSERT INTO structured_data
            (image_id, given_names, family_name, birth_date, father_name,
mother_name)
            VALUES (%s, %s, %s, %s, %s, %s)
        """, (image_id, person['given_names'], person['family_name'],
              birth_date, person['father_name'], person['mother_name']))

    conn.commit()  # zapis wszystkich zmian
    return {"saved_records": len(people_list)}
```

Listing 7. Punkt końcowy do zapisywania danych z plików OCR
Listing 7. Endpoint for saving data from OCR files

Rysunek 7. pokazuje logi modułu roboczego OCR podczas przetwarzania pliku image_id=27. Widoczne są etapy „Start ID: 27” (OCR_IN_PROGRESS) oraz „Gotowe. Osób: X” (OCR_COMPLETED).

```
2026-02-27 12:56:04,699 - WORKER - INFO - Worker gotowy.
2026-02-27 13:04:01,492 - WORKER - INFO - Start ID: 27
2026-02-27 13:04:03,981 - WORKER - INFO - Gotowe. Osób: 22
2026-02-27 13:24:39,088 - WORKER - INFO - Start ID: 28
2026-02-27 13:24:41,694 - WORKER - INFO - Gotowe. Osób: 16
2026-02-27 13:32:43,332 - WORKER - INFO - Worker gotowy.
2026-02-27 13:32:54,588 - WORKER - INFO - Start ID: 29
2026-02-27 13:32:57,249 - WORKER - INFO - Gotowe. Osób: 16
2026-02-27 13:33:33,009 - WORKER - INFO - Start ID: 28
2026-02-27 13:33:35,476 - WORKER - ERROR - Błąd: duplicate key value violates unique constraint "uq_ocr_results_file_id"
DETAIL:  Key (file_id)=(28) already exists.

2026-02-27 13:34:38,058 - WORKER - INFO - Start ID: 30
2026-02-27 13:34:40,105 - WORKER - INFO - Gotowe. Osób: 6
2026-02-27 13:35:49,655 - WORKER - INFO - Start ID: 31
2026-02-27 13:35:51,944 - WORKER - INFO - Gotowe. Osób: 6
2026-02-27 13:36:03,223 - WORKER - INFO - Start ID: 32
2026-02-27 13:36:05,491 - WORKER - INFO - Gotowe. Osób: 6
```

Rys. 7. Logi modułu roboczego OCR podczas przetwarzania pliku image_id=27. Widoczne etapy: „Start ID: 27” (OCR_IN_PROGRESS), **„Gotowe. Osób: X”(OCR_COMPLETED)

Fig. 7. OCR worker logs during processing of image_id=27. Visible stages: “Start ID: 27” (OCR_IN_PROGRESS), “Done. Persons: X” (OCR_COMPLETED)

Widok przetłumaczonego tekstu metrykalnego w panelu Tłumaczenie AI aplikacji, pokazujący oryginalny tekst łaciński obok polskiego odpowiednika, przedstawiono na rys. 8.



Rys. 8. Widok przetłumaczonego tekstu metrykalnego w panelu Tłumaczenie AI aplikacji

Fig. 8. View of translated parish record text in the application’s AI Translation panel

Konteneryzacja i konfiguracja środowiska

Aby ułatwić uruchamianie systemu na różnych komputerach, wszystkie jego główne elementy zostały „spakowane” w kontenery Docker. Każdy kontener zawiera wszystko, czego dana część potrzebuje do działania: odpowiednią wersję *Python*, bibliotek, konfigurację połączenia z bazą danych itp. Plik *docker-compose* opisuje, jak uruchomić razem kontener z bazą, z serwisem API, z modułem roboczym OCR oraz ewentualnie z dodatkowymi komponentami. Kluczowe fragmenty pliku *docker-compose.*yml definiujące zależności między bazą danych, API i workerem OCR przedstawiono w **listingu 8**.

```
services:
  db:
    image: postgres:15-alpine

  ocr-api:
    build: .
    depends_on: [db]

  ocr-worker:
    build: .
    depends_on: [db]
    environment:
```

Listing 8. Kluczowe serwisy docker-compose.yml z zależnościami
Listing 8. Key services in docker-compose.yml with dependencies

4. Studium przypadku i ewaluacja systemu

Aby sprawdzić, czy zaproponowane rozwiązanie naprawdę pomaga w pracy z księgami, przygotowano prosty eksperyment. Polegał on na przetworzeniu zestawu stron z jednej księgi parafialnej, a następnie porównaniu wyników aplikacji z ręcznie przygotowaną transkrypcją.

Metodyka badań

Do badań wybrano jedną księgę zawierającą kilkadziesiąt wpisów urodzeń zapisanych w podobny sposób. Skanowane obrazy stron wgrywano do aplikacji przez stronę w przeglądarce i uruchamiano dla nich proces OCR. Rysunek 7. pokazuje log błędu bazy danych (duplicate key dla ocr_results.file_id=28) oraz próbę ponownego startu zadania ID=30. Dla każdej strony mierzono czas przetwarzania, sprawdzano poprawność rozpoznanego tekstu oraz to, na ile dobrze udało się odtworzyć tabelę osób.

Jako główną miarę jakości przyjęto odsetek wpisów, dla których system poprawnie rozpoznał komplet podstawowych danych: imię i nazwisko dziecka, co

najmniej jedno imię rodzica (ojca lub matki) oraz datę urodzenia. Wyniki porównywano z ręcznie sporządzoną transkrypcją, tak jak robi się to w typowych badaniach nad OCR i HTR w dokumentach historycznych. Dodatkowo analizowano rodzaje najczęstszych błędów, na przykład zamianę łacińskich nazw miesięcy, łączenie kilku słów w jedno nazwisko albo mylne przypisywanie imion do niewłaściwego rodzica [[1], [6]]

Scenariusz wysokiego obciążenia i reakcja systemu

W scenariuszu większego obciążenia użytkownik wgrywał od razu większy zestaw skanów w formacie PNG. System konwertował je do JPEG i zapisywał w katalogu współdzielonym, a w bazie danych rejestrował nowe zadania ze stanem „oczekuje na OCR". Moduł roboczy OCR sukcesywnie pobierał kolejne pliki, zmieniał im status, wywoływał usługę *Vision* i zapisywał wyniki. Użytkownik mógł na bieżąco śledzić postęp prac w interfejsie oraz po zakończeniu zobaczyć listę wszystkich przetworzonych stron.

Dla przykładowej strony z kilkunastoma wpisami urodzeń moduł ekstrakcji struktury budował listę obiektów „osoba" wraz z przypisanymi polami. Lista ta była prezentowana w tabeli, z której dane można było wyeksportować do dalszych analiz lub poprawić ręcznie. Taki tryb pracy odpowiada potrzebom badaczy genealogii opisanym w literaturze: punktem wyjścia jest skan trudny do przeszukania, a efektem – edytowalny i łatwy do przeszukania zbiór rekordów, który w razie potrzeby ktoś może jeszcze dopracować [[4], [8]].

Dyskusja wyników

Uzyskane wyniki pokazują, że połączenie usługi *Google Vision* z warstwą heurystycznej ekstrakcji tabel umożliwia poprawne odtworzenie podstawowej struktury wpisów metrykalnych w znacznej części przypadków, przy akceptowalnym poziomie błędów możliwych do wychwycenia w procesie korekty eksperta. Szczególnie dobrze rozpoznawane są daty w formacie cyfrowym (DD.MM.RRRR) oraz imiona zapisane w miarę czytelnie, natomiast większe trudności pojawiają się przy nieregularnych wariantach pisma odręcznego i silnie skróconych formułach łacińskich [[1], [3]].

Porównanie z wynikami uzyskiwanymi przez specjalistyczne modele HTR wskazuje, że choć rozwiązanie oparte na *Google Vision* nie osiąga równie niskich współczynników błędów na poziomie pojedynczych znaków, to jednak jest znacznie prostsze do uruchomienia i nie wymaga czasochłonnego uczenia modelu na lokalnym materiale. Widok przetłumaczonego tekstu metrykalnego w panelu „Tłumaczenie AI" aplikacji, pokazujący oryginalny tekst łaciński obok polskiego odpowiednika, przedstawiono na rys. 8. W połączeniu z możliwością ręcznej korekty na stronie daje to rozsądną równowagę między nakładem pracy, a jakością danych wyjściowych [[6], [9]].

Ograniczenia i możliwe błędy

Do głównych ograniczeń zaproponowanego rozwiązania należy zależność od zewnętrznej usługi chmurowej, co rodzi kwestie kosztów, prywatności danych oraz wymaga stabilnego połączenia sieciowego. Kolejnym problemem jest wrażliwość heurystyk ekstrakcji tabeli na zmienność układu ksiąg – przesunięcie kolumn, zmiana szerokości marginesów czy dodatkowe notatki na marginesach mogą prowadzić do błędów w przypisaniu słów do odpowiednich pól [[2]].

Ważnym ograniczeniem zaproponowanego rozwiązania jest brak jednoznacznej i czytelnej dla operatora informacji o pewności rozpoznania każdego słowa lub fragmentu tekstu. W praktyce oznacza to, że nawet jeśli system poprawnie odczyta większość wpisu, użytkownik nie zawsze otrzymuje jasny sygnał, które elementy wymagają szczególnej kontroli. Ma to istotne znaczenie zwłaszcza w przypadku łaciny kościelnej, skrótów oraz mniej czytelnego pisma odręcznego. Z tego względu wyniki działania systemu powinny być traktowane jako materiał wstępnie opracowany, który w razie wątpliwości wymaga weryfikacji przez badacza znającego specyfikę dawnych ksiąg metrykalnych.

Ponadto obecna wersja systemu nie wykorzystuje zaawansowanych metod statystycznych ani uczenia maszynowego do automatycznej korekty imion i nazwisk, np. w oparciu o lokalne słowniki nazwisk czy modele językowe, które są coraz częściej stosowane w projektach przetwarzania dokumentów historycznych.

Wreszcie, mimo zastosowania walidacji dat i długości pól, jakość danych końcowych nadal w dużym stopniu zależy od zaangażowania ekspertów w proces korekty, co jest typowe dla projektów digitalizacji ksiąg parafialnych, ale ogranicza pełną automatyzację procesu [[4], [7], [8]].

5. Podsumowanie

W rozdziale opisano aplikację internetową, która automatyzuje dużą część pracy związanej z odczytywaniem i porządkowaniem wpisów z ksiąg metrykalnych urodzeń. Rozwiązanie opiera się na usłudze *Google Vision API*, osobnym module roboczym do przetwarzania skanów oraz relacyjnej bazie danych z modelami sprawdzającymi poprawność informacji osobowych. W praktyce narzędzie to wypełnia lukę pomiędzy całkowicie ręcznym przepisywaniem ksiąg, a bardzo zaawansowanymi, ale trudniejszymi w użyciu platformami HTR [[3], [6], [9]].

Przedstawiony system może stanowić użyteczne narzędzie wspierające badania genealogiczne i projekty humanistyki cyfrowej, zwłaszcza jako rozwiązanie pośrednie między ręczną transkrypcją a bardziej złożonymi platformami rozpoznawania pisma odręcznego []. Jednocześnie wyniki działania systemu wymagają ostrożnej interpretacji, ponieważ jakość ekstrakcji zależy od czytelności skanu, układu tabeli oraz poprawności rozpoznania tekstu przez usługę OCR []. W dalszych pracach zasadne wydaje się rozwinięcie mechanizmów oznaczania

niepewnych wyników, integracja z wyspecjalizowanymi modelami HTR oraz zastosowanie metod rozpoznawania nazw własnych i innych encji w dokumentach historycznych.

Literatura

[1] Enhancing *OCR in historical documents with complex layouts*, Dostępny: https://www.econstor.eu/bitstream/10419/319163/1/00799_2025_Article_413.pdf (dostęp: 19 lutego 2026).

[2] Google Cloud, *Detect and extract text from images – Cloud Vision API*, Dostępny: https://docs.cloud.google.com/vision/docs/ocr/ (dostęp: 19 lutego 2026).

[3] Gribomont I., *OCR with Google Vision API and Tesseract*, Programming Historian, *Dostępny*: https://programminghistorian.org/en/lessons/ocr-with-google-vision-and-tesseract (dostęp: 19 lutego 2026).

[4] *Handwritten Parish Registers OCR*, [Online]. Dostępny: https://www.handwritingocr.com (dostęp: 19 lutego 2026).

[5] *Historical Geography Research Series (HGRG),* [Online]. Dostępny: https://hgrg.org.uk/wp-content/uploads/2022/11/hgrs-no.7-feb-1981.pdf (dostęp: 19 lutego 2026).

[6] Kahle P. i in., *Transkribus: Handwritten Text Recognition technology for historical documents*, [Online]. Dostępny: https://dh2017.adho.org/abstracts/649/649.pdf (dostęp: 19 lutego 2026).

[7] Orasmaa S. i in., *Named Entity Recognition in Estonian 19th Century Parish Court Records*, Dostępny: https://aclanthology.org/2022.lrec-1.568.pdf (dostęp: 19 lutego 2026).

[8] *Reading and Transcribing Old Handwritten Documents: Transkribus*, Discover Genealogy *Blog*, Dostępny: https://discovergenealogy.blogspot.com/2025/05/reading-and-transcribing-old.html (dostęp: 19 lutego 2026).

[9] *Transkribus - Historical Documents with AI*, Dostępny: https://www.ghentcdh.ugent.be/transkribus-historical-documents-ai (dostęp: 19 lutego 2026).

Design and Implementation of a Text Recognition System for Parish Birth Registers using *Google Vision API*

Summary

This paper presents a web-based system supporting the transcription of historical parish birth registers using Google Vision API. The application allows users to upload scanned pages, perform OCR, store the full recognition result, and extract selected personal data such as given names, family names, birth dates, and parents' names. The proposed solution combines a general OCR service with a custom table-structure extraction layer adapted to Latin-Polish parish records. Extracted data are stored in a PostgreSQL database and validated with Pydantic models. The paper discusses the system architecture, selected implementation details, and a case study based on real parish register scans. Particular attention is given to the limitations of the approach, including OCR errors and the need for expert verification of ambiguous records. The presented system may serve as a useful tool for genealogy-oriented research and digital humanities projects.

Keywords: OCR, Google Vision API, PostgreSQL, Python, Parish registers, Data digitization, Genealogy, Archival studies, Handwriting recognition.

Zuzanna Dörfler, Ewelina Nycz
Studenckie Koło Naukowe Komunikacji Marketingowej „Brief"

MOJA MARKA, MOJA SIŁA: ZNACZENIE MARKI OSOBISTEJ W ROZWOJU KARIERY ZAWODOWEJ

Streszczenie

Współczesna nauka o zarządzaniu wskazuje, że marka osobista (personal branding) stanowi jedno z kluczowych narzędzi strategicznego zarządzania karierą zawodową jednostki. W warunkach dynamicznie zmieniającego się rynku pracy, postępującej cyfryzacji oraz rosnącej roli modelu gig economy, świadome kształtowanie wizerunku profesjonalnego staje się istotnym elementem budowania przewagi konkurencyjnej. Celem niniejszego rozdziału jest wieloaspektowe przedstawienie istoty marki osobistej, identyfikacja jej kluczowych determinantów oraz krytyczna analiza korzyści i zagrożeń związanych z procesem jej kreowania. W pierwszej części opracowania przybliżono definicję marki osobistej oraz jej genezę, osadzając ją w kontekście marketingu relacyjnego i nowoczesnego zarządzania zasobami ludzkimi. Następnie dokonano dekompozycji architektury marki, omawiając jej fundamentalne komponenty: autentyczność, system wartości, kompetencje typu T-shaped oraz spójność komunikacyjną. Kolejna sekcja koncentruje się na operacjonalizacji działań brandingowych, analizując efektywność narzędzi online (w tym roli serwisu LinkedIn i storytellingu) oraz znaczenie aktywności offline, takich jak networking i udział w projektach branżowych. W dalszej części rozdziału wykazano wpływ silnej marki na trajektorię rozwoju zawodowego, ze szczególnym uwzględnieniem wzrostu wskaźnika zatrudnialności (employability) oraz akumulacji kapitału społecznego. Autorzy wskazują również na potencjalne wyzwania i bariery, w tym ryzyko niespójności wizerunkowej, kryzysy reputacyjne w przestrzeni cyfrowej oraz etyczne dylematy związane z nadmierną autopromocją. Podsumowanie rozdziału podkreśla, że budowanie marki osobistej jest procesem ciągłym i długofalowym, wymagającym nie tylko biegłości technologicznej, ale przede wszystkim refleksji strategicznej i etycznego podejścia do komunikacji. Wskazano także na perspektywy dalszych badań, uwzględniające wpływ sztucznej inteligencji na autentyczność profesjonalnego wizerunku.

Słowa kluczowe: autoprezentacja, rozwój kariery, komunikacja marketingowa, wizerunek, networking.

1. Geneza i aktualność tematyki personal brandingu w naukach o zarządzaniu

Współczesna nauka o zarządzaniu coraz częściej odchodzi od postrzegania zasobów ludzkich jedynie przez pryzmat kosztów operacyjnych, skłaniając się ku traktowaniu jednostki jako kluczowego aktywa o charakterze strategicznym. W dobie gospodarki opartej na wiedzy marka osobista (personal brand) staje się mechanizmem sygnalizacyjnym, który pozwala na redukcję asymetrii informacji między pracownikiem a pracodawcą. Świadome kreowanie marki osobistej nie jest już jedynie działaniem wizerunkowym, lecz procesem optymalizacji komunikacji między specjalistą a rynkiem pracy. W warunkach dynamicznie zmieniającego się rynku pracy, postępującej cyfryzacji oraz rosnącej roli modelu gig economy, świadome kształtowanie wizerunku profesjonalnego staje się istotnym elementem budowania przewagi konkurencyjnej [9].

Niniejszy rozdział ma na celu osadzenie koncepcji marki osobistej w realiach dynamicznie zmieniającego się otoczenia gospodarczego oraz wskazanie naukowych fundamentów tego zjawiska w obszarze nauk o zarządzaniu i jakości.

1.1. Ewolucja rynku pracy w kierunku gospodarki cyfrowej

Przejście od modelu gospodarki industrialnej do realiów gospodarki cyfrowej (*Digital Economy*) wymusiło fundamentalną redefinicję sposobu, w jaki jednostki zarządzają swoją ścieżką zawodową. Kluczowym czynnikiem stymulującym te zmiany jest postępująca digitalizacja procesów rekrutacyjnych oraz wzrost znaczenia modelu Gig Economy, gdzie praca oparta jest na krótkoterminowych kontraktach i zadaniach projektowych. W takim schemacie profesjonalista musi funkcjonować jako samodzielny podmiot, którego głównym produktem są unikalne kompetencje i mierzalna zdolność do dostarczania wartości dodanej. W gospodarce cyfrowej tradycyjne bariery wejścia na rynek pracy uległy zatarciu, co z jednej strony zwiększyło mobilność zawodową, a z drugiej – drastycznie podniosło poziom konkurencyjności [18].

Z perspektywy systemowej, rynek pracy stał się środowiskiem o wysokim poziomie szumu informacyjnego. W tym kontekście marka osobista pełni funkcję filtra i nośnika wiarygodności. Cyfrowe ślady, takie jak profil w profesjonalnych serwisach społecznościowych, publikacje branżowe czy udział w projektach merytorycznych, są obecnie poddawane analizie nie tylko przez rekruterów, ale i przez algorytmy automatyzujące procesy selekcji kadr (ATS - Applicant Tracking Systems). Sprawia to, że zarządzanie marką staje się procesem wymagającym precyzyjnego modelowania danych o własnych osiągnięciach i umiejętnościach w przestrzeni wirtualnej [18].

2. Teoretyczne podstawy koncepcji marki osobistej

Zrozumienie mechanizmów budowania marki osobistej wymaga odniesienia się do fundamentów nauk o zarządzaniu oraz marketingu relacyjnego. W literaturze przedmiotu wskazuje się, że marka osobista nie jest jedynie zbiorem cech wizualnych, lecz złożonym systemem obietnic, wartości oraz kompetencji, które generują określoną użyteczność w otoczeniu rynkowym [1].

Niniejszy tekst analizuje proces kształtowania się tego pojęcia oraz jego osadzenie w nowoczesnych strukturach organizacyjnych.

2.1. Geneza i ewolucja pojęcia personal brandingu w literaturze przedmiotu

Początki teoretycznych rozważań nad marką osobistą wywodzą się z nurtu autoprezentacji oraz psychologii społecznej. Kluczowym momentem dla nauk o zarządzaniu było jednak upowszechnienie tezy, że techniki brandingowe, stosowane dotychczas wobec produktów masowych, mogą zostać zaimplementowane na poziomie jednostki. Ewolucja ta przebiegała od prostego modelu autopromocji w latach 90. XX wieku aż po współczesne, holistyczne podejście skoncentrowane na zarządzaniu reputacją [12, 16]. Kluczowe etapy zmian w postrzeganiu tego zjawiska przez badaczy przedmiotu wskazuje tabela 1., prezentująca przejście od perspektywy stricte rynkowej do relacyjnej.

Tabela 1. Ewolucja definicyjna marki osobistej w literaturze zarządzania

Table 1. Evolution of Personal Branding Definitions in Management Literature

Autor / Perspektywa	Główne założenia	Cel procesu
Podejście klasyczne (T. Peters)	Jednostka jako „Ja Sp. z o.o.”	Wyróżnienie się na tle konkurencji
Podejście strategiczne (P. Montoya)	Zarządzanie percepcją otoczenia	Budowanie pozycji lidera opinii
Podejście relacyjne (P. Kotler)	Kreowanie wartości dla odbiorcy	Budowanie długotrwałego zaufania

Źródło: opracowanie własne na podstawie: T. Peters, The Brand Called You, Fast Company, vol. 10, 1997, pp. 83-88, https://personalimpact.ca/wp-content/uploads/2015/08/thebrandcalledyou.pdf (dostęp: 16.02.2026 r.).

Współczesne definicje podkreślają, że marka osobista stanowi sumę doświadczeń, jakie inni mają w kontakcie z daną osobą [3]. Nie jest ona zatem jedynie tym, co jednostka komunikuje na swój temat, lecz przede wszystkim tym, jak jest postrzegana przez pryzmat osiąganych rezultatów oraz prezentowanych postaw.

2.2. Relacja między marką korporacyjną a marką osobistą (analiza porównawcza)

Analiza porównawcza marki osobistej i marki korporacyjnej wskazuje na istnienie silnego sprzężenia zwrotnego między tymi dwoma kategoriami. W tradycyjnym modelu zarządzania to marka organizacji stanowiła swoisty „parasol ochronny" dla pracowników. Obecnie obserwuje się jednak zjawisko częściowego odwrócenia tego mechanizmu – profesjonaliści posiadający silną markę osobistą stają się ambasadorami firm, a ich wiarygodność bezpośrednio wpływa na sposób postrzegania całej organizacji.

Kluczowe różnice oraz punkty wspólne między tymi kategoriami obejmują:

a) własność i kontrolę – o ile marka korporacyjna należy do organizacji, o tyle marka osobista stanowi niezbywalny kapitał jednostki, który pozostaje z nią niezależnie od miejsca zatrudnienia;
b) humanizację biznesu – marki osobiste pracowników pozwalają na humanizację przekazu firmy, co w warunkach ograniczonego zaufania do instytucji ma istotne znaczenie dla skuteczności komunikacji biznesowej;
c) synergię wizerunkową – optymalna sytuacja występuje wówczas, gdy wartości marki osobistej są spójne z wartościami organizacji (tzw. *person–organization fit*). W przeciwnym razie może dochodzić do dysonansu, który w skrajnych przypadkach prowadzi do kryzysu reputacyjnego zarówno po stronie jednostki, jak i organizacji [5].

2.3. Personal branding jako element strategii zarządzania kapitałem ludzkim

Z perspektywy współczesnego zarządzania zasobami ludzkimi (HRM) wspieranie pracowników w budowaniu ich marek osobistych staje się istotnym elementem strategii retencyjnej oraz rozwojowej. Organizacje, które umożliwiają swoim specjalistom budowanie statusu ekspertów funkcjonujących także poza strukturą przedsiębiorstwa, osiągają wyższy poziom zaangażowania pracowników oraz łatwiejszy dostęp do nowych talentów, co wzmacnia działania z zakresu employer brandingu [13].

Włączenie marki osobistej do strategii zarządzania kapitałem ludzkim przejawia się między innymi poprzez:

- promowanie thought leadership, czyli zachęcanie pracowników do dzielenia się wiedzą w mediach społecznościowych oraz podczas konferencji branżowych;
- indywidualizację ścieżek rozwoju, polegającą na dostosowaniu szkoleń i programów rozwojowych do unikalnych cech marki osobistej pracownika;
- wzmacnianie kapitału relacyjnego, poprzez wykorzystanie sieci kontaktów pracowników w realizacji celów biznesowych organizacji.

Takie podejście sprawia, że personal branding przestaje być traktowany jako potencjalne zagrożenie dla organizacji, a zaczyna być postrzegany jako inwestycja w jakość oraz rozpoznawalność kapitału ludzkiego przedsiębiorstwa.

3. Architektura marki osobistej

Skuteczność marki osobistej w procesach rynkowych nie wynika jedynie z intensywności działań promocyjnych, lecz przede wszystkim z jej wewnętrznej struktury. Z perspektywy nauk o zarządzaniu markę osobistą można postrzegać jako wielowymiarową architekturę, w której poszczególne elementy składowe muszą tworzyć spójną całość.

W niniejszym rozdziale pracy dokonano dekompozycji marki na jej fundamentalne czynniki: tożsamość opartą na wartościach, zasoby kompetencyjne oraz mechanizmy komunikacyjne.

3.1. Autentyczność i system wartości jako fundament tożsamości profesjonalnej

W literaturze przedmiotu coraz silniej akcentuje się, że autentyczność stanowi warunek konieczny trwałości marki osobistej. W przeciwieństwie do marek produktowych, które mogą być kreowane w sposób całkowicie sztuczny, marka osobista jest nierozerwalnie związana z realną osobowością jednostki [4].

Autentyczność definiowana jest jako stopień zbieżności między tożsamością wewnętrzną (tym, kim dana osoba jest) a wizerunkiem zewnętrznym (tym, jak jest postrzegana). Fundamentem tej architektury jest system wartości, który determinuje cele zawodowe, sposób podejmowania decyzji oraz etykę pracy [14].

Z punktu widzenia zarządzania karierą jasne zdefiniowanie wartości pozwala na selekcję ofert zawodowych poprzez wybór organizacji, których kultura organizacyjna pozostaje zbieżna z wartościami jednostki (person–organization fit). Ponadto spójność wartości sprzyja budowaniu wiarygodności zawodowej. Odbiorcy – zarówno rekruterzy, jak i klienci – intuicyjnie dostrzegają brak autentyczności w sytuacji, gdy deklarowane wartości nie znajdują odzwierciedlenia w działaniach. Budowanie marki osobistej na fundamencie sprzecznym z własnymi przekonaniami może natomiast prowadzić do spadku motywacji, wypalenia zawodowego oraz kryzysów wizerunkowych.

3.2. Kompetencje twarde i miękkie w strukturze zasobów marki

Jeśli autentyczność stanowi fundament marki osobistej, to kompetencje są „produktem", który marka oferuje rynkowi. W architekturze marki osobistej kluczowe jest rozróżnienie oraz odpowiednie zbalansowanie dwóch rodzajów zasobów:

a) kompetencje twarde (hard skills) – są to specyficzne, mierzalne umiejętności techniczne, takie jak znajomość metodologii zarządzania projek-

tami, obsługa systemów ERP czy analiza danych. W strukturze marki pełnią one funkcję dowodu eksperckości (*reason to believe*). Bez solidnego zaplecza twardych kompetencji marka osobista staje się jedynie „wydmuszką" marketingową;

b) kompetencje miękkie (soft skills) – obejmują zdolności interpersonalne, inteligencję emocjonalną, krytyczne myślenie oraz umiejętność adaptacji do zmieniających się warunków. To właśnie one decydują o tym, w jaki sposób wiedza techniczna jest wykorzystywana w praktyce biznesowej [17].

Współczesny rynek pracy premiuje model kompetencji typu T-shaped**,** w którym pionowe ramię litery T symbolizuje głęboką wiedzę specjalistyczną, natomiast ramię poziome oznacza szerokie kompetencje miękkie oraz zdolność do współpracy z przedstawicielami innych dziedzin [10].

Silna marka osobista powinna precyzyjnie komunikować oba te wymiary, wskazując na unikalną kombinację umiejętności, która stanowi o przewadze konkurencyjnej jednostki.

3.3. Rola spójności komunikacyjnej w budowaniu zaufania rynkowego

Ostatnim elementem architektury marki osobistej jest spójność komunikacyjna (*consistency*). Z perspektywy zarządzania informacją oznacza ona dostarczanie zbieżnych komunikatów we wszystkich punktach kontaktu z rynkiem [4].

Przykładowo, jeśli inżynier procesów deklaruje w serwisie LinkedIn dbałość o detale i profesjonalizm, a jego komunikacja mailowa pozostaje chaotyczna i niespójna, dochodzi do erozji zaufania odbiorców.

Spójność komunikacyjna może być analizowana w trzech wymiarach:

1) Spójność horyzontalna – zgodność komunikatów w różnych mediach i kanałach komunikacji (np. LinkedIn, Twitter, wystąpienia publiczne);
2) Spójność wertykalna – zgodność deklaracji z rzeczywistymi działaniami oraz wynikami pracy;
3) Spójność czasowa – utrzymywanie stabilnego poziomu jakości komunikacji i obecności w przestrzeni publicznej, co umożliwia budowanie trwałej reputacji, a nie jedynie chwilowej rozpoznawalności.

Budowanie zaufania rynkowego ma charakter kumulatywny. Każdy spójny komunikat wzmacnia kapitał zaufania, który w sytuacjach kryzysowych pełni funkcję bufora bezpieczeństwa dla marki osobistej.

4. Wielokanałowość w procesie kreowania wizerunku zawodowego

Skuteczna marka osobista we współczesnych warunkach rynkowych powinna opierać się na strategii omnichannel (wielokanałowości). Oznacza to, że proces budowania wizerunku nie może ograniczać się do jednego medium, lecz powinien

tworzyć spójną sieć punktów styku (*touchpoints*) między jednostką a jej otoczeniem zawodowym.

W niniejszym rozdziale przeanalizowano synergię między narzędziami cyfrowymi a aktywnością w świecie rzeczywistym.

4.1. Wykorzystanie mediów społecznościowych i platform profesjonalnych

W warunkach gospodarki cyfrowej serwisy społecznościowe o charakterze zawodowym stały się podstawową bazą danych o kapitale ludzkim. Platforma LinkedIn, będąca liderem w tym segmencie, przestała pełnić funkcję jedynie „cyfrowego życiorysu", ewoluując w kierunku zaawansowanego narzędzia do budowania autorytetu eksperckiego oraz social sellingu własnych kompetencji [1, 7].

Z perspektywy zarządzania wizerunkiem kluczowe aspekty wykorzystania tych platform obejmują:

- optymalizację pod kątem algorytmów (SEO personal branding) – dobór odpowiednich słów kluczowych w nagłówku oraz sekcji „O mnie" wpływa na widoczność profilu w wynikach wyszukiwania rekruterów i potencjalnych kontrahentów;
- społeczny dowód słuszności (*social proof*) – funkcje rekomendacji oraz potwierdzania umiejętności stanowią zewnętrzną walidację deklarowanych kompetencji, co znacząco obniża poziom niepewności odbiorców marki;
- aktywność merytoryczną – regularne publikowanie aktualizacji branżowych oraz komentarzy eksperckich pozycjonuje jednostkę jako osobę aktywnie śledzącą zmiany rynkowe.

Pomimo marginalizacji serwisów lokalnych, takich jak GoldenLine, na rzecz globalnych platform społecznościowych, wybór odpowiedniego kanału komunikacji powinien być uzależniony od specyfiki branży oraz grupy docelowej, do której marka osobista chce docierać.

4.2. Strategie content marketingu i storytellingu w budowaniu autorytetu eksperckiego

Samo posiadanie profilu w mediach społecznościowych jest niewystarczające do zbudowania silnej marki osobistej. Niezbędne jest wdrożenie strategii content marketingu, polegającej na dostarczaniu wartościowych treści, które rozwiązują problemy odbiorców lub edukują rynek [2].

Kluczowym narzędziem w tym obszarze jest storytelling, czyli narracyjne przedstawianie doświadczeń zawodowych. Inżynier czy menedżer nie powinien ograniczać się jedynie do wymieniania projektów, w których uczestniczył, lecz przedstawiać proces rozwiązywania problemów oraz kontekst podejmowanych decyzji.

Skuteczna struktura storytellingu w personal brandingu obejmuje:

- kontekst – z jakim wyzwaniem biznesowym się mierzono;

- działanie – jakie konkretne kroki podjęto (zastosowane metody, narzędzia);
- wynik – jaki wymierny efekt osiągnięto (np. wzrost wydajności, oszczędności kosztowe).

Takie podejście buduje autorytet ekspercki znacznie skuteczniej niż prosta lista faktów, ponieważ pozwala odbiorcy zrozumieć sposób myślenia oraz warsztat pracy danej osoby.

4.3. Znaczenie networkingu i relacji interpersonalnych w środowisku offline

Pomimo dominacji kanałów cyfrowych marka osobista w dalszym ciągu weryfikowana jest w bezpośrednich kontaktach interpersonalnych. Środowisko offline – obejmujące konferencje branżowe, warsztaty czy projekty międzyorganizacyjne – pozostaje kluczowe dla budowania relacji opartych na zaufaniu, których nie da się w pełni odtworzyć w przestrzeni wirtualnej [6].

Networking w kontekście rozwoju kariery można rozumieć jako:

- zarządzanie siecią kontaktów – świadome budowanie relacji z osobami o podobnych zainteresowaniach zawodowych oraz z mentorami;
- wymianę wartości – skuteczny networking nie polega na „zbieraniu wizytówek", lecz na wzajemnym wspieraniu się w rozwiązywaniu problemów zawodowych;
- udział w projektach studenckich i branżowych – dla osób wchodzących na rynek pracy (np. studentów uczelni technicznych) stanowi to jedną z najszybszych dróg zdobycia tzw. doświadczenia relacyjnego, które wzmacnia wiarygodność marki w oczach pierwszych pracodawców.

Podsumowując, nowoczesna architektura komunikacji marki osobistej opiera się na cyklu: zauważenie w sieci (online) – weryfikacja kompetencji poprzez treści (content) – pogłębienie relacji i budowanie zaufania (offline).

5. Wpływ personal brandingu na trajektorię kariery zawodowej

Współczesne modele kariery zawodowej charakteryzują się wysokim stopniem nieliniowości. Marka osobista pełni w tym kontekście rolę stabilizatora, który pozwala jednostce na sprawne poruszanie się między różnymi organizacjami, branżami czy rolami zawodowymi.

W niniejszym podrozdziale pracy przeanalizowano, w jaki sposób świadome zarządzanie wizerunkiem przekłada się na realne wskaźniki sukcesu na rynku pracy.

5.1. Marka osobista a wzrost wskaźnika zatrudnialności (employability)

Pojęcie zatrudnialności (*employability*) definiowane jest w naukach o zarządzaniu jako zdolność jednostki do zdobycia i utrzymania pracy, a także do uzyskania nowego zatrudnienia w przypadku zmiany sytuacji zawodowej.

Silna marka osobista znacząco podnosi ten wskaźnik poprzez redukcję tzw. ryzyka rekrutacyjnego. Z perspektywy pracodawcy kandydat posiadający rozpoznawalną markę osobistą jest w pewnym sensie „zweryfikowany" przez rynek. Dowody eksperckości dostępne w przestrzeni cyfrowej (np. publikacje, case studies czy wystąpienia branżowe) pełnią funkcję wstępnej selekcji kompetencyjnej [15]. W rezultacie:

- skraca się czas poszukiwania zatrudnienia – osoby posiadające silną markę osobistą często otrzymują propozycje współpracy w ramach tzw. rekrutacji odwróconej, w której to organizacje aktywnie zabiegają o specjalistę;
- wzrasta elastyczność zawodowa – marka oparta na jasno zdefiniowanych wartościach oraz uniwersalnych kompetencjach miękkich ułatwia tzw. *pivoting*, czyli zmianę branży lub specjalizacji przy zachowaniu dotychczasowego statusu zawodowego [15].

5.2. Budowanie przewagi konkurencyjnej na nowoczesnym rynku pracy

W warunkach dużej podaży specjalistów o zbliżonych kompetencjach twardych marka osobista staje się kluczowym czynnikiem różnicującym (*differentiator*). Przewaga konkurencyjna budowana poprzez personal branding opiera się na jasno zdefiniowanej unikalnej propozycji wartości (Unique Value Proposition) [11].

Jednostka, która potrafi precyzyjnie zakomunikować, jaki problem biznesowy rozwiązuje oraz w jaki sposób to robi, przestaje być postrzegana wyłącznie przez pryzmat kosztu zatrudnienia (wynagrodzenia), a zaczyna być traktowana jako inwestycja w rozwój organizacji.

Silna marka osobista pozwala również na wyjście z tzw. pułapki towarowej (*commoditization*), w której pracownicy są postrzegani jako zasoby łatwo zastępowalne. Przejście do kategorii unikalnych talentów przekłada się bezpośrednio na większą stabilność zatrudnienia oraz szybsze tempo awansów zawodowych.

5.3. Personal branding jako narzędzie zwiększania kapitału społecznego jednostki

Kapitał społeczny, rozumiany jako sieć relacji oraz wynikające z niej zasoby takie jak informacje, zaufanie czy wsparcie stanowi jedno z najcenniejszych aktywów w rozwoju kariery zawodowej [1].

Marka osobista działa w tym kontekście jak swoisty magnes – przyciąga osoby o podobnych celach i wartościach, ułatwiając budowanie wysokiej jakości sieci kontaktów.

Wpływ marki osobistej na rozwój kapitału społecznego przejawia się między innymi poprzez m.in.:

- dostęp do tzw. ukrytego rynku pracy - znacząca część wysokospecjalistycznych stanowisk obsadzana jest poprzez rekomendacje i sieci kontaktów,

zanim ogłoszenie zostanie opublikowane w oficjalnych kanałach rekrutacyjnych. Silna marka sprawia, że jednostka znajduje się w kręgu zainteresowania decydentów;

- efekt autorytetu - osoby posiadające ugruntowaną markę ekspercką łatwiej nawiązują relacje z mentorami oraz liderami opinii, co przyspiesza transfer wiedzy i otwiera dostęp do prestiżowych projektów;
- budowanie ekosystemu współpracy – marka osobista sprzyja przechodzeniu od relacji czysto transakcyjnych (praca za wynagrodzenie) do relacji partnerskich i wspólnotowych w ramach danej branży lub środowiska zawodowego.

6. Ekonomiczne i psychologiczne korzyści posiadania silnej marki

Efektywne zarządzanie marką osobistą generuje szereg korzyści, które można sklasyfikować w dwóch głównych wymiarach: obiektywnym (ekonomicznym) oraz subiektywnym (psychologicznym). W naukach o zarządzaniu podkreśla się, że silna marka osobista redukuje koszty transakcyjne na rynku pracy oraz zwiększa efektywność operacyjną jednostki jako „dostawcy" usług profesjonalnych.

6.1. Weryfikowalność kompetencji a wzrost rozpoznawalności i wiarygodności rynkowej

W tradycyjnym modelu rekrutacji weryfikacja kompetencji kandydata jest procesem kosztownym i obarczonym wysokim ryzykiem błędu. Silna marka osobista zmienia tę sytuację poprzez dostarczenie tzw. dowodów społecznych i merytorycznych jeszcze przed formalnym spotkaniem rekrutacyjnym [9, 15].

Rozpoznawalność w środowisku branżowym sprawia, że proces budowania zaufania zachodzi znacznie szybciej. Marka działa w tym kontekście jak swoisty certyfikat jakości, który może być weryfikowany poprzez publikacje, wystąpienia branżowe czy rekomendacje innych ekspertów.

Osoba posiadająca ugruntowaną reputację rzadziej spotyka się z koniecznością ciągłego udowadniania swoich kompetencji, co pozwala na szybsze przechodzenie do realizacji zadań strategicznych zamiast koncentrowania się na potwierdzaniu własnej wartości zawodowej.

6.2. Wpływ wizerunku na negocjacje płacowe i dostęp do ukrytego rynku pracy

Z ekonomicznego punktu widzenia marka osobista umożliwia uzyskanie tzw. premii za markę. Zjawisko to, znane z rynku dóbr i usług, znajduje swoje bezpośrednie odzwierciedlenie także na rynku pracy [5].

Specjaliści posiadający silną markę osobistą rzadziej konkurują poziomem wynagrodzenia, a częściej unikalną wartością, jaką wnoszą do organizacji. Pozwala to na negocjowanie wyższych stawek, ponieważ pracodawca postrzega taką osobę jako inwestycję obarczoną mniejszym ryzykiem błędu.

Ponadto wiele ofert pracy na poziomie eksperckim lub doradczym nigdy nie trafia do publicznego obiegu. Marka osobista umożliwia dostęp do tego tzw. ukrytego rynku pracy, w którym warunki finansowe oraz możliwości rozwoju zawodowego są często znacznie korzystniejsze.

6.3. Wymiar psychologiczny: pewność siebie i odporność zawodowa

Poza aspektem ekonomicznym marka osobista silnie oddziałuje również na dobrostan psychiczny profesjonalisty. Świadomość własnej wartości rynkowej oraz posiadanie spójnej tożsamości zawodowej przekładają się na:

a) wzrost poczucia sprawstwa – jednostka ma poczucie, że to ona zarządza swoją karierą zawodową, a nie jest jedynie przedmiotem decyzji kadrowych organizacji;
b) odporność na stres (*resilience*) – w sytuacjach kryzysowych, takich jak utrata pracy czy restrukturyzacja przedsiębiorstwa, osoba posiadająca silną markę osobistą wykazuje wyższy poziom bezpieczeństwa psychicznego, wynikający z przekonania, że jej reputacja umożliwi stosunkowo szybki powrót na rynek pracy;
c) redukcję dysonansu poznawczego – budowanie marki osobistej w oparciu o autentyczność (o czym wspomniano w podrozdziale 3.1. niniejszego rozdziału) sprzyja zachowaniu spójności między życiem zawodowym a prywatnym, co stanowi fundament długofalowej satysfakcji zawodowej.

7. Bariery, dylematy etyczne i ryzyka w procesie autopromocji

Pomimo licznych korzyści proces budowania marki osobistej generuje również szereg ryzyk, które mogą negatywnie oddziaływać zarówno na jednostkę, jak i na jej otoczenie zawodowe. W naukach o zarządzaniu podkreśla się, że niewłaściwie prowadzony personal branding może prowadzić do erozji kapitału zaufania i stać się barierą w rozwoju kariery zamiast czynnikiem ją stymulującym.

7.1. Pułapka niespójności wizerunkowej i kryzysy reputacyjne w sieci

W warunkach gospodarki cyfrowej każda informacja pozostawia w sieci trwały ślad (zjawisko tzw. digital footprint). Jednym z największych zagrożeń dla marki osobistej jest dysonans pomiędzy deklarowanym wizerunkiem a rzeczywistymi działaniami jednostki. Szczególnie istotna jest erozja wiarygodności. Jeżeli profesjonalista promuje wartości takie jak rzetelność czy terminowość, a jego działania projektowe wskazują na brak tych cech, dochodzi do gwałtownego spadku zaufania. Naprawa nadszarpniętej reputacji w środowisku cyfrowym jest procesem znacznie trudniejszym i bardziej kosztownym niż budowanie jej od podstaw [8].

Istotnym czynnikiem jest także efekt skali w sytuacjach kryzysowych. Media społecznościowe sprawiają, że błąd popełniony przez jednostkę może w bardzo krótkim czasie osiągnąć zasięg globalny, prowadząc do zjawisk takich jak cancel culture lub trwałe wykluczenie z określonych środowisk branżowych [8].

7.2. Granica między profesjonalizmem a nadmierną autopromocją

Częstym błędem w procesie personal brandingu jest przesunięcie środka ciężkości z merytoryki na samą formę przekazu. Może to prowadzić do zjawiska określanego w literaturze jako narcyzm zawodowy [13, s. 12-50].

Granica między budowaniem autorytetu eksperckiego a bezproduktywną autopromocją jest bardzo cienka. Nadmierna autoprezentacja, pozbawiona wartości merytorycznej dla odbiorcy, prowadzi do wizerunku osoby skoncentrowanej przede wszystkim na autopromocji, co może zniechęcać potencjalnych partnerów biznesowych oraz rekruterów.

Warto również zwrócić uwagę na zjawisko *imposter syndrome* (syndromu oszusta). Presja budowania idealnego wizerunku w przestrzeni cyfrowej może prowadzić do problemów psychologicznych. Jednostka, starając się sprostać wykreowanemu przez siebie wizerunkowi eksperta, może odczuwać lęk przed zdemaskowaniem swoich rzeczywistych – często wystarczających, lecz nie idealnych – kompetencji.

7.3. Ryzyka związane z cyfrowym śladem i ochroną prywatności

Personal branding wymaga od jednostki wysokiego poziomu ekspozycji publicznej, co niesie ze sobą określone zagrożenia w obszarze prywatności oraz bezpieczeństwa danych.

Jednym z nich jest zacieranie granic między sferą prywatną a zawodową. W środowisku mediów społecznościowych utrzymanie pełnej separacji tych obszarów staje się coraz trudniejsze [8]. Nieprzemyślane treści prywatne mogą negatywnie wpływać na postrzeganie profesjonalizmu jednostki. Z drugiej strony nadmierne skupienie na budowaniu marki zawodowej może prowadzić do wypalenia oraz utraty autentyczności w relacjach osobistych.

Kolejnym zagrożeniem jest profilowanie algorytmiczne. Dane udostępniane w procesie budowania marki osobistej są przetwarzane przez algorytmy platform cyfrowych, co może prowadzić do automatycznego profilowania użytkowników. W niektórych przypadkach może to wpływać na procesy selekcji kandydatów lub sposób, w jaki dana osoba jest prezentowana w systemach rekrutacyjnych [7].

Podsumowując, świadome zarządzanie marką osobistą wymaga zachowania etycznego kompasu oraz stałej autorefleksji. Marka osobista powinna być narzędziem wzmacniającym kompetencje jednostki, a nie ich substytutem.

8. Podsumowanie

Przeprowadzona w niniejszym akapicie analiza pozwala na sformułowanie wniosku, że marka osobista przestała być jedynie domeną osób publicznych czy kadry zarządzającej najwyższego szczebla, stając się powszechnym i niezbędnym narzędziem zarządzania karierą w gospodarce cyfrowej. W warunkach nadpodaży informacji oraz postępującej automatyzacji procesów rekrutacyjnych świadome kształtowanie wizerunku profesjonalnego stanowi fundament budowania trwałej przewagi konkurencyjnej.

W toku przeprowadzonych rozważań wykazano, że:

- marka osobista jest aktywem strategicznym – stanowi niematerialny zasób jednostki, który bezpośrednio wpływa na jej rynkową wartość oraz odporność na turbulencje rynkowe;
- autentyczność determinuje trwałość marki – skuteczny personal branding nie może być procesem czysto marketingowym, lecz powinien opierać się na realnych kompetencjach (fundamencie merytorycznym) oraz spójnym systemie wartości;
- wielokanałowość komunikacji jest koniecznością – efektywne budowanie autorytetu wymaga umiejętnego łączenia aktywności w mediach społecznościowych (LinkedIn, content marketing) z bezpośrednimi relacjami interpersonalnymi (networking offline).

Na podstawie przeprowadzonej analizy sformułowano następujące rekomendacje dla profesjonalistów oraz studentów:

a) audyt cyfrowego śladu – konieczne jest regularne monitorowanie i optymalizowanie własnej obecności w sieci pod kątem spójności z obranym celem zawodowym;
b) inwestycja w kompetencje typu T-shaped – należy dążyć do równowagi między pogłębianiem wiedzy specjalistycznej a rozwijaniem kompetencji komunikacyjnych, które umożliwiają efektywny transfer tej wiedzy do otoczenia;
c) etyka komunikacji – budowanie marki osobistej powinno opierać się na dostarczaniu realnej wartości odbiorcom, co pozwala uniknąć pułapki nadmiernej autopromocji i narcyzmu zawodowego.

Niniejsze zagadnienie, ze względu na ograniczenia objętościowe, stanowi jedynie zarys szerokiej problematyki personal brandingu. Główne ograniczenie pracy wynika z dynamicznego tempa zmian technologicznych – algorytmy platform profesjonalnych ewoluują często szybciej niż literatura przedmiotu.

Wskazuje to na potrzebę dalszych badań naukowych, w szczególności w takich obszarach jak:

- wpływ generatywnej sztucznej inteligencji (AI) na proces tworzenia treści oraz autentyczność marek osobistych;

- analiza ilościowa korelacji między siłą marki osobistej a realnym poziomem wynagrodzeń w poszczególnych sektorach gospodarki (np. IT czy inżynierii);
- badania nad trwałością marek osobistych w warunkach długofalowych kryzysów gospodarczych.

Podsumowując, marka osobista w XXI wieku stanowi proces ciągły, wymagający nie tylko wiedzy marketingowej, lecz przede wszystkim wysokiego poziomu samoświadomości oraz konsekwencji w komunikacji. Jest to inwestycja, która w długim okresie może przynieść znaczącą stopę zwrotu w postaci stabilnej i satysfakcjonującej ścieżki zawodowej.

Literatura

[1] American Marketing Association: Personal Brand, https://www.ama.org/topics/personal-brand/ (dostęp: 16.02.2026 r.).

[2] Arruda W., Dixon K.: Career Distinction: Stand Out by Building Your Brand, John Wiley & Sons, Hoboken 2007.

[3] Bendisch, F., Larsen, G., Trueman, M.: Fame and fortune: A conceptual model of CEO brands, European Journal of Marketing, vol 47, no 3/4, 2013, pp. 596-614, https://www.researchgate.net/publication/263146907_Fame_and_fortune_a_conceptual_model_of_CEO_brands (dostęp: 16.02.2026 r.).

[4] Goffman, E.: Człowiek w teatrze życia codziennego, Wydawnictwo Aletheia, Warszawa 2021.

[5] Kotler P., Keller K. L.: Marketing Management, Pearson, Boston 2016.

[6] Kotler P., Keller K. L.: Marketing, Rebis, Poznań 2016.

[7] Labrecque L. I., Markos E., Milne G. R.: Online personal branding: Processes, challenges, and implications, Journal of Interactive Marketing, vol 25, no 1, 2011, pp. 37-50, https://www.researchgate.net/publication/251645346_Online_Personal_Branding_Processes_Challenges_and_Implications (dostęp: 16.02.2026 r.).

[8] Lair D. J., Sullivan K., Cheney G.: Marketization and the Recasting of the Professional Self. The Rhetoric and Ethics of Personal Branding, Management Communication Quarterly, vol 18, no 3, 2005, pp. 307-343, https://www.researchgate.net/publication/258170724_Marketization_and_the_Recasting_of_the_Professional_Self_The_Rhetoric_and_Ethics_of_Personal_Branding (dostęp: 16.02.2026 r.).

[9] LinkedIn: The Future of Recruiting 2025. How AI redefines recruiting excellence, https://business.linkedin.com/hire/resources/future-of-recruiting (dostęp: 16.02.2026 r.).

[10] McKinsey & Company: The digital future of work: What skills will employees need?, https://www.mckinsey.com/featured-insights/future-of-work/the-digital-future-of-work-what-skills-will-be-needed (dostęp: 16.02.2026 r.).

[11] Montoya P., Vandehey T.: The Brand Called You: Create a Personal Brand that Wins Your Business and Raises Your Income, McGraw-Hill, New York 2008.

[12] Peters T.: The Brand Called You, Fast Company, vol. 10, 1997, pp. 83-88, https://personalimpact.ca/wp-content/uploads/2015/08/thebrandcalledyou.pdf (dostęp: 16.02.2026 r.).

[13] Peters T.: The Brand You 50: Or: Fifty Ways to Transform Yourself from an "Employee" into a Brand That Shouts Distinction, Commitment, and Passion! Knopf, New York 1999.

[14] Rampersad H. K.: Authentic Personal Branding: A New Blueprint for Building and Transforming Your Brand, Information Age Publishing, Charlotte, NC 2008.

[15] Schawbel D.: Personal Branding 2.0. Cztery kroki do zbudowania osobistej marki, Wydawnictwo Helion, Gliwice 2012.

[16] Shepherd I. D., From Cattle and Coke to Charlie: Meeting the Challenge of Self Marketing and Personal Branding, Journal of Marketing Management, vol 21, no 5, 2010, pp. 589--606, https://www.researchgate.net/publication/261573431_From_Cattle_and_Coke_to_Charlie_Meeting_the_Challenge_of_Self_Marketing_and_Personal_Branding (dostęp: 16.02.2026 r.).

[17] Vallas, S. P., Cummins, E. R.: Personal Branding and Identity Norms in the Popular Business Press: Enterprise Culture in an Age of Precarity, Organization Studies, vol 36, no 3, 2015, pp. 293-319, https://journals.sagepub.com/doi/10.1177/0170840614563741 (dostęp: 16.02.2026 r.).

[18] World Economic Forum: The Future of Jobs Report 2023, https://www.weforum.org/publications/the-future-of-jobs-report-2023/ (dostęp: 16.02.2026 r.).

MY BRAND, MY STRENGTH: THE SIGNIFICANCE OF PERSONAL BRANDING IN PROFESSIONAL CAREER DEVELOPMENT

Summary

Modern management science indicates that personal branding is one of the key tools for the strategic management of an individual's professional career. In the conditions of a dynamically changing labor market, progressive digitalization, and the growing role of the gig economy model, the conscious shaping of a professional image becomes an essential element of building a competitive advantage. The aim of this article is a multifaceted presentation of the essence of personal branding, the identification of its key determinants, and a critical analysis of the benefits and risks associated with the process of its creation. The first part of the study introduces the definition of personal branding and its genesis, embedding it in the context of relationship marketing and modern human resource management. Subsequently, the brand architecture is decomposed, discussing its fundamental components: authenticity, value systems, T-shaped competencies, and communication consistency. The next section focuses on the operationalization of branding activities, analyzing the effectiveness of online tools (including the role of LinkedIn and storytelling) and the importance of offline activities, such as networking and participation in industry projects. Furthermore, the article demonstrates the impact of a strong brand on the trajectory of professional development, with particular emphasis on the increase in the employability rate and the accumulation of social capital. The author also points out potential challenges and barriers, including the risk of image inconsistency, reputation crises in the digital space, and ethical dilemmas related to excessive self-promotion. The summary of the article emphasizes that building a personal brand is a continuous and long-term process, requiring not only technological proficiency but, above all, strategic reflection and an ethical approach to communication. Perspectives for further research are also indicated, considering the impact of artificial intelligence on the authenticity of the professional image.

Keywords: self-presentation, career development, marketing communication, image, networking.

Zuzanna Dörfler, Ewelina Nycz
Studenckie Koło Naukowe Komunikacji Marketingowej „Brief"

INFLUENCER MARKETING JAKO NARZĘDZIE KOMUNIKACJI MARKETINGOWEJ I JEGO WPŁYW NA ZACHOWANIA KONSUMENTÓW

Streszczenie

Współczesne przedsiębiorstwa funkcjonują w dynamicznym i silnie nasyconym informacyjnie środowisku, w którym tradycyjne formy promocji stopniowo ustępują miejsca interaktywnym mediom społecznościowym. Rozwój technologii cyfrowych oraz spadek zaufania do reklamy masowej przyczyniły się do wzrostu znaczenia influencer marketingu jako istotnego narzędzia komunikacji marketingowej. Umożliwia on markom budowanie bezpośrednich i spersonalizowanych relacji z konsumentami w ramach systemu zintegrowanej komunikacji marketingowej (IMC). Celem rozdziału jest analiza influencer marketingu jako instrumentu kształtowania postaw oraz decyzji zakupowych konsumentów, ze szczególnym uwzględnieniem specyfiki pokolenia Z oraz milenialsów. W rozdziale omówiono kluczowe mechanizmy psychologiczne i społeczne determinujące skuteczność tego narzędzia, takie jak społeczny dowód słuszności, identyfikacja społeczna oraz efekt autorytetu. Podkreślono także znaczenie autentyczności, wiarygodności i spójności wizerunkowej twórców internetowych w procesie budowania zaufania odbiorców.
W części teoretycznej zdefiniowano istotę influencer marketingu oraz przedstawiono klasyfikację liderów opinii, obejmującą nano-, mikro-, makro- oraz mega-influencerów. Część empiryczna opiera się na metodzie studium przypadku (case study) i obejmuje analizę wybranych kampanii marketingowych realizowanych przez marki Deichmann oraz Samsung. Przeprowadzona analiza pozwoliła wskazać czynniki determinujące skuteczność działań influencer marketingowych, w szczególności znaczenie dopasowania influencera do wartości i wizerunku marki. W rozdziale zwrócono również uwagę na wyzwania związane z dalszym rozwojem tej formy komunikacji marketingowej, w tym konieczność profesjonalizacji branży, zachowania transparentności współprac oraz dostosowania działań marketingowych do obowiązujących regulacji prawnych dotyczących oznaczania treści sponsorowanych.

Słowa kluczowe: strategie marketingowe, media społecznościowe, liderzy opinii, autentyczność marki.

1. Geneza i istota influencer marketingu

1.1. Ewolucja komunikacji marketingowej w dobie cyfryzacji

Tradycyjny model komunikacji marketingowej, oparty na paradygmacie jednostronnego przekazu (nadawca/marka – odbiorca/konsument), uległ w ostatniej dekadzie nieodwracalnej dezaktualizacji. Rozwój technologii Web 2.0, a w szczególności dynamiczna ekspansja mediów społecznościowych, wymusiły przejście od monologu korporacyjnego w stronę wielokierunkowego dialogu i interakcji [18].

W tym nowym ekosystemie konsument przestał pełnić rolę jedynie biernego odbiorcy treści, stając się „prosumentem" – jednostką, która aktywnie współtworzy wizerunek marek, poszukuje niezależnych opinii oraz rygorystycznie weryfikuje komunikaty płynące z oficjalnych kanałów firmowych [24].

W tym kontekście influencer marketing wyłonił się jako naturalna ewolucja marketingu szeptanego (word-of-mouth), przeniesiona w sferę cyfrową i wzmocniona przez algorytmy platform społecznościowych. Stanowi on odpowiedź na narastające zjawisko „ślepoty bannerowej" (banner blindness) oraz powszechny spadek zaufania do tradycyjnych formatów reklamowych.

1.2. Definicja i ramy teoretyczne influencer marketingu

Influencer marketing definiowany jest jako strategiczna forma marketingu relacji, polegająca na wykorzystaniu wpływu osób posiadających znaczący autorytet, wiedzę lub kapitał społeczny w danej społeczności w celu promocji produktów, usług bądź marek [15]. Influencerzy, pełniący rolę cyfrowych liderów opinii, stają się kluczowymi pośrednikami w dwustopniowym przepływie informacji. Dzięki budowaniu relacji o charakterze parasocjalnym – czyli złudzeniu bliskości i osobistej więzi między twórcą a tysiącami odbiorców – influencer marketing skutecznie skraca dystans psychologiczny między marką a potencjalnym nabywcą [20].

W literaturze naukowej zjawisko to postrzegane jest jako istotny komponent zintegrowanej komunikacji marketingowej (IMC). Jego fundamentem nie jest wyłącznie techniczny zasięg, lecz przede wszystkim:

a) zaufanie i wiarygodność – odbiorcy traktują treści publikowane przez influencerów nie jako agresywną reklamę, lecz jako osobistą rekomendację lub poradę ekspercką,
b) autentyczność – postrzegana spójność życia prywatnego twórcy z promowanymi przez niego wartościami, co nadaje komunikatom unikalny, ludzki wymiar,
c) relacje interpersonalne – wykorzystanie mechanizmów identyfikacji społecznej, w których konsument poprzez zakup produktu dąży do upodobnienia się do wzorca osobowego prezentowanego przez lidera opinii.

W przeciwieństwie do tradycyjnej reklamy influencer marketing operuje na płaszczyźnie wartości dodanej – treści marketingowe są tu wplecione w kontekst

rozrywkowy, edukacyjny lub lifestylowy, co znacząco podnosi ich skuteczność i przyswajalność przez współczesny rynek [3].

2. Klasyfikacja influencerów i ich charakterystyka techniczna

Współczesna taksonomia twórców internetowych wykracza poza prosty podział liczbowy, stanowiąc fundament inżynierii procesów marketingowych. Kluczowa w tym kontekście jest segmentacja oparta nie tylko na skali oddziaływania, lecz przede wszystkim na parametrach efektywności, takich jak wskaźnik ER (Engagement Rate), CPM (Cost Per Mille) oraz potencjał konwersji. Wybór odpowiedniej kategorii influencera determinuje realizację konkretnych celów strategicznych przedsiębiorstwa – od budowania zasięgów masowych po precyzyjne dotarcie do nisz rynkowych. Na szczycie tej struktury znajdują się mega-influencerzy oraz celebryci, posiadający powyżej 1 miliona obserwujących. W tej grupie dominują globalne gwiazdy, sportowcy i pionierzy platform społecznościowych, którzy pełnią funkcję „cyfrowych billboardów". Ich strategiczna rola opiera się na budowaniu masowej świadomości marki, co czyni ich niezastąpionymi w fazie launchu produktów. Choć kampanie te charakteryzują się najniższym jednostkowym kosztem dotarcia (niski CPM), obarczone są zazwyczaj najniższym wskaźnikiem zaangażowania procentowego, a ich przekaz bywa postrzegany jako mniej autentyczny ze względu na jawnie komercyjny charakter [11].

Kolejny segment stanowią makro-influencerzy (100 tys. – 1 mln obserwujących), którzy wprowadzają do ekosystemu marketingu profesjonalizację i wysoką jakość przekazu. Często funkcjonują oni jako profesjonalne podmioty gospodarcze z własnym zapleczem produkcyjnym, co gwarantuje marce bezpieczeństwo wizerunkowe (brand safety) oraz wysokie standardy techniczne, np. w zakresie produkcji wideo czy jakości dźwięku. Stanowią oni idealny kompromis między znacznym zasięgiem a wiarygodnością ekspercką, potrafiąc przełożyć skomplikowane cechy produktu na język korzyści zrozumiały dla szerokiego grona odbiorców [1].

W procesach zorientowanych na efektywność i zwrot z inwestycji (ROI) kluczową rolę odgrywają mikro-influencerzy (10 tys. – 100 tys. obserwujących). Jako architekci zaangażowania skupiają się oni wokół konkretnych pasji i wyspecjalizowanych obszarów tematycznych, takich jak np. kosmetyki wegańskie czy nowe technologie. Ich rekomendacje nie są traktowane jak klasyczna reklama, lecz jak porada ekspercka, co przekłada się na silną relację z odbiorcami i wysoki poziom zaufania. W tej grupie obserwuje się także znaczący wzrost konwersji bezpośredniej, co czyni ich fundamentem kampanii sprzedażowych [7].

Dopełnieniem klasyfikacji są nano-influencerzy, posiadający mniej niż 10 tys. obserwujących. Choć dysponują najmniejszą skalą oddziaływania, ich przewaga tkwi w autentyczności – wskaźnik zaangażowania w tej grupie często przekracza 10%, co jest wynikiem trudnym do osiągnięcia dla większych kont. Określani mianem „architektów zaufania", posiadają relatywnie najwyższy wpływ na

jednostkowe decyzje zakupowe w bardzo wąskich społecznościach zainteresowań, skutecznie domykając strukturę współczesnego influencer marketingu [5].

Klasyfikacja zaprezentowana w tabeli 1. wskazuje na istnienie odwrotnej zależności między skalą dotarcia, a głębokością relacji z odbiorcą.

Tabela 1. Charakterystyka grup influencerów pod kątem parametrów kampanii

Table 1. Characteristics of influencer groups in terms of campaign parameters

Kategoria	Zasięg	Wskaźnik zaangażowania (ER)	Koszt kampanii (CPM)	Główny cel marketingowy
Mega-influencerzy	>1 000 000	Niski (1-2%)	Wysoki	Świadomość marki
Makro-influencerzy	100 tys. - 1 mln	Średni (2-4%)	Średni	Edukacja o produkcie, prestiż
Mikro-influencerzy	10 tys. - 100 tys.	Wysoki (4-8%)	Przystępny	Konwersja i sprzedaż
Nano-influencerzy	<10 000	Bardzo wysoki (>8%)	Niski	Budowanie zaufania w niszy

Źródło: Opracowanie własne na podstawie: https://influencermarketinghub.com/influencer-marketing-benchmark-report/

Z perspektywy inżynierii marketingu optymalizacja kampanii wymaga zatem precyzyjnego doboru profilu twórcy do zdefiniowanych KPI – od maksymalizacji zasięgu przy niskim CPM (mega-influencerzy) po maksymalizację wskaźnika konwersji w grupach niszowych (nano- i mikro-influencerzy). Skuteczna strategia nowoczesnego przedsiębiorstwa coraz rzadziej opiera się na pojedynczym kanale, dążąc raczej do budowy hybrydowych ekosystemów komunikacyjnych, które łączą autorytet profesjonalnych makro-twórców z wysokim poziomem zaangażowania mniejszych liderów opinii [14].

3. Rola i mechanizmy wpływu influencer marketingu na zachowania konsumentów

Analiza wpływu liderów opinii na procesy decyzyjne współczesnych nabywców wymaga interdyscyplinarnego podejścia, łączącego wiedzę z zakresu marketingu, socjologii oraz psychologii poznawczej. W dobie nadmiaru informacji klasyczne modele zakupowe ulegają transformacji, a kluczowym czynnikiem determinującym wybór konkretnej marki staje się rekomendacja pochodząca z wiarygodnego źródła [13].

Influencerzy, pełniąc rolę cyfrowych przewodników, nie tylko dostarczają informacji o produkcie, lecz przede wszystkim nadają mu określony kontekst emocjonalny i społeczny [13].

3.1. Influencer marketing w systemie zintegrowanej komunikacji marketingowej

Współczesne przedsiębiorstwa coraz częściej rezygnują z izolowanych działań promocyjnych na rzecz zintegrowanej komunikacji marketingowej (IMC), w której influencer marketing pełni funkcję strategicznego łącznika. Jego nadrzędnym zadaniem jest nie tylko budowanie świadomości marki (brand awareness), lecz przede wszystkim kształtowanie pożądanego wizerunku oraz generowanie realnego zaangażowania odbiorców. Kluczową cechą tej formy komunikacji jest jej dwukierunkowy charakter [17].

W odróżnieniu od mediów tradycyjnych influencer marketing umożliwia natychmiastową interakcję poprzez komentarze, wiadomości bezpośrednie czy reakcje w czasie rzeczywistym. Dzięki temu narzędzie to wpisuje się w nowoczesne koncepcje marketingu relacji oraz marketingu doświadczeń (experience marketing), w których konsument nie kupuje jedynie produktu, lecz także cały kontekst emocjonalny i społeczny z nim związany.

3.2. Psychologiczne mechanizmy oddziaływania na procesy decyzyjne

Skuteczność liderów opinii w przestrzeni cyfrowej nie jest dziełem przypadku, lecz efektem precyzyjnego wykorzystania mechanizmów psychologii społecznej, które modelują zachowania zakupowe współczesnych konsumentów. Kluczowym fundamentem tej skuteczności jest kapitał zaufania oraz transfer autorytetu. Influencerzy budują swoją pozycję poprzez demonstrowanie autentyczności i eksperckości, co w psychologii wiąże się bezpośrednio z efektem aureoli. Mechanizm ten sprawia, że jeśli odbiorca uznaje twórcę za wiarygodnego w jednej sferze (np. stylu życia), podświadomie przenosi tę wiarygodność na polecane przez niego produkty. Dzięki temu rekomendacje są postrzegane jako obiektywne rady, co znacząco obniża opór psychologiczny konsumenta, typowy dla kontaktu z tradycyjną, agresywną reklamą [20].

Istotnym elementem inżynierii decyzji w mediach społecznościowych jest społeczny dowód słuszności (social proof). Opiera się on na założeniu, że w sytuacjach niepewności jednostka podejmuje decyzje, obserwując zachowania innych [4]. Widząc produkt polecany przez influencera oraz śledząc pozytywne reakcje jego społeczności – wyrażone poprzez polubienia czy komentarze – konsument uzyskuje społeczne potwierdzenie, że dany wybór jest wartościowy i bezpieczny. Proces ten skutecznie redukuje postrzegane ryzyko zakupowe i znacząco przyspiesza przejście użytkownika przez kolejne etapy lejka sprzedażowego [9].

Wpływ ten jest dodatkowo wzmacniany przez mechanizmy identyfikacji społecznej i aspiracyjności. Konsumenci wykazują silną tendencję do traktowania influencerów jako cyfrowych wzorców zachowań, dążąc do zmniejszenia dystansu między „ja realnym" a „ja idealnym". Poprzez naśladowanie preferencji zakupo-

wych i postaw prezentowanych przez lidera opinii konsument buduje własną tożsamość społeczną, a sam akt zakupu staje się formą przynależności do określonej, prestiżowej grupy skupionej wokół twórcy [8].

Dopełnieniem tych procesów jest emocjonalne zaangażowanie oraz storytelling. Treści tworzone przez influencerów rzadko ograniczają się do suchej informacji; zazwyczaj opierają się na rozbudowanej narracji, która angażuje emocjonalnie odbiorcę. Taki narracyjny charakter przekazu sprawia, że informacje o produkcie są lepiej zapamiętywane, a marka zyskuje pożądaną „ludzką twarz". W rezultacie emocjonalna więź z twórcą przekłada się na długofalową lojalność wobec polecanych przez niego marek, co buduje znacznie trwalszą przewagę rynkową niż efekty uzyskiwane za pomocą krótkotrwałych obniżek cenowych czy tradycyjnych komunikatów promocyjnych [12].

4. Analiza przypadków

W celu zweryfikowania teoretycznych założeń dotyczących skuteczności influencer marketingu dokonano analizy dwóch kontrastujących ze sobą kampanii na polskim rynku w formie studium przypadku (case study). Analiza ta pozwala na identyfikację kluczowych czynników sukcesu oraz potencjalnych ryzyk operacyjnych.

4.1. Kampania Genzie x Deichmann – synergia i spójność wizerunkowa

Współpraca grupy projektowej Genzie, działającej pod przewodnictwem Karola „Friza" Wiśniewskiego, z marką Deichmann stanowi modelowy przykład sukcesu opartego na wysokim poziomie dopasowania produktowego (product-fit).

Analiza mechanizmu wskazuje, że produkty (obuwie) zostały zintegrowane z codziennym procesem twórczym grupy. Nie stanowiły jedynie tła, lecz stały się integralnym elementem narracji wideo (tzw. daily vlogów). Dzięki współtworzeniu dedykowanej kolekcji fani grupy poczuli osobisty związek z marką. Kampania skutecznie wykorzystała entuzjazm pokolenia Z, co przełożyło się na wysoką konwersję sprzedażową.

Kluczowym czynnikiem sukcesu była natywność przekazu – produkt pojawiał się w naturalnym środowisku odbiorcy, nie przerywając jego doświadczenia rozrywkowego, jak wynika z analizy materiałów prasowych oraz internetowych [6, 10, 23].

4.2. Współpraca Wersow x Samsung – kryzys wiarygodności i „efekt najemnika"

Przypadek Weroniki Sowy (Wersow) promującej markę Samsung (seria Galaxy Z Flip) stanowi istotne studium ryzyka utraty autentyczności w influencer marketingu. Kampania spotkała się z oporem społeczności, która wskazała na brak spójności historycznej. Twórczyni przez lata była silnie kojarzona z ekosystemem

marki Apple, co w oczach odbiorców uczyniło nagłą zmianę technologiczną niewiarygodną [22].

Negatywny odbiór został pogłębiony przez działania po zakończeniu kontraktu, kiedy to influencerka zdecydowała się na wyprzedaż akcesoriów marki Samsung. Zjawisko to, definiowane w literaturze jako „efekt najemnika", uderza w fundament zaufania odbiorców. Pokazuje ono, że nawet zasięgi liczone w milionach nie są w stanie zrekompensować braku autentycznej sympatii twórcy do promowanego produktu.

Z punktu widzenia inżynierii marki był to błąd w doborze ambasadora, który nie przeszedł pomyślnie testu spójności długofalowej.

5. Znaczenie strategiczne i wyzwania dla przedsiębiorstw

Wdrożenie influencer marketingu do struktury działań promocyjnych przedsiębiorstwa nie może być traktowane jako odizolowany projekt, lecz jako element długofalowej strategii budowania przewagi konkurencyjnej. Z punktu widzenia inżynierii zarządzania kluczowe jest zrozumienie, w jaki sposób relacje z twórcami internetowymi przekładają się na konkretne wskaźniki biznesowe oraz jakie ryzyka operacyjne wiążą się z tą formą komunikacji. Skuteczne wykorzystanie liderów opinii wymaga od organizacji nie tylko nakładów finansowych, lecz przede wszystkim zmiany podejścia do kontroli nad komunikatem marki na rzecz większej elastyczności i autentyczności. Wpływ tej formy promocji na efektywność biznesową przejawia się przede wszystkim w możliwości realizacji celów trudnych do osiągnięcia za pomocą mediów tradycyjnych. Dzięki współpracy z twórcami przedsiębiorstwa zyskują narzędzie do precyzyjnego targetowania, pozwalające na dotarcie do mikronisz rynkowych o bardzo specyficznych profilach zainteresowań [19].

Personalizacja komunikacji, przejawiająca się w dostosowaniu języka i estetyki do danej społeczności, pozwala na budowanie trwałych relacji, które wykraczają poza jednorazową transakcję. Ponadto wdrożenie mierzalnych parametrów, takich jak linki afiliacyjne czy unikalne kody rabatowe, umożliwia precyzyjne monitorowanie wskaźnika ROAS (Return on Ad Spend), dostarczając twardych danych analitycznych niezbędnych do optymalizacji budżetów marketingowych. Mimo wysokiego potencjału branża ta mierzy się z istotnymi ograniczeniami oraz wyzwaniami o charakterze zarówno inżynierskim, jak i etycznym. Krytycznym ryzykiem pozostaje utrata autentyczności – nadmierna komercjalizacja profilu influencera niemal zawsze prowadzi do spadku zaangażowania odbiorców i erozji zaufania. Równie istotnym problemem jest zjawisko fałszywych wskaźników (fake followers), które wymusza na przedsiębiorstwach stosowanie zaawansowanych narzędzi audytowych, takich jak HypeAuditor, w celu weryfikacji realności deklarowanych zasięgów. Współczesna inżynieria marketingu musi również uwzględniać aspekty transparentności prawnej. Zgodnie z aktualnymi wytycznymi UOKiK brak wyraźnego i jednoznacznego oznaczenia współpracy reklamowej

grozi nie tylko wysokimi karami finansowymi, lecz także poważnym kryzysem wizerunkowym marki [21]. Ostatnim istotnym wyzwaniem pozostaje trudność w atrybucji, czyli precyzyjnym określeniu, w jakim stopniu dany influencer przyczynił się do ostatecznej decyzji zakupowej w złożonym, wielokanałowym procesie sprzedażowym. Rozwiązanie tych problemów wymaga od współczesnych menedżerów łączenia kompetencji z zakresu psychologii społecznej z zaawansowaną analityką danych.

6. Ograniczenia i wyzwania influencer marketingu

Mimo wysokiej skuteczności oraz dynamicznego wzrostu nakładów finansowych na ten kanał komunikacji branża influencer marketingu boryka się z szeregiem systemowych wyzwań. Z punktu widzenia zarządzania procesami marketingowymi kluczowe znaczenie ma identyfikacja ryzyk, które mogą negatywnie wpływać na wskaźnik ROI oraz bezpieczeństwo wizerunkowe marki (brand safety) [2].

Największym zagrożeniem dla długofalowej efektywności tych działań jest postępująca erozja zaufania, wynikająca z kryzysu autentyczności oraz zjawiska SCAM-u. Nadmierna liczba publikacji sponsorowanych sprawia, że profil influencera zaczyna być postrzegany jako „słup ogłoszeniowy", co prowadzi do drastycznego spadku zaangażowania (engagement rate). Szczególnie destrukcyjne dla kapitału społecznego twórcy jest promowanie produktów o wątpliwej jakości (np. nieuczciwy dropshipping), co negatywnie rzutuje na wizerunek wszystkich marek kojarzonych z daną osobą.

Wymiar etyczny i biznesowy uzupełniają kwestie braku transparentności oraz coraz bardziej rygorystyczne regulacje prawne. Przez lata branża funkcjonowała w sferze niedomówień w zakresie oznaczania treści reklamowych, co zmieniło wprowadzenie wytycznych przez Urząd Ochrony Konkurencji i Konsumentów. Obecnie wymóg jasnego deklarowania współprac (np. poprzez tagi #reklama) jest nie tylko standardem etycznym, lecz także prawnym. Brak transparentności w tym zakresie naraża przedsiębiorstwa na kary finansowe oraz oskarżenia o wprowadzanie konsumentów w błąd, co w strukturze inżynierii biznesowej traktowane jest jako błąd krytyczny w zarządzaniu komunikacją [21].

Z perspektywy inżynierii danych i analityki influencer marketing wciąż napotyka istotne bariery w precyzyjnym pomiarze skuteczności. Powszechny problem fałszywych obserwujących (fake followers) oraz botów generujących sztuczne zaangażowanie wymusza na organizacjach stosowanie zaawansowanych narzędzi audytowych przed podjęciem współpracy. Dodatkowo w środowisku wielokanałowym (omnichannel) niezwykle trudna staje się precyzyjna atrybucja sprzedaży. Oddzielenie wpływu pojedynczej publikacji influencera od efektów działań SEO czy kampanii Google Ads pozostaje poważnym wyzwaniem technicznym.

Nawet stosowanie dedykowanych parametrów UTM czy unikalnych kodów rabatowych nie pozwala na pełne uchwycenie długofalowego wpływu wizerunkowego, co utrudnia pełną optymalizację procesów zakupowych w oparciu o twarde dane.

7. Podsumowanie i wnioski końcowe

Influencer marketing stanowi obecnie jedno z najistotniejszych narzędzi współczesnej komunikacji marketingowej, a jego znaczenie systematycznie rośnie wraz z postępującą cyfryzacją życia społecznego oraz zmieniającymi się modelami konsumpcji mediów. Niniejszy rozdział wykazał, że skuteczność liderów opinii wynika przede wszystkim z ich zdolności do budowania głębokich, relacyjnych więzi z odbiorcami, co pozwala przedsiębiorstwom realizować cele promocyjne na poziomie trudnym do osiągnięcia przy wykorzystaniu mediów tradycyjnych.

Analiza przypadków rynkowych, w szczególności kampanii marek Deichmann oraz Samsung, dostarcza istotnych wniosków dla praktyki zarządzania marketingowego. Dowodzi ona, że o ostatecznym sukcesie działań nie decyduje jedynie surowy zasięg techniczny liczony w milionach obserwujących, lecz przede wszystkim spójność wizerunkowa, autentyczność przekazu oraz tzw. product-fit.

Przypadek współpracy Weroniki Sowy (Wersow) z marką Samsung stanowi ważną przestrogę dla przedsiębiorstw – współcześni konsumenci, zwłaszcza przedstawiciele pokolenia Z, stają się coraz bardziej krytyczni i potrafią szybko zweryfikować szczerość intencji twórcy. Brak historycznej spójności między stylem życia influencera a promowaną marką może prowadzić do efektu odwrotnego od zamierzonego, osłabiając kapitał zaufania zarówno do twórcy, jak i do marki.

Z perspektywy inżynierii marketingu przyszłość tego narzędzia będzie zorientowana na profesjonalizację analityki danych oraz wykorzystanie algorytmów sztucznej inteligencji w procesie optymalnego doboru twórców do wartości marki. Kluczowym wyzwaniem pozostaje również rygorystyczne przestrzeganie zasad etycznych oraz nowych regulacji prawnych, które wymuszają transparentność komunikacji.

Podsumowując, influencer marketing nie powinien być traktowany przez przedsiębiorstwa jako jednorazowy, taktyczny kanał dystrybucji treści, lecz jako długofalowe partnerstwo strategiczne. Skuteczna kampania w tym obszarze wymaga nie tylko odpowiedniego budżetu, lecz przede wszystkim pogłębionej analizy jakościowej i etycznej, która pozwala budować trwałe i autentyczne doświadczenia współczesnych konsumentów.

Literatura

[1] Audrezet A., de Kerviler G., Guidry Moulard J.: Authenticity under threat: When social media influencers need to go beyond self-presentation, Journal of Business Research, vol. 117,

2020, pp. 557–569, https://doi.org/10.1016/j.jbusres.2018.07.008 (dostęp: 28 lutego 2026 r.).

[2] Bączyk A.: Influencer marketing – narzędzie XXI wieku, Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu, 2017, t. 74, nr 3, https://journals.wsb.poznan.pl/index.php/znwsb (dostęp: 28 lutego 2026 r.).

[3] Brown D., Hayes N.: Influencer Marketing: Who Really Influences Your Customers?, Routledge, London 2008.

[4] Cialdini R. B.: Wywieranie wpływu na ludzi. Teoria i praktyka, Gdańskie Wydawnictwo Psychologiczne (GWP), Gdańsk 2016, s. 112–115.

[5] De Veirman M., Cauberghe V., Hudders L.: Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude, International Journal of Advertising, vol. 36, no. 5, 2017, pp. 798–828, https://doi.org/10.1080/02650487.2017.1348035 (dostęp: 28 lutego 2026 r.).

[6] Deichmann SE: Press, https://corpsite.deichmann.com/en-DE/press (dostęp: 28 lutego 2026 r.).

[7] Djafarova E.; Rushworth C.: Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users, Computers in Human Behavior, vol. 68, 2017, pp. 1–7, https://doi.org/10.1016/j.chb.2016.11.009 (dostęp: 28 lutego 2026 r.).

[8] Escalas J. E.; Bettman J. R.: Self-construal, reference groups, and brand meaning, Journal of Consumer Research, vol. 32, no. 3, 2005, pp. 378–389, https://doi.org/10.1086/497549 (dostęp: 28 lutego 2026 r.).

[9] Freberg K.; Graham K.; McGaughey K.; Freberg L. A.: Who are the social media influencers? A study of public perceptions of personality, Public Relations Review, vol. 37, no. 1, 2011, pp. 90–92, https://doi.org/10.1016/j.pubrev.2010.11.001 (dostęp: 28 lutego 2026 r.).

[10] Genzie – kanał YouTube, https://www.youtube.com/@GENZIEE (dostęp: 28 lutego 2026 r.).

[11] Haenlein M., Libai B.: Seeding, referral, and recommendation: Creating profitable word-of-mouth programs, California Management Review, vol. 59, no. 3, 2017, pp. 5–26, https://doi.org/10.1177/0008125617697943 (dostęp: 28 lutego 2026 r.).

[12] Herskovitz S., Crystal M.: The essential brand persona: Storytelling and branding, Journal of Business Strategy, vol. 31, no. 3, 2010, pp. 21–28, https://doi.org/10.1108/02756661011036673 (dostęp: 28 lutego 2026 r.).

[13] Hovland C. I., Weiss W.: The influence of source credibility on communication effectiveness, Public Opinion Quarterly, vol. 15, no. 4, 1951, pp. 635–650, https://doi.org/10.1086/266350 (dostęp: 28 lutego 2026 r.).

[14] HypeAuditor, Global insights on influencer marketing, 2026, https://hypeauditor.com/resources/whitepapers/ (dostęp: 28 lutego 2026 r.).

[15] IAB Polska, Przewodnik po influencer marketingu, https://influencermarketing.org.pl/ (dostęp: 28 lutego 2026 r.).

[16] Influencer Marketing Hub, The State of Influencer Marketing 2024: Benchmark Report, 2024, https://influencermarketinghub.com/influencer-marketing-benchmark-report/ (dostęp: 28 lutego 2026 r.).

[17] Kliatchko J.: Towards a new definition of integrated marketing communications (IMC), International Journal of Advertising, vol. 24, no. 1, 2005, pp. 7–34, https://doi.org/10.1080/02650487.2005.11072902 (dostęp: 28 lutego 2026 r.).

[18] Kotler P., Keller K. L.: Marketing, Wydawnictwo Rebis, Poznań 2016.
[19] Kumar V., Mirchandani R.: Increasing the ROI of social media marketing, MIT Sloan Management Review, vol. 54, no. 1, 2012, pp. 55–61, https://sloanreview.mit.edu/article/increasing-the-roi-of-social-media-marketing/ (dostęp: 28 lutego 2026 r.).
[20] Lou C., Yuan S.: Influencer marketing: How message value and credibility affect consumer trust of branded content on social media, Journal of Interactive Advertising, vol. 19, no. 1, 2019, pp. 58–73, https://doi.org/10.1080/15252019.2018.1533501 (dostęp: 28 lutego 2026 r.).
[21] Prezes Urzędu Ochrony Konkurencji i Konsumentów: Rekomendacje Prezesa UOKiK dotyczące oznaczania treści reklamowych przez influencerów w mediach społecznościowych, Warszawa 2022, https://www.iab.org.pl/wp-content/uploads/2022/09/Rekomendacje_Prezesa_UOKiK.pdf (dostęp: 28 lutego 2026 r.).
[22] Schouten A. P., Janssen L., Verspaget M.: Celebrity vs. influencer endorsements in advertising: The role of identification, credibility, and product-endorser fit, International Journal of Advertising, vol. 39, no. 2, 2020, pp. 258–281, https://doi.org/10.1080/02650487.2019.1634898 (dostęp: 28 lutego 2026 r.).
[23] SocialPress.pl: Fenomen Genzie – jak budować zaangażowanie wśród Gen Z, https://socialpress.pl/ (dostęp: 28 lutego 2026 r.).
[24] Sznajder A.: Technologie mobilne w marketingu, Wolters Kluwer, Warszawa 2014.

INFLUENCER MARKETING AS A MARKETING COMMUNICATION TOOL AND ITS IMPACT ON CONSUMER BEHAVIOR

Summary

Contemporary enterprises operate in a dynamic and highly information-saturated environment in which traditional forms of promotion are gradually being replaced by interactive social media. The development of digital technologies and the declining level of trust in mass advertising have contributed to the growing importance of influencer marketing as a significant tool of marketing communication. This approach enables brands to build direct and personalized relationships with consumers within the framework of integrated marketing communication (IMC). The aim of this article is to analyze influencer marketing as an instrument shaping consumer attitudes and purchasing decisions, with particular emphasis on Generation Z and Millennials. The article discusses the key psychological and social mechanisms determining the effectiveness of this tool, including social proof, social identification, and authority bias. It also highlights the importance of authenticity, credibility, and the image consistency of content creators in the process of building audience trust. The theoretical section defines the essence of influencer marketing and presents a classification of opinion leaders, including nano-, micro-, macro-, and mega-influencers. The empirical section is based on the case study method and includes an analysis of selected marketing campaigns conducted by the brands Deichmann and Samsung. The analysis identified factors determining the effectiveness of influencer marketing activities, particularly the importance of aligning the influencer with the brand's values and image. The article also highlights challenges related to the further development of this form of marketing communication, including the need for the professionalization of the industry, maintaining transparency in brand collaborations, and adapting marketing activities to applicable legal regulations concerning the labeling of sponsored content.

Keywords: marketing strategies, social media, opinion leaders, brand authenticity.

Aleksandra Brożyna, Dorota Głowacz-Czerwonka
Koło Naukowe Studentów Chemii „Esprit”

WPŁYW PŁATKÓW RÓŻY I GRAFITU EKSPANDOWANEGO NA WŁAŚCIWOŚCI FIZYCZNE SZTYWNYCH PIANEK POLIURETANOWYCH

Streszczenie

Niniejszy rozdział przedstawia wyniki badań dotyczące właściwości fizycznych sztywnych pianek poliuretanowych, które nzostały zmodyfikowane poprzez wprowadzenie naturalnego napełniacza w postaci płatków róży oraz uniepalniacza. W procesie otrzymywania poliuretanów wykorzystano poliol RF 151V, natomiast rolę środka ograniczającego palność pełnił grafit ekspandowany EG 290. Zakres przeprowadzonych badań obejmował analizę właściwości fizycznych, takich jak gęstość pozorna, masowa oraz objętościowa chłonność wody, stabilność wymiarów, a także współczynnik przewodzenia ciepła. Wyniki analiz dowiodły, że zarówno obecność biomasy kwiatowej, jak i grafitu, w istotny sposób wpłynęła na morfologię otrzymanych materiałów. Pianki charakteryzowały się obniżoną gęstością pozorną w porównaniu do próbek kontrolnych, co było jednoznaczne z uzyskaniem struktury o wysokim stopniu spienienia. Ponadto próbki wykazały zwiększoną chłonność wody. W przypadku badania stabilności wymiarów próbki zawierające tylko dodatek płatków róży wykazywały najmniejszy ubytek masy, gdy ilość napełniacza nie przekraczała 10%, podczas gdy w piankach modyfikowanych grafitem optymalna wartość wyniosła 5%. Analiza współczynnika przewodzenia ciepła wykazała, iż kompozyty zawierające uniepalniacz charakteryzują się wyższymi wartościami tego parametru w porównaniu do próbek referencyjnych. Zjawisko to jest bezpośrednio związane z obecnością grafitu w strukturze materiału, który cechuje się naturalnie wysokim przewodnictwem cieplnym. Wszystkie uzyskane materiały cechowały się wynikami pozwalającymi uznać je za skuteczne izolatory termiczne. Otrzymane poliuretany mogą stanowić alternatywę dla powszechnie dostępnych tworzyw sztucznych na rynku.

Słowa kluczowe: biomasa, uniepalniacz, poliuretany, płatki róży, gęstość pozorna, stabilność wymiarów.

1. Wprowadzenie

Pianki poliuretanowe należą do jednej z najszerzej rozwiniętych grup polimerów wykorzystywanych w produkcji przemysłowej. Rosnące zainteresowanie tymi

materiałami wynika z ich wszechstronnych właściwości oraz możliwości zastosowania w wielu sektorach przemysłu. Materiały testowane są jako różnego rodzaju kleje, powłoki oraz uszczelniacze [4]. Równocześnie ze względu na swoją hemokompatybilność oraz biozgodność często wykorzystywane są w medycynie w produkcji protez, cewników czy zastawek sercowych. Pianki poliuretanowe stanowią także kluczowy komponent w produkcji mebli oraz elementów wyposażenia pojazdów [2].

Polimery te zyskują coraz szersze zainteresowanie ze względu na łatwość modyfikacji struktury. Pozwala to nie tylko na dostosowanie budowy danego poliuretanu pod konkretną implementację ale także na wprowadzenie surowców naturalnych do procesu wytwarzania. Wykorzystanie tego typu składników w procesie produkcji daje możliwość zastąpienia wykorzystywanych do tej pory złóż paliw kopalniach na surowce z odnawialnych źródeł [7]. Temat ten przyciąga uwagę ze względu wyzwania pojawiające się przed producentami tworzyw sztucznych. Występujący problem dotyczy w głównej mierze podstawowych surowców wykorzystywanych w produkcji poliuretanów. Powszechnie stosowane poliole oraz izocyjaniany otrzymywane są z produktów pochodzenia ropopochodnego. Stale malejąca ilość paliw kopalnianych skutkuje systematycznym ograniczeniem ich dostępności co w konsekwencji prowadzi do wzrostu cen rynkowych. Drugim aspektem jest negatywne oddziaływanie tych surowców na środowisko naturalne. Zwiększony popyt na materiały polimerowe wiążę się z intensywną eksploatacją tych złóż co przyczynia się do wzrostu emisji szkodliwych substancji oraz do silnej ingerencji w ekosystem. W konsekwencji obydwa te czynniki odpowiedzialne są za szukanie przez naukowców alternatywnych rozwiązań [6]. Pojawiająca się możliwość modyfikacji pianek poliuretanowych materiałami pochodzenia naturalnego może stanowić istotny krok w kierunku bardziej zrównoważonych technologii. Dodatek naturalnych napełniaczy zyskuję coraz większą popularność, a próby zastosowania go opierają się głównie na surowcach takich jak kwiaty, nasiona, liście, łodygi czy łupiny [5]. Wprowadzenie do struktury polimeru biomasy może umożliwić poprawienie właściwości danego materiału, a także zminimalizować rosnący problem zagospodarowania bioodpadów. Ponadto obróbka tych surowców pod zastosowanie w produkcji wiąże się z niewielkim nakładem energetycznym.

Obecnie literatura opisuje próby zastosowania naturalnych składników takich jak pestki śliwek czy fusy z kawy. Napełniacze te oddziałują na matrycę poliuretanową przede wszystkim w sposób fizyczny. Ich obecność wpływa na przebieg procesu spieniania, morfologię komórkową oraz wybrane właściwości mechaniczne i termoizolacyjne otrzymanych pianek. Analizy te głównie przedstawiają pozytywne wyniki oraz poprawę właściwości otrzymywanych materiałów [1,3].

Celem badania było otrzymanie sztywnych pianek poliuretanowych zawierających dodatek płatków róży oraz uniepalniacza, a także określenie wpływu tych dodatków na właściwości fizyczne materiału.

2. Cześć doświadczalna

2.1. Aparatura laboratoryjna

Wyposażenie laboratoryjne używane w procesie otrzymywania i badania pianek poliuretanowych: suszarka SLW 53 STD (Pol-Eko Aparatura, Polska), waga analityczna E42 (Radwag, Polska), waga techniczna PM 600 (Szwajcaria), suwmiarka elektroniczna (Niteo, Polska), mieszadło mechaniczne (Meramont WE27T, Polska), aparat do oznaczania przewodnictwa cieplnego IZOMET 2114 (Applied Precision, Słowacja).

2.2. Otrzymanie pianek poliuretanowych

W pierwszej kolejności przygotowano pianki referencyjne: bez dodatków oraz z grafitem ekspandowanym stosowanym jako uniepalniacz. Następnie wytworzono dwie serie pianek z dodatkiem płatków róży. Pierwszą grupę stanowiły materiały zawierające płatki róży w ilości odpowiednio 5, 10 i 15% masowych. Druga seria zawierała analogiczną ilość biomasy różanej oraz grafit ekspandowany w stałej ilości we wszystkich próbkach.

Do procesu otrzymywania pianek poliuretanowych wykorzystano poliol Rokopol RF151V (PCC Rokita, Polska). Następnie, dodawano kolejno silikon L-6900 (Momentive Performance Materials, Belgia), trietyloaminę TEA (Chempur, Polska), wodę dejonizowaną, płatki roży (Greenok Place, Egipt) oraz w przypadku pianek z grafitem EG290 (Sinograf, Polska). W kolejnym kroku wprowadzano 4,4’-diizocyjanian difenylometanu Ongronat 2100 (BorsodChem, Węgry).

Pianki otrzymywano w kubkach polipropylenowych, odważając do nich określone ilości surowców: poliolu (100 g), TEA (0,1 g), silikonu (0,1 g) wody (2% mas.), płatków róży (5-15% mas.) oraz EG 290. Całość wymieszano przy użyciu mieszadła mechanicznego i obserwowano wzrost pianki. Po otrzymaniu pozostawiano je na 72 godz., aby uległy dosieciowaniu.

Uzyskano 8 kompozycji piankowych:

- RF 151V (RF) Ref.
- RF 151V (RF) Ref. + EG 290
- RF+ róża 5%
- RF+ róża 10%
- RF+ róża 15%
- RF+ róża 5% + EG 290
- RF+ róża 10% + EG 290
- RF+ róża 15% + EG 290

2.3. Metodologia

2.3.1. Gęstość pozorna

W celu określenia gęstości pozornej bryły sześcienne o wymiarach 3 x 3 x 3 cm zważono oraz zmierzono przy użyciu suwmiarki. Procedura została wykonana zgodnie z normą PN-EN ISO 845:2010.

2.3.2. Chłonność wody

Badanie chłonności przeprowadzono umieszczając próbki w wodzie, a parametry odczytywano po 5 min, 3 i 24 godzinach zanurzenia. Analiza została wykonana zgodnie z normą PN-93/C-89084. Określono objętościową chłonność wody zgodnie z poniższym wzorem.

$$W_V = \frac{m_2 - m_1}{V \cdot \rho} \cdot 100\% \tag{1}$$

gdzie: W_v – objętościowa chłonność wody
m_1 – masa próbki przed zanurzeniem w wodzie [g]
m_2 – masa próbki po zakończeniu procesu nasiąkania [g]
V – objętość próbki [cm^3]
ρ – gęstość wody [g/cm^3]

2.3.3. Stabilność wymiarów

Badanie stabilności wymiarów przeprowadzono zgodnie z normą PN-92/C89083. W tym celu wykorzystano po dwie beleczki dla każdej z serii badanego materiału o wymiarach równych 10 x 1 x 1 cm. Próbki poddano działaniu temperatury 150°C i monitorowano zmianę wymiarów po upływie 20 i 40 godzin.

2.3.4. Współczynnik przewodzenia ciepła

Współczynnika przewodzenia ciepła oznaczono zgodnie z normą PN-EN 993-15:2006. W tym celu wykorzystano aparaturę IZOMET 2114 (Applied Precision). Analizę przeprowadzono umieszczając sondę igłową w badanej piance i dokonywano pomiaru, poprzez przepuszczenie prądu o znanym natężeniu. Na podstawie zmian temperatury rejestrowanych przez urządzenie wyznaczano wartość współczynnika przewodzenia ciepła materiału.

3. Wyniki i dyskusja

3.1. Gęstość pozorna

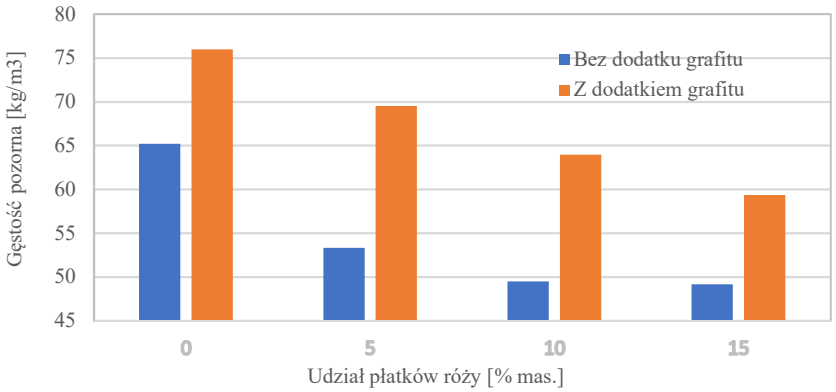
Kompozycje piankowe (tab. 1.) zawierające dodatek płatków róży wykazały niższe wartości gęstości pozornej niż próbka referencyjna. Zaobserwowano spadek gęstości pozornej wraz z rosnącą ilością napełnienia w składzie. Takie wyniki świadczą o występowaniu wysokorozwiniętej struktury porowatej.

Tabela 1. Wpływ składu kompozycji piankowych na wartość gęstości pozornej

Table 1. The effect of the composition of foam compositions on the apparent density value

Lp.	Nazwa kompozycji	Gęstość pozorna [kg/m^3]
1.	RF 151V (RF) Ref.	65,2
2.	RF+ róża 5%	53,3
3.	RF+ róża 10%	49,5
4.	RF+ róża 15%	49,2
5.	RF 151V (RF) + EG290 Ref.	76,0
6.	RF+ róża 5% + EG290	69,5
7.	RF+ róża 10% + EG290	64,0
8.	RF+ róża 15% + EG290	59,3

Rysunek 1. przedstawia wpływ zawartości płatków róży na strukturę otrzymanych pianek. Kompozycje zawierające dodatek uniepalniacza wykazują wyższe wartości niż w kompozycje zawierających tylko dodatek róży. Otrzymane rezultaty świadczą o tym, że dodatek surowców naturalnych w bezpośredni sposób oddziałuje na mechanizm spieniania oraz morfologię kompozycji.



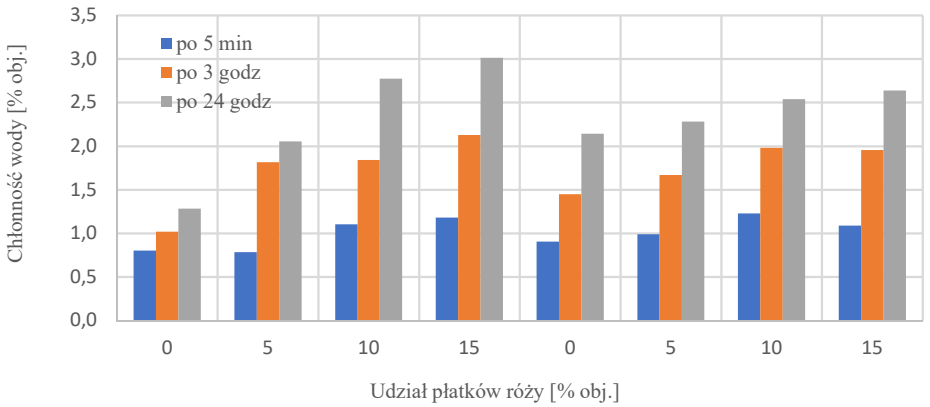
Rys. 1. Wpływ składu kompozycji na jej gęstość pozorną

Pic. 1. The effect of composition on apparent density

3.2. Chłonność wody

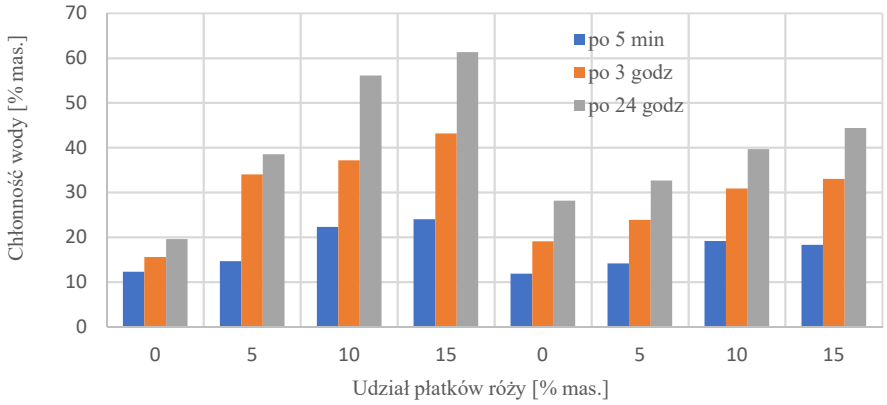
Analizując poniższe wykresy (rys. 2. i rys. 3.) zaobserwowano wzrost masowej oraz objętościowej chłonności wody wraz z rosnącą ilością surowca roślinnego w składzie. Ponadto wszystkie z badanych pianek wykazały zwiększoną sorpcję w porównaniu do próbek kontrolnych. Nietypową zależność zaobserwowano w przypadku odczytywania wyników po 5 min dla pianki modyfikowanej grafitem zawierającej 15% mas. płatków róży. Podczas gdy pozostałe pianki dla

tego czasu charakteryzowały się rosnącą chłonnością wody wraz z rosnąca ilością napełnienia, analizowana próbka początkowo wykazywała niższe wartości tego parametru. W kolejnych pomiarach jej chłonność wody wzrosła i była zgodna z ogólnym trendem obserwowanym dla pozostałych próbek. Powodem mogła być ograniczona dostępność porów w matrycy poliuretanowej spowodowana lokalną aglomeracją kwiatów co wiązało się z niewystarczającą ilością czasu na powolną penetrację wody do mikroporów lub błędem pomiarowym.



Rys. 2. Wpływ składu pianek poliuretanowych na objętościową chłonność wody. Od lewej: bez dodatku grafitu, z dodatkiem grafitu

Pic. 2. The effect of the composition of polyurethane foams on the volumetric water absorption. From left to right: without graphite addition, with graphite addition



Rys. 3. Wpływ składu pianek poliuretanowych na masową chłonność wody. Od lewej: bez dodatku grafitu, z dodatkiem grafitu

Pic. 3. The effect of the composition of polyurethane foams on mass water absorption. From left to right: without graphite addition, with graphite addition

Badanie to w istotny sposób wiąże się z uzyskanymi wartościami gęstości pozornej. Pianki charakteryzujące się najniższą wartością gęstości pozornej przedstawiają największe wartości chłonności wody. Tak jak opisano wcześniej wykazują one wysoki stopień spienienia, który z kolei sprzyja sorpcji wody. Pianki wykorzystane w badaniu zaprezentowano poniżej rys. 4.



Rys. 4. Pianki poliuretanowe wykorzystane w badaniu chłonności wody

Pic. 4. Polyurethane foams used in the water absorption test

3.3. Stabilność wymiarów

Wyniki dotyczące badania stabilności wymiarów zaprezentowano w tab. 2.

Tabela 2. Ubytek masy oraz zmiana wymiarów po ekspozycji w suszarce laboratoryjnej

Table 2. Loss of mass and change in dimensions after exposure in a laboratory dryer

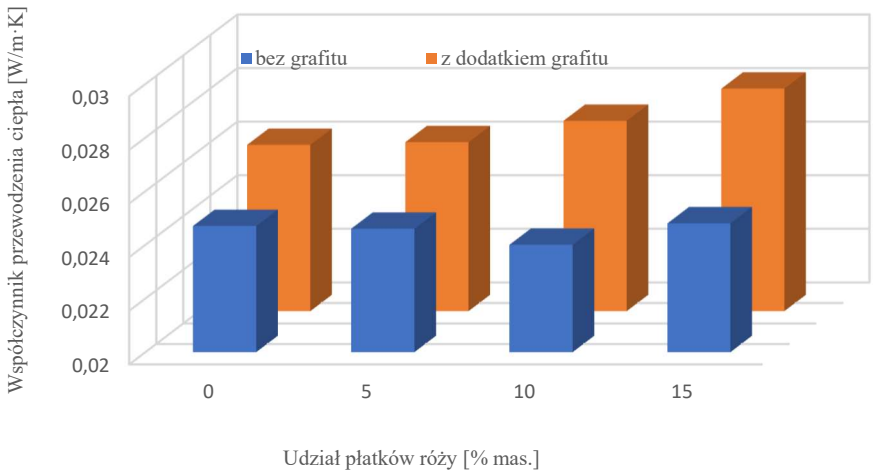
Lp.	Nazwa kompozycji	Ubytek masy [% mas.] po:		Zmiana długości [% liniowy] po:	
		20 godz.	40 godz.	20 godz.	40 godz.
1.	RF 151V (RF) Ref.	2,4	3,0	5,1	5,2
2.	RF+ róża 5%	1,9	2,2	1,7	5,0
3.	RF+ róża 10%	1,6	1,8	1,4	2,8
4.	RF+ róża 15%	3,3	4,1	1,1	2,7
5.	RF 151V (RF) + EG290 Ref.	2,4	3,0	10,9	12,5
6.	RF+ róża 5% + EG290	2,1	2,7	3,5	5,6
7.	RF+ róża 10% + EG290	2,7	3,4	1,5	3,3
8.	RF+ róża 15% + EG290	3,8	4,5	1,1	1,5

Kompozycje zawierające tylko dodatek płatków róży w ilości 5% oraz 10% charakteryzowały się mniejszym ubytkiem masy w porównaniu do próbek referencyjnych. W przypadku pianek zawierających grafit taka tendencja jest zachowana tylko gdy róża występuje w najmniejszej zawartości. Zastosowanie surowca naturalnego zarówno w jednym jak i drugim przypadku spowodowało

zmniejszenie zmiany wymiarów liniowych w porównaniu do referencji. Zaobserwowano także malejącą zależności pomiędzy wrastającą ilością wypełnienia, a liniową zmianą wymiarów. Najkorzystniejsze wyniki pod względem ograniczenia ubytku masy przedstawiały pianki zawierające płatki róży w ilości 10% mas., natomiast w materiałach modyfikowanych grafitem najlepszy efekt osiągnięto przy 5% zawartości dodatku.

3.4. Współczynnik przewodzenia ciepła

Analizując rysunek 5. zaobserwowano, iż wprowadzenie do kompozycji płatków róży w ilości 5 i 10% mas. spowodowało zmniejszenie przewodności cieplnej w porównaniu do próbek kontrolnych, natomiast przy ilości 15% wartość ta wzrasta. Prawdopodobnie jest to wynikiem występowania zbyt dużej ilości wypełnienia powodującego zaburzenia strukturalne. Mimo to wartość ta nie odbiega w znaczący sposób od referencji.



Rys. 5. Wpływ składu pianek poliuretanowych na współczynnik przewodzenia ciepła

Pic. 5. The effect of the composition of polyurethane foams on the thermal conductivity coefficient

Kompozycje zawierające unieplaniacz wykazują wzrastające przewodnictwo cieplne wraz z rosnącą ilością surowca roślinnego w składzie. Ponadto pianki te charakteryzowały się wyższymi wartościami niż próbki referencyjne. Związane jest to bezpośrednio z obecnością grafitu, który przedstawia wysoki współczynniki przewodzenia ciepła. Wszystkie z otrzymanych wyników charakteryzują analizowane pianki jako dobre izolatory cieplne.

4. Podsumowanie

Na podstawie przeprowadzonych badań zaobserwowano, iż wprowadzenie do kompozycji zarówno płatków róży jak i grafitu ekspandowanego w istotny sposób wpływa na strukturę oraz właściwości otrzymanych pianek. Napełniacz naturalny odpowiadał za uzyskanie struktur charakteryzujących się wyższym stopniem spienienia co bezpośrednio wiąże się z otrzymanymi wartościami masowej oraz objętościowej chłonności wody. Pianki zawierające największą ilość surowca naturalnego wykazały najniższe wartości gęstości pozornej i jednocześnie najwyższe wartości dotyczące chłonności wody. Obecność kwiatów w układzie odpowiedzialna jest za wysoki stopień spienienia sprzyjający zwiększonej chłonności wody. Ponadto otrzymane pianki przedstawiały wyniki na poziomie dobrych izolatorów cieplnych. Najlepsze właściwości wykazały pianki zawierające płatki róży w ilości 5% oraz 10% zarówno te bez grafitu jak i z jego dodatkiem. Analiza ta może być odpowiedzią na stale rosnący problem gromadzenia odpadów naturalnych w środowisku. Otrzymane kompozycje wykazywały dobre właściwości co stwarza możliwość do dalszego rozpatrywania tematu ze względu na możliwość zastąpienia podstawowych składników wykorzystywanych w produkcji poliuretanów stanowiących surowce ropopochodne. Zastosowanie płatków róży jako napełniacza kompozycji spienionych może być istotnym krokiem w kierunku stworzenia ekologicznych poliuretanów.

Literatura

[1] Auguścik-Królikowska M.: Struktura i właściwości lepkosprężystych pianek poliuretanowych napełnionych fusami kawy, Polimery, nr 65 (10), 2020, s. 708-718.

[2] Haponiuk J.T., Hejna A., Piszczyk Ł.: Wykorzystanie surowców odnawialnych i odpadowych w syntezie poliuretanów, Elastomery, 18 (3), 2014, s. 21-30.

[3] Miedzińska K.: Biobased polyurethane composite foams reinforced with plum stones and silanized plum stones, International Journal of Molecular Sciences, 22 (9), 2021, p. 6-8.

[4] Prociak A., Rokicki G., Ryszkowska J.: Materiały poliuretanowe, Wydawnictwo Naukowe PWN, Warszawa 2016, s. 250-264.

[5] Ryszkowska J.: Materiały poliuretanowe wytwarzane z zastosowaniem surowców odnawialnych, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2019, s. 82-90.

[6] Zakrzewska P., Kuźnia M., Telejko T.: Nowe perspektywy w produkcji pianek poliuretanowych: alternatywne źródła dla polioli petrochemicznych, Materiały konferencyjne Sympozjum Inżynierii Materiałowej, AGH, Kraków 2024, s. 71.

[7] Żółtowska K., Sobczak M., Olędzka E.: View of polyurethanes in pharmacy – current state and perspectives of the development, Polimery, 10, 2014, p. 687-760.

THE EFFECT OF ROSE PETALS AND EXPANDED GRAPHITE ON THE PHYSICAL PROPERTIES OF RIGID POLYURETHANE FOAMS

Summary

This article presents the results of research on the physical properties of rigid polyurethane foams that have been modified by the introduction of a natural filler in the form of rose petals and a flame retardant. In the process of obtaining polyurethanes, polyol RF 151V was used, while expanded graphite EG 290 served as a flame retardant. The scope of the conducted research included the analysis of physical properties, such as apparent density, mass and volume water absorption, dimensional stability, and thermal conductivity coefficient. Analysis results proved that both the presence of floral biomass and graphite significantly influenced the morphology of the materials obtained. The foams had a lower apparent density compared to the control samples, which meant that they had a highly foamed structure. Additionally, the samples showed increased water absorption. In the case of dimensional stability testing, samples containing only a rose petal additive showed the best properties when the filler amount did not exceed 10%, while in graphite-modified foams the optimal value was 5%.Analysis of the thermal conductivity coefficient showed that composites containing flame retardant have higher values of this parameter compared to reference samples. This phenomenon is directly related to the presence of graphite in the material structure, which is characterized by naturally high thermal conductivity. All the obtained materials were characterized by results that allowed them to be considered as effective thermal insulators. The obtained polyurethanes can be an alternative to commonly available plastics on the market.

Keywords: biomass, flame retardant, polyurethanes, rose flower, apparent density, dimensional stability.

Gabriela Bober, Dorota Głowacz-Czerwonka
Koło Naukowe Studentów Chemii „Esprit”

WPŁYW DODATKU SUSZONYCH KWIATÓW GOŹDZIKA NA WŁAŚCIWOŚCI FIZYCZNE I OGNIOWE SZTYWNYCH PIANEK POLIURETANOWYCH

Streszczenie

W pracy przedstawiono wyniki badań dotyczących wpływu dodatku suszonych kwiatów goździka na właściwości fizyczne oraz ogniowe sztywnych pianek poliuretanowych. W ramach badań otrzymano serię kompozycji obejmującą materiał referencyjny oraz kompozycje zawierające 5, 10 oraz 15% mas. dodatku bio-wypełniacza roślinnego. Wytworzone materiały poddano analizie w celu określenia wpływu dodatku roślinnego na przebieg procesu spieniania oraz wybrane właściwości fizyczne i ogniowe pianek poliuretanowych. Dla otrzymanych materiałów zbadano gęstość pozorną, chłonność wody, stabilność wymiarów oraz współczynnik przewodzenia ciepła. Dodatkowo oceniono właściwości związane z zachowaniem materiałów w warunkach oddziaływania płomienia, w tym indeks tlenowy oraz test poziomy. Wprowadzenie dodatku roślinnego wpływało na przebieg procesu spieniania, co mogło oddziaływać na budowę komórkową materiału i znajdowało odzwierciedlenie w zmianach jego właściwości użytkowych. Wraz ze wzrostem zawartości kwiatów goździka obserwowano zmniejszenie gęstości materiałów oraz wzrost ich chłonności wody, co może być związane z obecnością porowatej struktury bio-wypełniacza oraz zmianą morfologii komórek pianek. W przeprowadzonych badaniach analizowano również wpływ dodatku kwiatów goździka na przewodnictwo cieplne pianek, które stanowi istotny parametr w kontekście potencjalnych zastosowań izolacyjnych. Uzyskane wyniki wskazują, że suszone kwiaty goździka mogą stanowić potencjalny naturalny modyfikator właściwości pianek poliuretanowych. Zastosowanie tego typu dodatku może przyczyniać się do modyfikacji właściwości fizycznych materiałów poliuretanowych oraz stanowić interesujący kierunek rozwoju kompozytów polimerowych zawierających składniki pochodzenia naturalnego.

Słowa kluczowe: kompozycje polimerowe, bio-wypełniacze roślinne, gęstość pozorna, chłonność wody, współczynnik przewodzenia ciepła.

1. Wprowadzenie

Pianki poliuretanowe należą do jednej z najważniejszych grup materiałów polimerowych o strukturze komórkowej. Dzięki zróżnicowanej gęstości, dobrym właściwościom termoizolacyjnym oraz możliwości szerokiej modyfikacji składu chemicznego znajdują one zastosowanie w wielu gałęziach przemysłu, w szczególności w budownictwie, przemyśle meblarskim oraz w produkcji materiałów izolacyjnych [7-9]. Sztywne pianki poliuretanowe charakteryzują się przede wszystkim niskim współczynnikiem przewodzenia ciepła oraz stosunkowo dobrą stabilnością wymiarową, co sprawia, że są one powszechnie stosowane jako materiały izolacyjne w przegrodach budowlanych i instalacjach technicznych [7-8].

W ostatnich latach obserwuje się rosnące zainteresowanie wykorzystaniem surowców odnawialnych w technologii materiałów polimerowych. Zastosowanie dodatków pochodzenia naturalnego pozwala częściowo ograniczyć zużycie surowców petrochemicznych, a jednocześnie może prowadzić do modyfikacji właściwości fizycznych i strukturalnych otrzymywanych materiałów [2, 4]. W szczególności bio-wypełniacze pochodzenia roślinnego są coraz częściej stosowane jako dodatki do pianek poliuretanowych, ponieważ mogą wpływać na przebieg procesu spieniania oraz kształtowanie budowy komórkowej powstających materiałów [2].

Materiały roślinne zawierają w swojej strukturze lignocelulozę, której głównymi składnikami są celuloza, hemiceluloza oraz lignina. Składniki te zawierają liczne grupy hydroksylowe, które mogą uczestniczyć w tworzeniu wiązań wodorowych oraz wpływać na oddziaływania występujące na granicy faz pomiędzy biowypełniaczem a matrycą poliuretanową. Obecność cząstek lignocelulozowych może również modyfikować przebieg procesu spieniania i kształtowanie struktury komórkowej materiału. W rezultacie obserwowane mogą być zmiany właściwości pianek poliuretanowych, takich jak gęstość pozorna, chłonność wody, właściwości mechaniczne czy właściwości termoizolacyjne [2, 4].

Jednym z potencjalnych bio-wypełniaczy stosowanych w materiałach poliuretanowych mogą być suszone kwiaty roślin. Materiały roślinne charakteryzują się złożoną strukturą komórkową, obejmującą komórki oraz przestrzenie międzykomórkowe, które zgodnie z danymi literaturowymi mają głównie charakter makroporowaty [2]. Obecność takiej struktury oraz naturalnych związków organicznych może wpływać na absorpcję składników mieszaniny reakcyjnej, a tym samym oddziaływać na przebieg procesu spieniania i formowanie struktury komórkowej pianek poliuretanowych. Zastosowanie tego typu dodatków wpisuje się w aktualny trend rozwoju materiałów polimerowych o bardziej zrównoważonym charakterze oraz zwiększonym udziale surowców pochodzenia naturalnego.

Pomimo licznych zalet pianek poliuretanowych ich istotnym ograniczeniem jest podatność na działanie wysokiej temperatury oraz ognia. Materiały te należą do grupy polimerów palnych i podczas spalania mogą wydzielać znaczną ilość energii cieplnej oraz produktów rozkładu termicznego [1, 3, 5]. Z tego względu

ważnym elementem badań nad nowymi kompozycjami poliuretanowymi jest analiza ich właściwości ogniowych oraz określenie wpływu różnych dodatków na przebieg procesu spalania [6].

Celem niniejszej pracy było określenie wpływu dodatku suszonych kwiatów goździka na właściwości fizyczne oraz ogniowe sztywnych pianek poliuretanowych.

2. Część doświadczalna

2.1. Otrzymywanie sztywnych pianek poliuretanowych

Pianki otrzymywano w polipropylenowych kubkach, odważając do nich określone ilości surowców i odczynników:

- poliol Rokopol RF151V, PCC Rokita S.A., Brzeg Dolny, Polska (100 g),
- trietyloamina, Chempur, Piekary Śląskie, Polska (0,1 g),
- silikon L6900, Momentive Performance Materials, Belgia (0,1 g),
- suszone kwiaty goździka, Ziołowy Dwór, Polska (5-15% mas.),
- woda dejonizowana (2% mas.)
- MDI, BorsodChem, Węgry (200,0 g)

Całość wymieszano przy użyciu mieszadła mechanicznego. Po wyschnięciu otrzymane pianki pozostawiono je w warunkach laboratoryjnych, aby uległy dosieciowaniu. Przygotowano następujące kompozycje: piankę referencyjną oraz serię pianek z dodatkiem suszonych kwiatów goździka w ilości 5, 10 i 15% mas.

W celu określenia wpływu dodatku suszonych kwiatów goździka na właściwości sztywnych pianek poliuretanowych zbadano: gęstość pozorną, chłonność wody, stabilność wymiarów, współczynnik przewodzenia ciepła, indeks tlenowy, test poziomy.

2.2. Gęstość pozorna

Badanie gęstości pozornej przeprowadzono zgodnie z normą PN-EN ISO 845-2010. Z każdej pianki wycięto sześciany o wymiarach 3×3×3 cm. Kształtki zważono na wadze z dokładnością do 0,0001 g i zmierzono suwmiarką. Na podstawie otrzymanych wartości obliczono gęstość pozorną każdej pianki.

2.3. Chłonność wody

Badanie wykonano zgodnie z normą PN-93/C-890084. Wykorzystano sześcienne kształtki, które zanurzono w wodzie destylowanej i sprawdzono ich masę oraz wymiary kolejno po 5 min., 3 godz. i 24 godz. ekspozycji w wodzie. Na podstawie otrzymanych wartości obliczono objętościową chłonność wody.

2.4. Stabilność wymiarów

Badanie przeprowadzone zostało zgodnie z normą PN-EN 2796:1986. Przygotowano kształtki o wymiarach 10×10×100 mm. Zważono je z dokładnością do 0,0001 g i zmierzono za pomocą suwmiarki. Tak przygotowane próbki umieszczono w temperaturze 150°C na czas 20 i 40 godz. Po każdorazowym wyjęciu z suszarki kształtki sezonowano przez 30 minut i ponownie ważono i mierzono. Szczególną uwagę poświęcono zmianom długości pianek, jako iż był to ich najdłuższy wymiar.

2.5. Współczynnik przewodzenia ciepła

Badanie współczynnika przewodzenia ciepła wykonano za pomocą przyrządu IZOMET 2114 firmy Applied Precision według normy PN-EN 993-15:2006. Do każdej z badanych pianek poliuretanowych wprowadzono sondę pomiarową i dokonano pomiaru, poprzez przepuszczenie prądu o znanym natężeniu. Na podstawie zmian temperatury rejestrowanych przez urządzenie wyznaczano wartość współczynnika przewodzenia ciepła materiału.

2.6. Indeks tlenowy

Pracę wykonano zgodnie z normą PN-EN ISO 4589-3:2017-06. Do badania przygotowano kształtki o wymiarach 12×12×100 mm, które umieszczono wewnątrz szklanej rury i rozpoczęto analizę poprzez przepuszczenie mieszaniny tlenu i azotu. Następnie pianki podpalano przy użyciu palnika gazowego i odczytując wskazania aparatu, obserwowano zachowanie próbek w celu ustalenia wartości indeksu tlenowego.

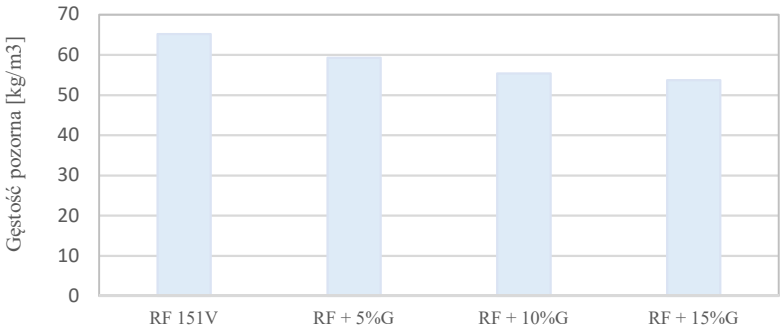
2.7. Test poziomy

Postępowano zgodnie z normą PN-EN ISO 3582:2002. Przygotowano próbki o wymiarach 13×50×150 mm, które zważono i zaznaczono linię na 2,5 cm od góry. Umieszczono je w łapie, a następnie podpalono płomieniem palnika gazowego. Zmierzono czas i drogę palenia, a na podstawie otrzymanych wyników obliczono szybkość palenia i ubytek masy.

3. Część dyskusyjna

Otrzymano pianki poliuretanowe z udziałem suszonych kwiatów goździka i zbadano ich wybrane właściwości fizyczne i ogniowe.

Pierwszym badaniem była gęstość pozorna. Stanowi ona jeden z podstawowych parametrów opisujących strukturę pianek poliuretanowych. Parametr ten wpływa na wiele właściwości użytkowych materiału, w tym wytrzymałość mechaniczną, właściwości termoizolacyjne oraz przebieg procesu spieniania. Wyniki zebrane zostały w sposób graficzny na poniższym wykresie (rys. 1.).

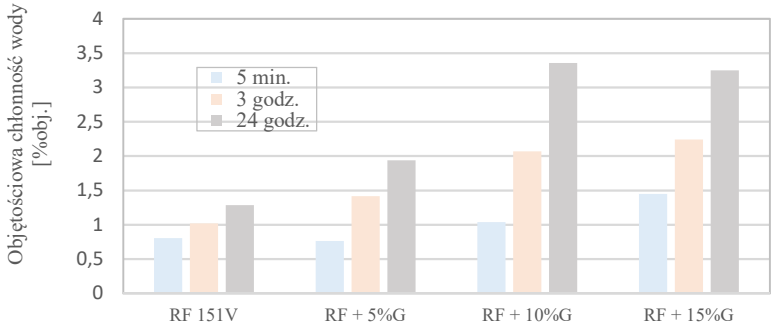


Rys. 1. Gęstość pozorna badanych pianek poliuretanowych w zależności od zawartości suszonych kwiatów goździka

Fig. 1. Apparent density of polyurethane foams as a function of dried clove flower content

Na podstawie uzyskanych wyników stwierdzono, że wprowadzenie suszonych kwiatów goździka wpływa na wartość tego parametru. Największą gęstość pozorną wykazała pianka referencyjna, natomiast wraz ze wzrostem zawartości bio-wypełniacza obserwowano stopniowe zmniejszenie gęstości badanych materiałów. Zjawisko to może być związane z wpływem dodatku roślinnego na przebieg procesu spieniania oraz formowanie struktury komórkowej pianek. Pomimo zauważalnej tendencji spadkowej, uzyskane wartości gęstości nie różnią się znacząco od wartości uzyskanej dla materiału referencyjnego.

Następnie przeprowadzono badanie chłonności wody. Jest to istotny parametr określający odporność materiałów porowatych na działanie wilgoci. W przypadku sztywnych pianek poliuretanowych parametr ten ma szczególne znaczenie ze względu na ich zastosowanie jako materiałów izolacyjnych. Wyniki badań objętościowej chłonności wody przedstawiono na rys. 2.



Rys. 2. Objętościowa chłonność wody badanych pianek poliuretanowych

Fig. 2. Volumetric water absorption of the tested polyurethane foams

Na podstawie przeprowadzonych badań stwierdzono, że chłonność wody pianek poliuretanowych wzrasta wraz ze zwiększaniem zawartości suszonych kwiatów goździka w ich strukturze. Kompozycja zawierająca 5% dodatku wykazała wartości chłonności wody zbliżone do pianki referencyjnej. Natomiast w przypadku pianek zawierających 10% oraz 15% bio-wypełniacza zaobserwowano wyraźny wzrost tego parametru po 24 godzinach ekspozycji w wodzie. Zjawisko to może być związane z higroskopijnym charakterem materiałów roślinnych oraz zaburzeniem ciągłości struktury komórkowej pianek, co sprzyja wnikaniu wody do ich wnętrza.

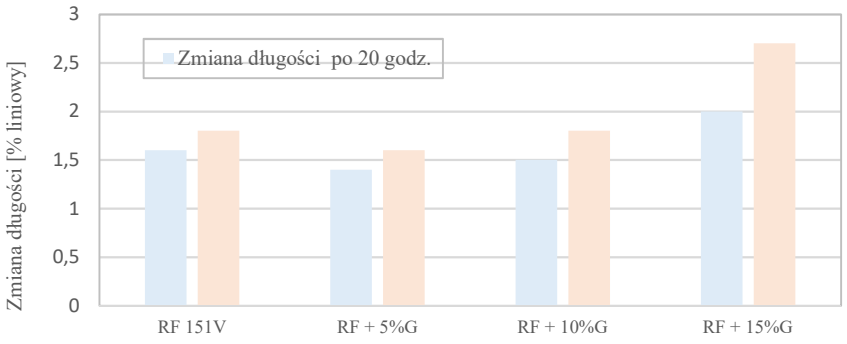
Do badania chłonności wody i gęstości pozornej wykorzystano kształtki przedstawione na rys. 3.



Rys. 3. Próbki pianek PUR wykorzystane do badań gęstości pozornej i chłonności wody

Fig. 3. Polyurethane foam samples used for apparent density and water absorption measurements

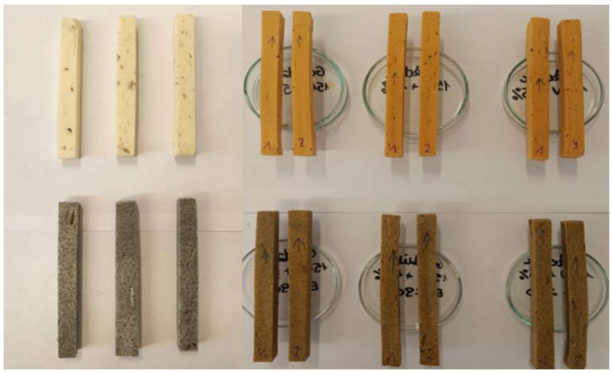
Stabilność wymiarów określa zdolność materiału do zachowania pierwotnych wymiarów podczas działania podwyższonej temperatury. Parametr ten jest szczególnie ważny w przypadku materiałów izolacyjnych eksploatowanych w zmiennych warunkach termicznych. Wyniki zmian długości badanych pianek przedstawiono na rys. 4.



Rys. 4. Zmiana długości badanych pianek poliuretanowych po ekspozycji w temperaturze 150°C

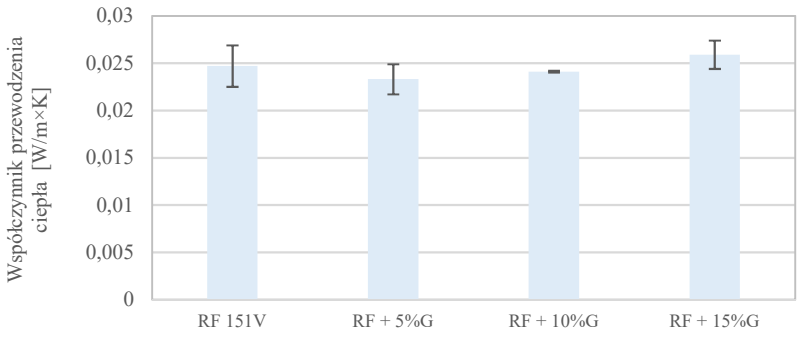
Fig. 4. Change in the length of polyurethane foams after exposure at 150°C

Zauważyć można, że długość pianek zmienia się nieznacznie, a zmiany po 40 godz. są większe niż po 20 godz. Pianka z 5% dodatkiem kwiatów wykazała mniejsze zmiany długości w porównaniu z pianką referencyjną, a pianka z 10% dodatkiem kwiatów bardzo zbliżone do pianki referencyjnej. Stabilność wymiarów pogarsza się wraz ze wzrostem udziału dodatków w strukturze pianek. Na zdjęciu (rys. 5.) przedstawiono wygląd pianek odpowiednio przed i po 40 godzinach ekspozycji w suszarce, gdzie zauważyć można nieznaczne odkształcenia i zmianę barwy.



Rys. 5. Wygląd pianek PUR przed i po 40 godzinach ekspozycji w temperaturze 150°C
Fig. 5. Appearance of polyurethane foams before and after 40 h exposure at 150°C

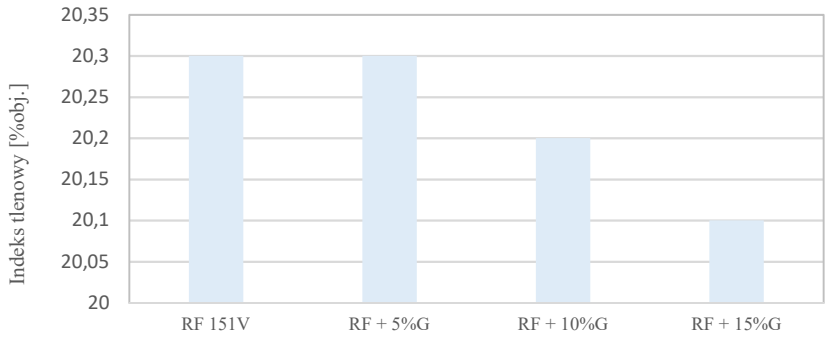
Współczynnik przewodzenia ciepła (λ) jest związany ze strukturą pianek, ich gęstością pozorną oraz składem chemicznym. Wprowadzenie do składu pianek dodatków nieznacznie wpłynęło na zmiany wartości λ, a dokładne wyniki przedstawiono graficznie na rys. 6.



Rys. 6. Współczynnik przewodzenia ciepła badanych pianek poliuretanowych
Fig. 6. Thermal conductivity coefficient of the tested polyurethane foams

Współczynnik przewodzenia ciepła jest jednym z najważniejszych parametrów określających właściwości termoizolacyjne pianek poliuretanowych. Wartość tego parametru dla pianki referencyjnej wyniosła 0,0247 W/m·K. Wprowadzenie niewielkiej ilości dodatku roślinnego (5% mas.) spowodowało nieznaczne obniżenie współczynnika przewodzenia ciepła, co może wskazywać na korzystny wpływ niewielkiej ilości dodatku roślinnego na przebieg procesu spieniania oraz właściwości termoizolacyjne materiału. W przypadku pianek zawierających większe ilości bio-wypełniacza wartości współczynnika przewodzenia ciepła były zbliżone lub nieco wyższe od wartości uzyskanej dla materiału referencyjnego.

Indeks tlenowy jest powszechnie stosowanym parametrem służącym do oceny palności materiałów polimerowych. Określa on minimalne stężenie tlenu w mieszaninie tlenowo-azotowej niezbędne do podtrzymania procesu spalania próbki. Wyniki badań przedstawiono na rys. 7.



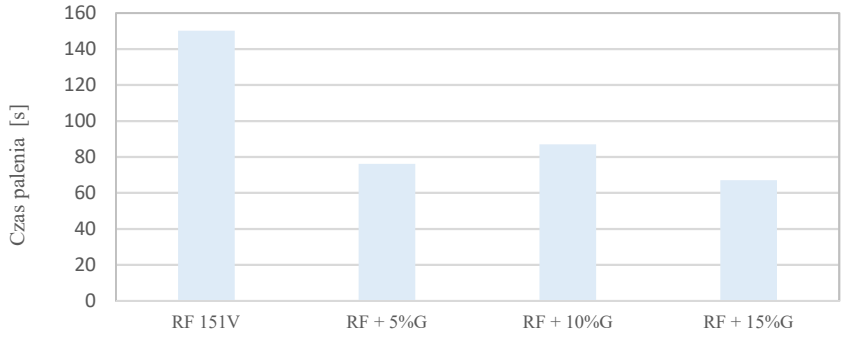
Rys. 7. Indeks tlenowy badanych pianek poliuretanowych

Fig. 7. Limiting oxygen index of the tested polyurethane foams

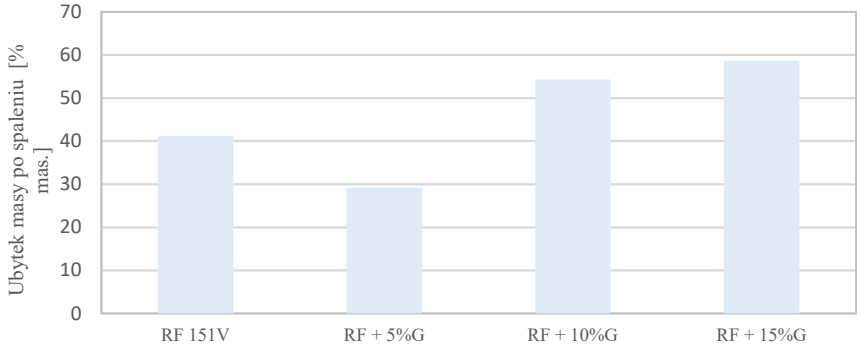
Na podstawie przeprowadzonych badań stwierdzono, że wartość indeksu tlenowego dla pianki referencyjnej wynosiła 20,3%. Wprowadzenie do struktury pianek dodatku suszonych kwiatów goździka nie spowodowało istotnych zmian tego parametru. Wraz ze wzrostem zawartości bio-wypełniacza obserwowano jedynie nieznaczny spadek wartości indeksu tlenowego. Uzyskane wyniki wskazują, że zarówno pianka referencyjna, jak i kompozycje zawierające dodatek roślinny należą do materiałów palnych.

Kolejnym i ostatnim badaniem właściwości otrzymanych pianek był test poziomy czyli test poziomego palenia HB. Umożliwia ocenę zachowania materiału podczas spalania oraz określenie takich parametrów jak czas spalania, szybkość rozprzestrzeniania płomienia i ubytek masy. Badanie to pozwala na porównanie wpływu zastosowanych modyfikatorów na właściwości ogniowe pianek poliuretanowych. Wyniki przedstawiono na rys. 8.-10.

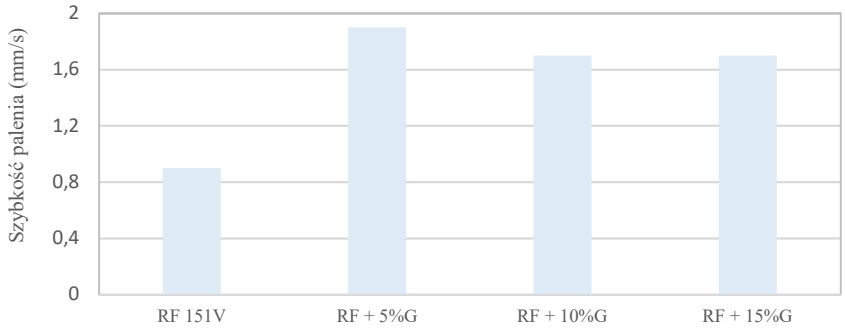
Wszystkie pianki spaliły się w całości. Czas palenia pianki referencyjnej to 150 sekund, natomiast czas palenia pianek z dodatkiem suszonych kwiatów goździka mieścił się w zakresie 67-87 sekund (rys. 8.). Ubytek masy pianki referencyjnej wyniósł 41,1% mas., a jej szybkość palenia 0,87 mm/s. Pianka, do której dodano 5% kwiatów pomimo szybszego palenia wykazała mniejszy ubytek masy w porównaniu z pianką referencyjną i podczas jej palenia można było zauważyć tylko nieznaczne wygięcie. Pianki z 10 i 15% udziałem kwiatów wykazały zarówno większy ubytek masy jak i szybkość palenia w odniesieniu do kompozycji referencyjnej (rys. 9. i 10.).



Rys. 8. Czas spalania badanych pianek poliuretanowych w teście poziomym
Fig. 8. Burning time of polyurethane foams in the horizontal flammability test



Rys. 9. Ubytek masy pianek poliuretanowych po teście poziomym
Fig. 9. Mass loss of polyurethane foams after the horizontal burning test



Rys. 10.
Fig. 10.

4. Podsumowanie

Przeprowadzone badania miały na celu określenie wpływu dodatku suszonych kwiatów goździka na właściwości fizyczne oraz ogniowe sztywnych pianek poliuretanowych. W ramach pracy otrzymano serię materiałów obejmującą piankę referencyjną oraz kompozycje zawierające 5, 10 i 15% masowych dodatku bio-wypełniacza pochodzenia roślinnego.

Wytworzone materiały poddano analizie obejmującej wybrane właściwości fizyczne oraz właściwości palne. Oznaczono gęstość pozorną, chłonność wody, stabilność wymiarów w podwyższonej temperaturze, współczynnik przewodzenia ciepła, a także indeks tlenowy oraz zachowanie materiałów podczas spalania w teście poziomym.

Na podstawie uzyskanych wyników stwierdzono, że dodatek suszonych kwiatów goździka wpływał na przebieg procesu spieniania oraz wybrane właściwości fizyczne pianek poliuretanowych. Zakres zaobserwowanych zmian zależał od zawartości bio-wypełniacza w materiale. Wraz ze wzrostem zawartości dodatku roślinnego obserwowano stopniowe zmniejszenie gęstości pozornej pianek oraz wzrost ich chłonności wody. Zjawisko to można tłumaczyć obecnością naturalnego bio-wypełniacza o rozwiniętej strukturze komórkowej oraz jego wpływem na przebieg procesu spieniania.

Analiza stabilności wymiarów wykazała, że niewielki dodatek kwiatów goździka nie powoduje istotnych zmian tego parametru, natomiast przy większych zawartościach dodatku obserwuje się pogorszenie stabilności materiału w podwyższonej temperaturze. Badania współczynnika przewodzenia ciepła wskazały, że niewielki udział bio-wypełniacza może korzystnie wpływać na właściwości termoizolacyjne pianek, jednak przy większych zawartościach dodatku efekt ten ulega osłabieniu.

W zakresie właściwości ogniowych stwierdzono, że dodatek suszonych kwiatów goździka nie powoduje istotnych zmian wartości indeksu tlenowego badanych

pianek, a wszystkie analizowane materiały należą do grupy materiałów palnych. Jednocześnie w teście poziomym zaobserwowano skrócenie czasu spalania oraz zwiększenie szybkości palenia pianek zawierających większe ilości bio-wypełniacza.

Uzyskane wyniki wskazują, że suszone kwiaty goździka mogą stanowić interesujący naturalny dodatek modyfikujący właściwości pianek poliuretanowych. Zastosowanie tego typu bio-wypełniaczy może stanowić jeden z kierunków rozwoju materiałów polimerowych o częściowo odnawialnym charakterze, jednak ich udział w kompozycji powinien być odpowiednio dobrany w celu zachowania korzystnych właściwości użytkowych materiału.

Literatura

[1] Bourbigot S., Duquesne S., Fire retardant polymers: recent developments and opportunities, Journal of Materials Chemistry, 2007, 17, pp. 2283–2300.
[2] Członka S., Strąkowska A., Strzelec K., Natural fillers in polyurethane foams: A review, Journal of Cellular Plastics, 2019, 55, pp. 1–28.
[3] Hull T., Stec A., Fire Retardancy of Polymeric Materials, CRC Press, Boca Raton 2009.
[4] Kurańska M., Ryszkowska J., Polyurethane foams modified with natural fillers, Materials,
[5] Levchik S., Weil E., Flame retardancy of polyurethane foams, Polymer International, 2004,
[6] Modesti M., Lorenzetti A., Flame retardancy of polyurethane foams, Polymer Degradation and Stability, 2002, 77, pp. 195–202.
[7] Prociak A., Poliuretany i ich zastosowanie w materiałach izolacyjnych, Wydawnictwo PK,
[8] Prociak A., Rokicki G., Ryszkowska J., Chemia i technologia poliuretanów, PWN, Warszawa
[9] Sonnenschein M., Polyurethanes: Science, Technology, Markets and Trends, Wiley, 2015.

EFFECT OF DRIED CLOVE FLOWER ADDITION ON THE PHYSICAL AND FLAMMABILITY PROPERTIES OF RIGID POLYURETHANE FOAMS

Summary

The study presents the results of research on the influence of dried clove flower additives on the physical and fire properties of rigid polyurethane foams. A series of compositions was prepared, including a reference material and compositions containing 5, 10 and 15 wt.% of a plant-based bio-filler. The obtained materials were analyzed in order to determine the influence of the plant additive on the foaming process and the cellular structure of polyurethane foams.

For the prepared materials, apparent density, water absorption, dimensional stability and thermal conductivity coefficient were determined. In addition, properties related to the behaviour of the materials under flame exposure were evaluated, including the limiting oxygen index and the horizontal burning test.

The incorporation of the plant additive affected the foaming process and the cellular structure of the foams, which was reflected in changes in their functional properties. With increasing

content of clove flowers, a decrease in the density of the materials and an increase in their water absorption were observed, which may be associated with the porous structure of the bio-filler and changes in the morphology of the foam cells.

The influence of clove flower addition on the thermal conductivity of the foams was also analysed, as this parameter is important in the context of potential insulation applications. The obtained results indicate that dried clove flowers may serve as a potential natural modifier of polyurethane foam properties. The use of this type of additive may contribute to the modification of the physical properties of polyurethane materials and may represent an interesting direction for the development of polymer composites containing natural components.

Keywords: polymer composites, plant-based fillers, apparent density, water absorption, thermal conductivity.

Natalia Klocek, Julia Ziąbek, Maja Jaszowska
Dominik Dworak, Sebastian Cwynar
Koło Naukowe Elektroniki i Technologii Informacyjnych

dr inż. Bartosz Pawłowicz
opiekun Koła Naukowego

WYKORZYSTANIE KWANTYZACJI WEKTORÓW UCZĄCYCH DO ANALIZY I KLASYFIKACJI DANYCH EWALUACYJNYCH ASYSTENTÓW NAUCZANIA

Streszczenie

Celem jest przedstawienie procesu projektowania oraz implementacji sztucznej sieci neuronowej wykorzystującej algorytm kwantyzacji wektorów uczących (Learning Vector Quantization – LVQ) w zadaniu klasyfikacji danych. Przedmiotem badań był zbiór *„Teaching Assistant Evaluation"*, zawierający dane dotyczące oceny efektywności asystentów nauczania na podstawie atrybutów takich jak status rodzimego użytkownika języka, instruktor, kurs, wielkość klasy czy typ semestru. w części teoretycznej pracy przybliżono budowę neuronu McCullocha-Pittsa oraz szczegółowo scharakteryzowano mechanizmy działania wariantów algorytmu: LVQ1, LVQ2, LVQ2.1 i LVQ3, kładąc nacisk na różnice w metodach aktualizacji wektorów prototypowych. Część badawcza obejmuje implementację modelu w języku Python oraz przeprowadzenie serii eksperymentów walidacyjnych z użyciem techniki kroswalidacji. Kluczowym elementem badań była wieloaspektowa analiza wpływu doboru parametrów, w tym współczynnika uczenia, tempa jego degradacji oraz liczby epok, na zdolności generalizacyjne sieci. Wykazano, że dla analizowanego zbioru ewaluacyjnego maksymalna uzyskana dokładność wyniosła około 56%, co sugeruje trudności związane z subiektywnym charakterem danych wejściowych. Aby zweryfikować poprawność implementacji, przeprowadzono testy porównawcze na standardowym zbiorze *„Iris"*, osiągając wynik 94%, co potwierdziło skuteczność algorytmu dla danych o wyraźnej separacji klastrów. Podjęto również próbę zmiany atrybutu decyzyjnego, co jednak skutkowało problemem silnie niezbalansowanych klas. Wnioski końcowe podkreślają, że metoda LVQ jest wysoce wrażliwa na jakość i strukturę danych oraz wymaga precyzyjnego dostrojenia parametrów, a jej skuteczność maleje w przypadku danych o zatartej granicy decyzyjnej.

Słowa kluczowe: sieci neuronowe, LVQ, klasyfikacja, uczenie nadzorowane, analiza danych.

1. Wstęp

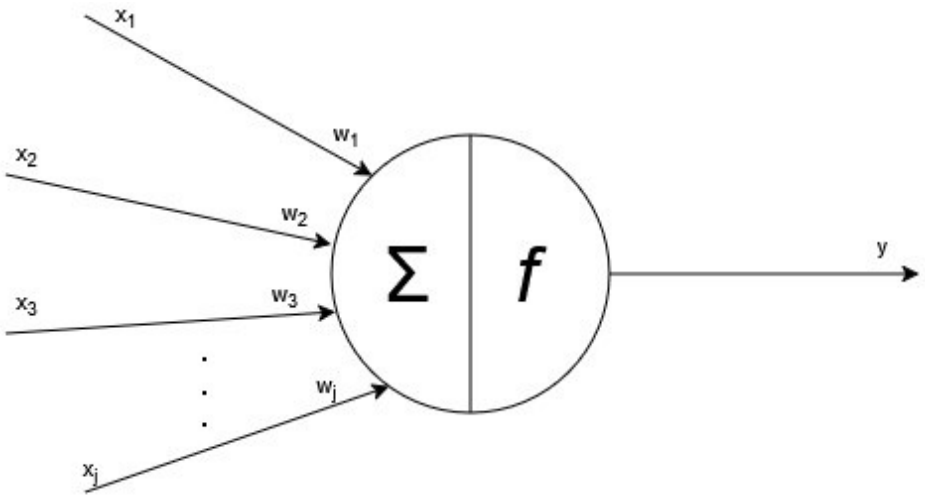
Celem projektu było stworzenie sieci neuronowej wykorzystującej algorytm kwantyzacji wektorów uczących (LVQ) w zadaniu klasyfikacji, aby ocenić jaka była ocena efektywności kursów na podstawie ich cech.

Prace nad projektem rozpoczęto od zapoznania się z teoretycznymi podstawami algorytmu LVQ oraz zagadnieniem klasyfikacji. Następnie przygotowano dane wejściowe oraz przeanalizowano dostarczony skrypt, który posłużył do implementacji modelu. w ostatnim etapie przeprowadzono serię eksperymentów, podczas których oceniano dokładność działania modelu, modyfikując jego wybrane parametry w celu uzyskania jak najlepszej dokładności klasyfikacji.

2. Wstęp teoretyczny

2.1. Model neuronu

Neuron McCullocha-Pittsa jest podstawową jednostką przetwarzania w sztucznych sieciach neuronowych. Jest to funkcja matematyczna, która przyjmuje wiele wejść wraz z wagami i na ich podstawie oblicza wyjście. Sieć neuronowa składa się z połączonych neuronów, podzielonych na warstwy, które współpracują w przetwarzaniu informacji i rozwiązywaniu problemów na podstawie dostarczonych danych.



Rys. 1. Model neuronu McCullocha-Pittsa

Fig. 1. McCulloch-Pitts neuron model

Źródło: Opracowanie własne.

Jak przedstawia rys. 1. neuron składa się z wielu wejść, którymi są dane wchodzące do bloku sumującego, wraz z wagami, które określają poziom ważności wejść. Blok sumujący jest odpowiedzialny za przyjęcie danych wejściowych wraz z ich wagami i przetworzenie ich za pomocą sumy ważonej, co później przekazuje do funkcji aktywacji. Funkcja aktywacji pobiera te dane i na podstawie swojego przebiegu przydziela odpowiednią wartość wyjściową dla otrzymanego argumentu. Sygnał wyjściowy neuronu *y* określony jest zależnością:

$$y = f\left(\sum_{j=1}^{L} w_j \cdot x_j + b\right) \tag{1}$$

gdzie: x_j – j-ty sygnał wejściowy, w_j – waga j-tego sygnału wejściowego, L – liczba sygnałów wejściowych, b – bias (liczba stała).

Powyższe równanie można uprościć, stosując zapis macierzowy:

$$\boldsymbol{x} = [x_1, x_2, \ldots, x_L]^T \tag{2}$$

$$\boldsymbol{w} = [w_1, w_2, \ldots, w_L]^T \tag{3}$$

Wtedy **x** jest wektorem sygnałów wejściowych, w jest macierzą wierszową wag, natomiast y i b są skalarami. w ten sposób można otrzymać równanie:

$$y = f(\mathbf{wx} + b) \tag{4}$$

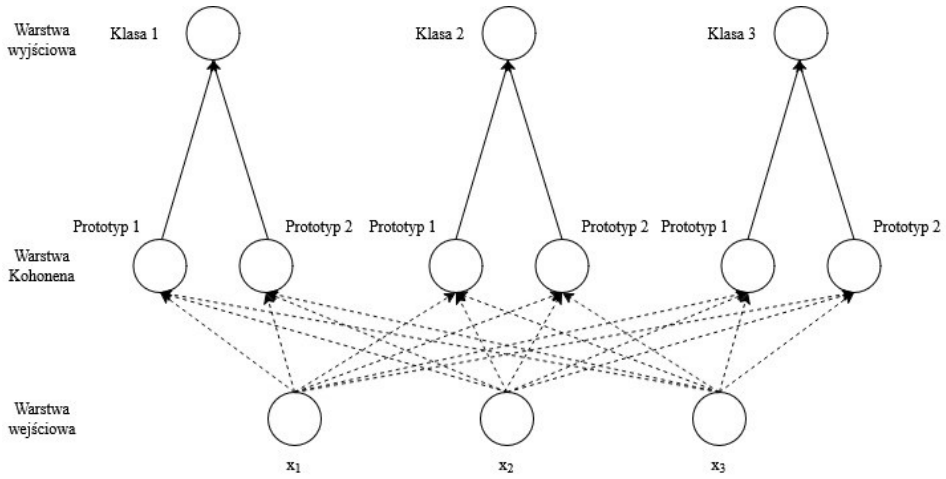
2.2. Learning Vector Quantization

Learning Vector Quantization (LVQ) to opracowana przez Teuvo Kohonena metoda uczenia oparta na prototypach, wykorzystywana w zadaniach klasyfikacyjnych w ramach uczenia nadzorowanego. Choć sam algorytm nie należy do najbardziej zaawansowanych pod względem skuteczności, jego zaletą jest prostota, intuicyjność oraz możliwość rozbudowy o bardziej złożone warianty.

Podczas procesu uczenia sieci LVQ dla każdego sygnału wejściowego obliczana jest odległość do wszystkich neuronów w warstwie i wybierany jest neuron leżący najbliżej. Jeśli ten neuron reprezentuje tę samą klasę co sygnał wejściowy, jego wektor wag jest modyfikowany w taki sposób, aby zbliżył się do danego sygnału. w przeciwnym przypadku, jeśli neuron zwycięzca reprezentuje inną klasę niż dane wejściowe, jego wektor wag zostaje oddalony od sygnału. Dzięki temu sieć uczy się, jak najlepiej rozróżniać klasy na podstawie dystansu w przestrzeni cech. Każdy neuron w sieci jest przypisany do konkretnej klasy, co oznacza, że znajduje się w obszarze decyzyjnym odpowiadającym tej kategorii. Podczas etapu testowania, dla nowego sygnału wejściowego obliczane są odległości do wszystkich neuronów, a wynik przypisuje się na podstawie klasy, do której należy neuron znajdujący się najbliżej.

W podstawowej wersji algorytmu LVQ do obliczania odległości między wektorem wejściowym a wagami neuronów stosowana jest metryka euklidesowa. Odległość ta jest liczona osobno dla każdego neuronu w sieci i stanowi podstawę do podjęcia decyzji o tym, który neuron zwycięża i jak należy zaktualizować jego wektory wagowe.

Algorytm LVQ występuje w kilku odmianach, takich jak LVQ 1, LVQ 2, LVQ 2.1 czy LVQ 3, z których każda różni się podejściem do inicjalizacji wektorów prototypowych oraz metodą aktualizacji wag w trakcie uczenia. Poszczególne wersje wprowadzają własne modyfikacje w mechanizmie uczenia, co wpływa na ich skuteczność i predyspozycje do rozwiązywania różnych typów problemów klasyfikacyjnych. Wybór konkretnego wariantu LVQ powinien być uzależniony od charakterystyki problemu oraz rodzaju i struktury danych, z którymi pracuje model.



Rys. 2. Architektura sieci LVQ

Fig. 2. LVQ network architecture

Źródło: Opracowanie własne.

Struktura sieci LVQ przedstawionej na rys. 2 składa się z:

1. Wektorów wejściowych – Opisują dane wejściowe.
2. Warstwy ukrytej (Kohonena) – Neurony reprezentują prototypy (czyli przyjęte wzorce). Każdy z prototypów jest przypisany do swojej konkretnej klasy i „rywalizuje" o dane wejściowe poprzez mierzenie odległości metryką euklidesową. Najbliższy prototyp jest neuronem „zwycięskim".
3. Warstwy wyjściowej – Przyporządkowuje wektory wejściowe do klasy, którą reprezentuje neuron "zwycięski".

2.3. LVQ 1

LVQ 1 jest pierwszym i najbardziej podstawowym wariantem algorytmu LVQ. Klasyfikuje dane na podstawie porównania wektora wejściowego z wektorami prototypowymi. Oblicza się odległość punktu treningowego do wszystkich prototypów miarą euklidesową, a następnie wybiera się najbliższy prototyp. w zależności, czy zarówno wektor wejściowy, jak i wektor prototypowy należą do tej samej klasy, wagi są odpowiednio zmieniane. Jeśli należą do tej samej klasy, to prototyp jest przesuwany bliżej punktu treningowego. Natomiast jeśli nie należą do tej samej klasy, to jest on odsuwany.

Wzór na obliczenie odległości:

$$d_i = \|w_i - x\| = \sqrt{\sum_{j=1}^{N}(w_{ij} - x_j)^2} \tag{5}$$

gdzie: d_i – odległość pomiędzy wektorem x a i-tym prototypem w_i, w_i – w ektor prototypów dla neuronu i, x – wektor wejściowy, N – liczba cech, w_{ij} – wartość j-tej cechy i-tego prototypu, x_j – wartość j-tej cechy wektora wejściowego.

Wektor wag, który zwyciężył oraz poprawnie zakwalifikował punkt treningowy, jest aktualizowany za pomocą wzoru:

$$w_i(t+1) = w_i(t) + \alpha(x - w_i(t)) \tag{6}$$

Natomiast, gdy punkt treningowy został niepoprawnie zaklasyfikowany, to wektor wagowy aktualizuje się za pomocą wzoru:

$$w_i(t+1) = w_i(t) - \alpha(x - w_i(t)) \tag{7}$$

gdzie: t – bieżąca epoka, w_i – wektor wagi w danej epoce, x – wartość wektora wejściowego, α – współczynnik uczenia.

Współczynnik uczenia może być zmniejszany za pomocą różnych wzorów. Wzorem wykorzystywanym w projekcie jest:

$$\alpha(t) = \alpha_0\beta \tag{8}$$

gdzie: t – bieżąca epoka, α_0 – początkowy współczynnik uczenia (Wybierany ręcznie), β – stała kontrolująca tempo spadku (Wybierana ręcznie). Przykładowo znane są również inne wzory, w tym zmniejszanie współczynnika uczenia zależne od obecnej epoki oraz maksymalnej ilości epok:

$$\alpha(t) = \alpha_0\left(1 - \frac{t}{T}\right) \tag{9}$$

gdzie: t – bieżąca epoka, α_0 – początkowy współczynnik uczenia (Wybierany ręcznie), T – maksymalna liczba epok (Wybierana ręcznie).

Lub współczynnik uczenia może też być zmniejszany logarytmicznie za pomocą wzoru:

$$\alpha(t) = \frac{\alpha_0}{1+\beta \cdot t} \tag{10}$$

gdzie: t – bieżąca epoka, α0 – początkowy współczynnik uczenia (Wybierany ręcznie), β – stała kontrolująca tempo spadku (Wybierana ręcznie).

2.4. LVQ 2

W algorytmie LVQ 2 wprowadzono zmiany mające na celu poprawę jakości klasyfikacji. Zamiast jednego najbliższego prototypu znajdowane są dwa, które następnie są aktualizowane tylko w określonym przypadku. Aktualizacja wektorów prototypowych zachodzi w momencie, gdy najbliższy prototyp ma inną klasę od wektora wejściowego, a drugi najbliższy prototyp ma taką samą klasę. Do tego punkt wejściowy musi leżeć w oknie określonym wzorem:

$$min(\frac{d_1}{d_2}, \frac{d_2}{d_1}) > \frac{1-\varepsilon}{1+\varepsilon} \tag{11}$$

gdzie: d_1, d_2 – odległości od dwóch najbliższych wektorów prototypowych, ε – stały parametr okna (Wybierany ręcznie).

Wprowadzono okno, aby ograniczyć aktualizację wag tylko do punktów leżących blisko granicy między klasami. Jeśli warunki są spełnione, prototyp z tą samą klasą co dane wejściowe jest przesuwany bliżej, a ten drugi dalej.

2.5. LVQ 2.1

Algorytm LVQ 2.1 działa bardzo podobnie do LVQ 2. Zachowane jest okno obliczane tym samym wzorem. Natomiast tym razem wystarczy spełnić jeden z dwóch przypadków:

1. Kiedy klasa wejścia jest równa klasie najbliższego prototypu i nie jest równa klasie drugiego najbliższego prototypu.
2. Kiedy klasa wejścia jest równa klasie drugiego najbliższego prototypu i nie jest równa klasie najbliższego prototypu.

Wtedy prototyp z tej samej klasy co punkt wejściowy zostaje aktualizowany, aby przybliżyć go do wejścia, a ten z innej klasy jest oddalany.

2.6. LVQ 3

LVQ 3 działa na podobnej zasadzie co poprzednie algorytmy. w przeciwieństwie do nich, ten wariant aktualizuje oba wektory prototypowe, nawet gdy oba mają tę samą klasę, jeśli dobrze zaklasyfikowały dane wejściowe. Wprowadza on również dodatkowy parametr, który kontroluje siłę aktualizacji. w przypadku, gdy oba prototypy wskazują na poprawną klasę, wzór na aktualizację wektorów wagowych wygląda następująco:

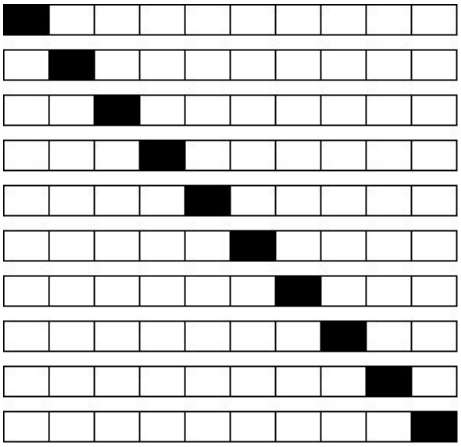
$$w_i(t+1) = w_i(t) + \varepsilon\alpha(x - w_i(t)) \quad (12)$$

gdzie: t – bieżąca epoka, w_i – wektor wagi w danej epoce, x – wartość wektora wejściowego, α – współczynnik uczenia, ε – stała stabilizująca (Wybierana ręcznie).

Jeśli są dwa poprawne prototypy, to wzór jest wykorzystywany do zaktualizowania wag na każdym z nim, w przeciwnym wypadku używany jest tylko na jedynym poprawnym prototypie, lub wcale jeśli żaden prototyp nie zaklasyfikował dobrze punktu treningowego.

2.7. Kroswalidacja

Kroswalidacja jest techniką oceny skuteczności modelu uczącego się, która pozwala sprawdzić, jak dobrze będzie działał na nieznanych danych. Polega ona na podzieleniu całego zbioru danych na k równych części (foldy), a następnie trenowaniu modelu k razy, za każdym razem wykorzystując jeden z foldów jako dane testowe i resztę k − 1 części jako dane treningowe. Średnia z wyników wszystkich przeprowadzonych prób daje ocenę skuteczności. Sprawia to, że wolniej dostaje się wyniki, ze względu na testowanie modelu kilka razy. Natomiast zalet kroswalidacji jest o wiele więcej niż wad. Pozwala ona na uzyskanie realistycznej oceny jakości modelu, dzięki temu, że skuteczność nie zależy od przypadkowego podziału zbioru danych na treningowy oraz testowy. Poprzez wykorzystanie każdej z części na testy i resztę na trening wykorzystywany jest cały zbiór danych. Można też łatwo porównać wyniki z innymi modelami, ponieważ w ten sposób każdy model ma takie same warunki testowe.



Rys. 3. Schemat podziału zbioru danych na zbiory uczące i testowe w iteracyjnym procesie 10-krotnej walidacji krzyżowej

Fig. 3. Scheme of dividing the dataset into training and test sets in the iterative process of 10-fold cross-validation

Źródło: Opracowanie własne.

Rysunek 3. przedstawia przykład kroswalidacji przy podzieleniu zbioru danych na 10 części. Kolor czarny oznacza dane testowe, natomiast kolor biały dane treningowe. Ponieważ jest 10 foldów, model wykonał 10 przebiegów, każdy na innych danych testowych i treningowych.

3. Analiza danych

Jako dane uczące został wykorzystany zbiór „Teaching Assistant Evaluation". Zbiór zawiera oceny efektywności pracy asystentów nauczania zebrane w ciągu trzech semestrów regularnych i dwóch letnich na Wydziale Statystyki Uniwersytetu Wisconsin-Madison. Zawiera on 151 rekordów, każdy po 6 atrybutów. Parametrem wyjściowym jest ostatni atrybut. Dane są złożone z następujących atrybutów:

1. Czy asystent jest rodzimym użytkownikiem języka angielskiego (wartości binarne: 1 – tak, 2 – nie).
2. Identyfikator instruktora prowadzącego kurs (wartości kategoryczne, 25 kategorii).
3. Identyfikator kursu (wartości kategoryczne, 26 kategorii).
4. Typ semestru (wartości binarne: 1 – letni, 2 – regularny).
5. Wielkość klasy (wartości numeryczne)
6. Ocena efektywności zajęć (wartości kategoryczne: 1 – niska, 2 – średnia, 3 – wysoka).

Przykładowy rekord: 1,23,3,1,19,3

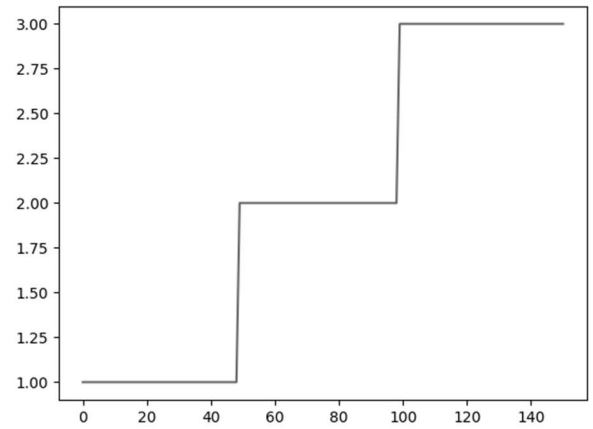
1. Asystent jest rodzimym użytkownikiem języka angielskiego.
2. Identyfikator instruktora prowadzącego kurs wynosi 23.
3. Identyfikator kursu wynosi 3.
4. Semestr letni.
5. Wielkość klasy ma wartość 19.
6. Ocena efektywności zajęć jest wysoka.

4. Opis implementacji

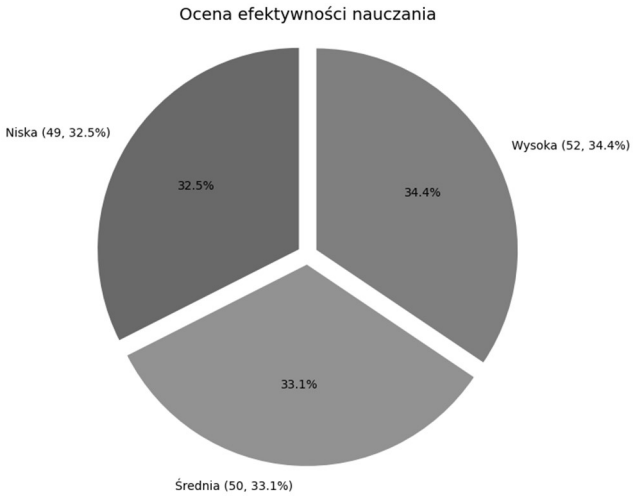
W celu zachowania przejrzystości i zwartości niniejszego rozdziału naukowego, zrezygnowano z umieszczania surowych fragmentów kodu źródłowego w tekście głównym. Pełna implementacja skryptów przygotowania danych, funkcji uczących oraz procedur walidacyjnych może zostać udostępniona zainteresowanym badaczom na bezpośrednią prośbę, po kontakcie mailowym z autorami.

Program został zrealizowany w języku Python. Wykorzystano biblioteki numpy do obliczeń numerycznych, hickle do zapisu danych oraz matplotlib do wizualizacji. Dane wczytano z pliku tekstowego. Zmienna x przechowuje transponowaną macierz danych wejściowych, która obejmuje wszystkie kolumny, oprócz pierwszej, która jest zbiorem etykiet. Dane zrzutowano na liczby zmiennoprzecinkowe i znormalizowano do przedziału [-1;1]. Normalizacja jest krytyczna dla

metod opartych na miarach odległości. Weryfikowana jest poprawność normalizacji, poprzez sprawdzenie zakresu wartości. Przetworzone dane, zarówno oryginalne, jak i znormalizowane, zapisano w formacie hkl. Opisana procedura stanowi niezbędny etap przygotowawczy przed zastosowaniem algorytmu LVQ, zapewniając odpowiedni format danych oraz ich normalizację, co jest krytyczne dla skuteczności metod opartych na miarach odległości.



Rys. 4. Rozkład etykiet klasowych po procesie sortowania indykujący strukturę zbioru TAE
Fig. 4. Label distribution after sorting indicating the structure of the TAE dataset



Rys. 5. Rozkład klas w zbiorze danych
Fig. 5. Class distribution in the dataset

Jak zilustrowano na rys. 5. w zbiorze danych liczących 151 próbek jest 49 kursów (32.5%), które mają niską ocenę efektywności, 50 (33.1%) kursów z średnią oceną oraz 52 (34.4%) z wysoką. Widać, że w każdej klasie jest po prawie tyle samo próbek, co wpłynie na wiarygodne prowadzenie eksperymentów.

Do prawidłowego działania struktury eksperymentu wykorzystano moduł StratifiedKFold z pakietu sklearn służący do podziału danych z zachowaniem proporcji klas, bibliotekę numpy do obliczeń numerycznych oraz funkcję sqrt do obliczania pierwiastków. Do wydajnego odczytu i zapisu dużych tablic użyto biblioteki hickle, natomiast za wizualizację danych i tworzenie wykresów odpowiada biblioteka matplotlib.

Funkcja lvq_train przyjmuje następujące dane wejściowe: X (dane treningowe), y (etykiety klas), a (współczynnik uczenia), b (współczynnik zmniejszania a w każdej epoce), max_ep (maksymalna liczba epok), min_a (minimalna wartość współczynnika uczenia) oraz e (parametr epsilon dla aktualizacji wag).

Na początku należy zainicjalizować wektory prototypowe. Wybierane są pierwsze unikalne wystąpienia każdej klasy. Następnie tworzy się nową tablicę, w której będą dane, które nie zostały wykorzystane do wektorów prototypowych. Dzięki temu zostały przygotowane próbki, w których nie ma danych wybranych na prototypy. Inicjalizowany zostaje licznik epok, który jest później wykorzystany w pętli, która zakończy się dopiero, gdy epoki przekroczą maksymalną liczbę epok (max_ep) lub gdy parametr a będzie poniżej minimalnego progu a (min_a).

Funkcja w pętli iteruje po wszystkich próbkach treningowych, które zostały wcześniej przygotowane. Obliczana jest odległość euklidesowa między bieżącą próbką a każdym prototypem. Do zmiennych zapisywane są odległości do dwóch najbliższych wektorów wagowych oraz ich indeksy. Następuje sprawdzenie, czy najbliższy prototyp ma tę samą klasę co sprawdzana próbka oraz czy najbliższy oraz drugi najbliższy prototyp mają różne klasy. W przypadku, gdy warunek jest prawdziwy, to najbliższy prototyp zostaje przybliżony do bieżącej próbki poprzez dodanie wektora różnicy przeskalowanego przez współczynnik uczenia a.

W innym przypadku, jeżeli najbliższy i drugi najbliższy prototyp należą do innych klas oraz sprawdzana próbka należy do tej samej klasy co drugi najbliższy prototyp, to zaktualizowane zostaną oba wektory prototypowe. Wcześniej jednak należy sprawdzić, czy minimalny stosunek odległości jest większy od szerokości okna. Jeżeli warunki zostaną spełnione, to poprawny prototyp zostanie przybliżony, a prototyp z inną klasą niż próbka zostanie oddalony. Aktualizacja odbędzie się tak samo jak w poprzednim przypadku, natomiast przy drugim prototypie przeskalowany wektor odległości zostanie odjęty, zamiast dodany.

Jeżeli żaden z poprzednich warunków nie został spełniony, to funkcja przechodzi do ostatniego z warunków. Jeżeli oba najbliższe prototypy mają taką samą klasę co próbka treningowa, to oba prototypy są przybliżane do próbki. Tym razem wektor różnicy jest skalowany nie tylko przez współczynnik uczenia, ale również przez epsilon e, który jest też używany w wyliczaniu szerokości okna. Na koniec

pętli współczynnik uczenia zostaje zmniejszony przez przemnożenie z współczynnikiem zmniejszania b oraz zostaje zwiększony licznik epok. Gdy pętla się zakończy, funkcja zwraca wytrenowane wektory wagowe oraz ich etykiety klas.

Funkcja lvq_test ta ma następujące dane wejściowe: wektor cech pojedynczej próbki testowej x, którą należy zaklasyfikować (jest to jednowymiarowa tablica zawierająca wartości poszczególnych cech danej próbki) oraz wyjście funkcji lvq_train (W), które jest krotką zawierającą macierz wektorów prototypowych i tablicę etykiet klas.

Funkcja najpierw rozpakowuje krotkę w na macierz wektorów prototypowych oraz wektor etykiet klas c. Dla każdego wektora prototypowego funkcja oblicza odległość euklidesową między tym prototypem a wektorem testowym x. Wynikiem jest lista odległości d między wektorem x a każdym z prototypów. Następnie funkcja znajduje indeks najbliższego prototypu do próbki testowej i zwraca etykietę klasy odpowiadającą temu najbliższemu prototypowi. Funkcja ta implementuje regułę najbliższego sąsiada, przypisując nowej próbce etykietę klasy tego prototypu, który jest do niej najbliżej w sensie odległości euklidesowej. Jest to kluczowy element procesu predykcji w algorytmie LVQ.

Za pomocą biblioteki hickle wczytywane są dane z pliku tae.hkl, w którym są oryginalne i znormalizowane dane wejściowe oraz etykiety klas. Posortowane znormalizowane dane wejściowe oraz etykiety klas zostały dodane do pliku, aby stworzyć wykresy, pokazujące między innymi rozkład klas. Sprawdzane jest, czy minimalna wartość etykiety klas ma wartość 0 i w przeciwnym wypadku ta minimalna wartość jest odejmowana od wszystkich wartości, aby etykiety klas zaczynały się od 0. Następnie znormalizowane dane są transponowane, aby wiersze oznaczały próbki, a kolumny pojedyncze cechy. Tablica etykiet klas jest zamieniana na jednowymiarową tablicę.

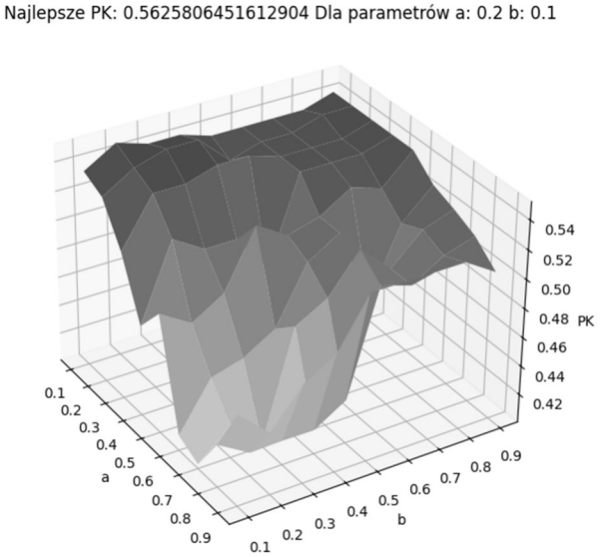
Ustawiana jest liczba foldów do walidacji krzyżowej (w eksperymentach będą wykorzystywane wartości 5, 10 oraz 20 dla tej zmiennej). Tworzony jest obiekt StratifiedKFold, który dzieli dane na części, które zostały zainicjalizowane wcześniej, i zachowuje oryginalny rozkład klas w każdej części. Dzięki temu kroswalidacja będzie zbalansowana. Zaczyna się pętla, która iteruje przez wszystkie wcześniej wyznaczone części danych. Model LVQ jest trenowany na wyznaczonych danych treningowych z określonymi parametrami, które wybiera osoba testująca model. Wybór odpowiednich parametrów jest bardzo ważnym elementem, który zostanie poruszony w eksperymentach. Na koniec obliczana jest dokładność klasyfikacji dla bieżącego foldu jako stosunek poprawnie sklasyfikowanych próbek do całkowitej liczby próbek testowych. Wyświetlane są informacje o klasyfikacji w każdym foldzie oraz jako ostatnie wyświetlana jest średnia dokładność klasyfikacji z wszystkich części.

5. Eksperymenty

5.1. Dobór najlepszego zestawu parametrów

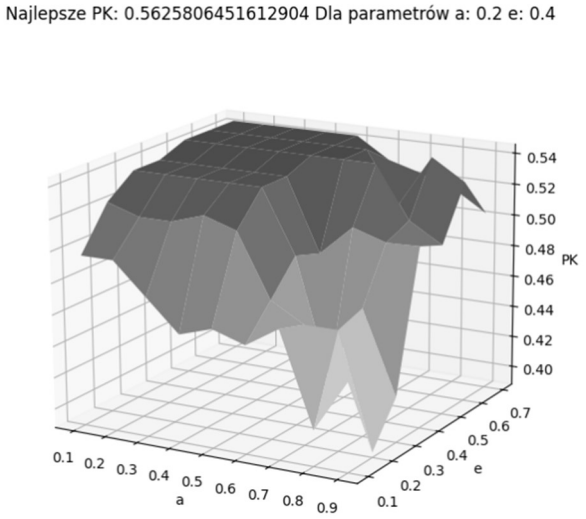
Najważniejszym aspektem w algorytmie LVQ jest odpowiednie wybranie zestawu parametrów. Zmiana nawet jednego z nich może skutkować znacznym zwiększeniem lub zmniejszeniem otrzymanej poprawności klasyfikacyjnej. Aby znaleźć najlepsze parametry przy nauce za pomocą zestawu danych „Teaching Assistant Evaluation", wyznaczono wartości dla każdego parametru, a następnie sprawdzano wynikową poprawność każdej kombinacji.

W eksperymentach przetestowano szeroki zakres wartości: współczynnik uczenia a oraz współczynnik jego degradacji b przyjmowały wartości ze zbioru {0.1, 0.2, …, 0.9}, parametr okna e ze zbioru {0.1, 0.2, …, 0.7}, maksymalna liczba epok max_ep przyjmowała wartości {10, 50, 100, 200, 400, 600}, natomiast minimalny próg uczenia min_a sprawdzano dla wartości {0.000015, 0.00001, 0.000005, 0.000001}. Łącznie sprawdzono 13 608 kombinacji tych pięciu parametrów. Czas trwania nauki modelu zależał od liczby foldów w walidacji krzyżowej – dla 5 foldów wynosił około godziny, dla 10 – dwie godziny, a dla 20 – do 4 godzin. Ze względu na niewielki rozmiar zbioru testowego (151 przypadków) możliwe było przetestowanie tak dużej liczby kombinacji.

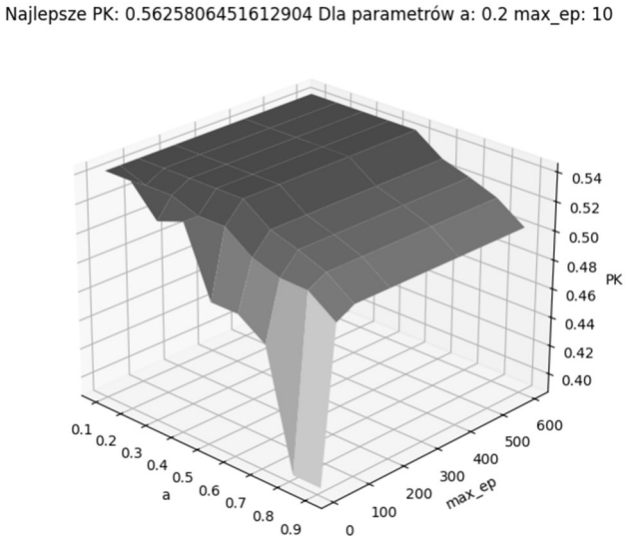


Rys. 6. Wykres przestawiający zależność parametrów a i b dla 5 foldów

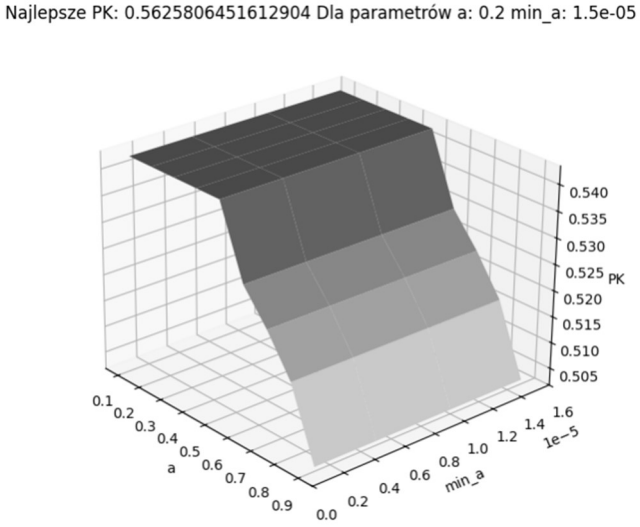
Fig. 6. Graph showing the dependence of parameters a and b for 5 folds



Rys. 7. Wykres przestawiający zależność parametrów a i e dla 5 foldów
Fig. 7. Graph showing the dependence of parameters a and e for 5 folds

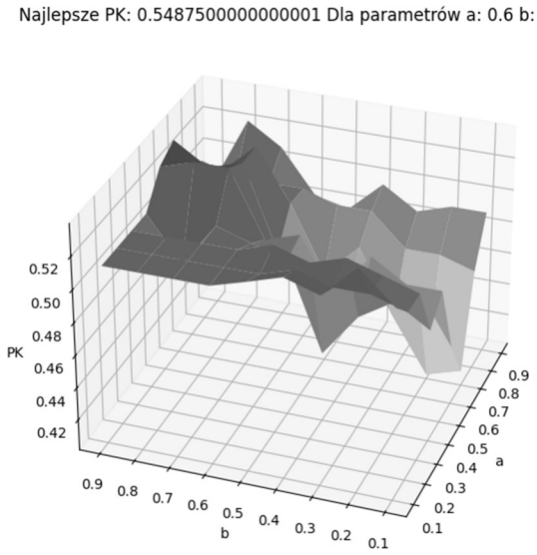


Rys. 8. Wykres przestawiający zależność parametrów a i max_ep dla 5 foldów
Fig. 8. Graph showing the dependence of parameters a and max_ep for 5 folds

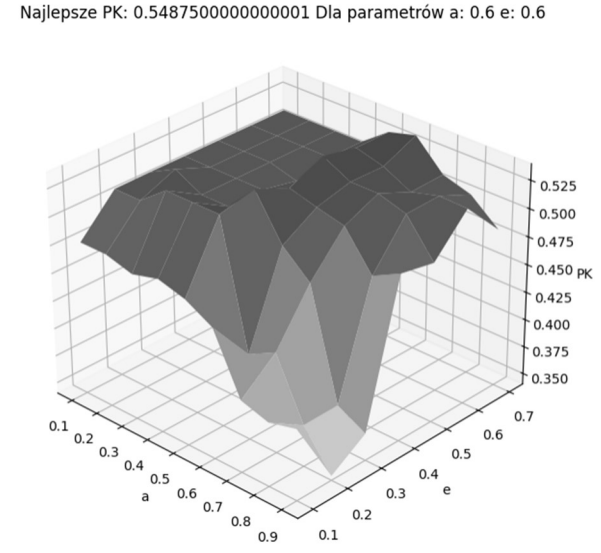


Rys. 9. Wykres przestawiający zależność parametrów a i min_a dla 5 foldów
Fig. 9. Graph showing the dependence of parameters a and min_a for 5 folds

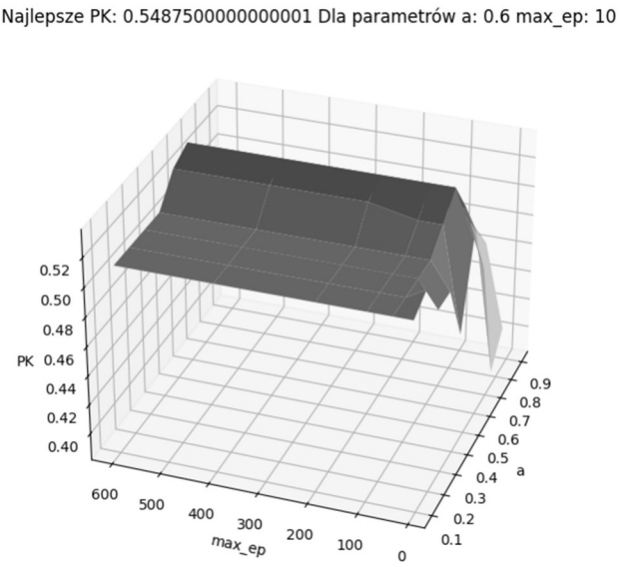
Na rysunkach 6.-9. przedstawiono wpływ doboru parametrów na skuteczność klasyfikacji przy zastosowaniu walidacji krzyżowej z 5 foldami. Najlepszy wynik (PK≈56.26%) uzyskano dla zestawu parametrów: a=0.2, b=0.1, e=0.4, max_ep=10 oraz min_a=0.000015.



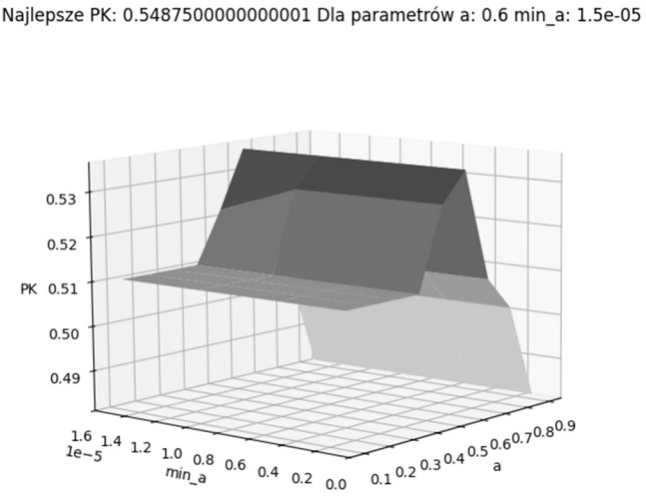
Rys. 10. Wykres przestawiający zależność parametrów a i b dla 10 foldów
Fig. 10. Graph showing the dependence of parameters a and b for 10 folds



Rys. 11. Wykres przestawiający zależność parametrów a i e dla 10 foldów
Fig. 11. Graph showing the dependence of parameters a and e for 10 folds

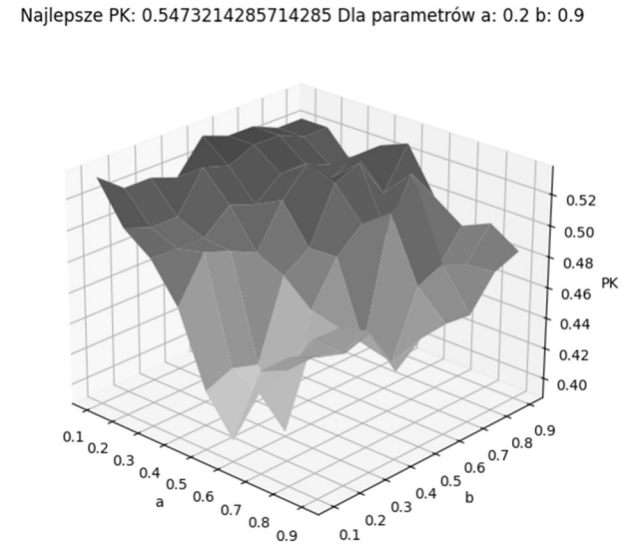


Rys. 12. Wykres przestawiający zależność parametrów a i max_ep dla 10 foldów
Fig. 12. Graph showing the dependence of parameters a and max_ep for 10 folds

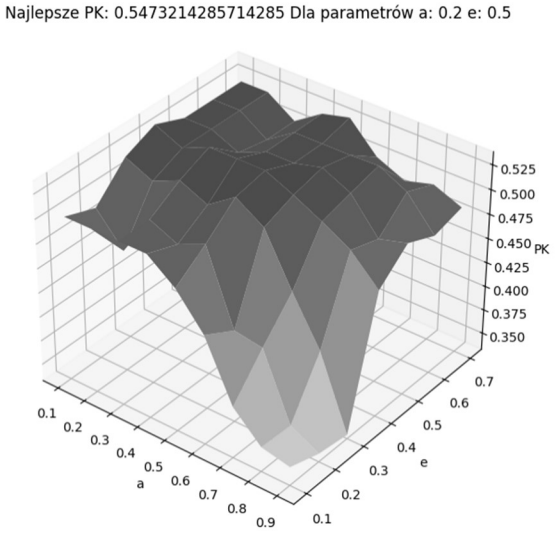


Rys. 13. Wykres przestawiający zależność parametrów a i min_a dla 10 foldów
Fig. 13. Graph showing the dependence of parameters a and min_a for 10 folds

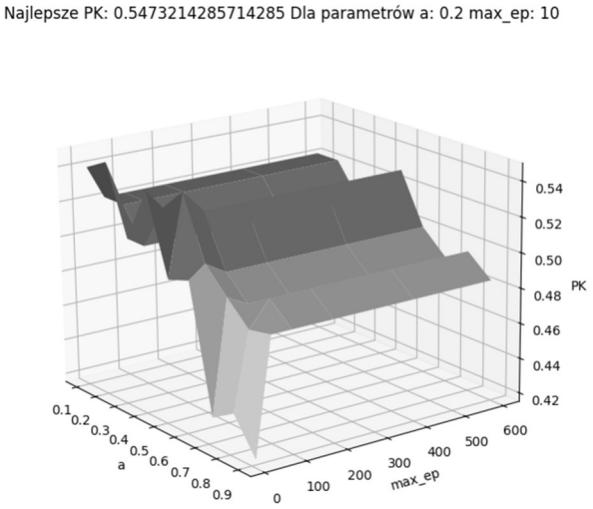
Rysunki 10.-13. ukazują, jak zmiana parametrów wpływa na poprawność klasyfikacji w przypadku 10-foldowej walidacji krzyżowej w tym przypadku najwyższą skuteczność (PK≈54.88%) osiągnięto dla parametrów: a=0.6, b=0.8, e=0.6, max_ep=10 oraz min_a=0.000015.



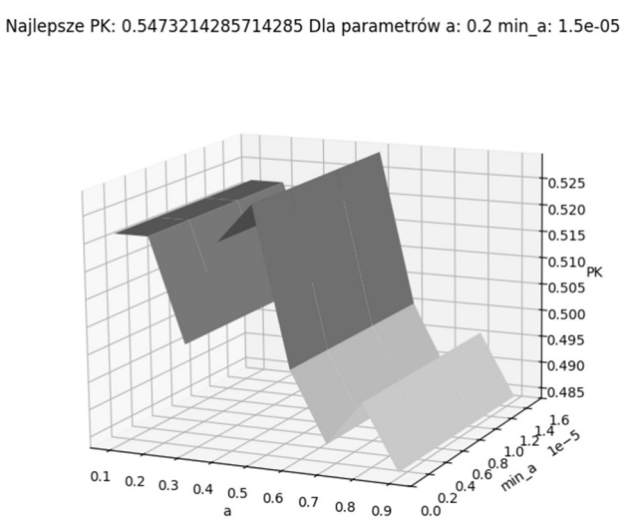
Rys. 14. Wykres przestawiający zależność parametrów a i b dla 20 foldów
Fig. 14. Graph showing the dependence of parameters a and b for 20 folds



Rys. 15. Wykres przestawiający zależność parametrów a i e dla 20 foldów
Fig. 15. Graph showing the dependence of parameters a and e for 20 folds



Rys. 16. Wykres przestawiający zależność parametrów a i max_ep dla 20 foldów
Fig. 16. Graph showing the dependence of parameters a and max_ep for 20 folds



Rys. 17. Wykres przestawiający zależność parametrów a i min_a dla 20 foldów

Fig. 17. Graph showing the dependence of parameters a and min_a for 20 folds

Analogicznie, na rys. 14.-17. zilustrowano wyniki dla walidacji 20-foldowej, co potwierdza powtarzalność zależności pomiędzy parametrami a skutecznością. Najlepszy wynik (PK≈54.73%) uzyskano przy ustawieniach: a=0.2, b=0.9, e=0.5, max_ep=10 oraz min_a=0.000015.

Niezależnie od liczby foldów w walidacji krzyżowej, uzyskana poprawność modelu utrzymuje się na podobnym poziomie (54-56%). Oznacza to, że algorytm LVQ jest stabilny dla tego zbioru danych. Model dobrze zaklasyfikował tylko połowę danych wejściowych, co nie jest najlepszym wynikiem. Natomiast wskazuje to na to, że sieć nie zgaduje cały czas jednej i tej samej wartości, ponieważ wtedy poprawność wynosiłaby w okolicach 33%. Maksymalna liczba epok oraz najmniejsza możliwa wartość współczynnika nauczania (max_ep i min_a) w każdym przypadku mają tą samą wartość. Sugeruje to, że model działa najlepiej, gdy nie ma dużej ilości iteracji. Dłuższy trening przekłada się na taki sam lub gorszy stopień poprawności.

Po wykresach widać, że współczynnik uczenia, współczynnik zmniejszania współczynnika uczenia oraz parametr do obliczania szerokości okna (a, b i e) są ze sobą związane. Oprócz tego parametry te różnią się w każdym foldzie, co wskazuje, że ich wartości są bardziej wrażliwe na konkretny podział danych. Patrząc na wykresy można zauważyć, że dla dużych wartości a z połączeniem z małymi wartościami b, poprawność jest najmniejsza. Wraz ze wzrostem a i maleniem b otrzymuje się lepsze wyniki. Zależność między a i e wygląda podobnie, natomiast w tym przypadku lepsze wyniki otrzymywane są przy zwiększaniu zarówno a i e.

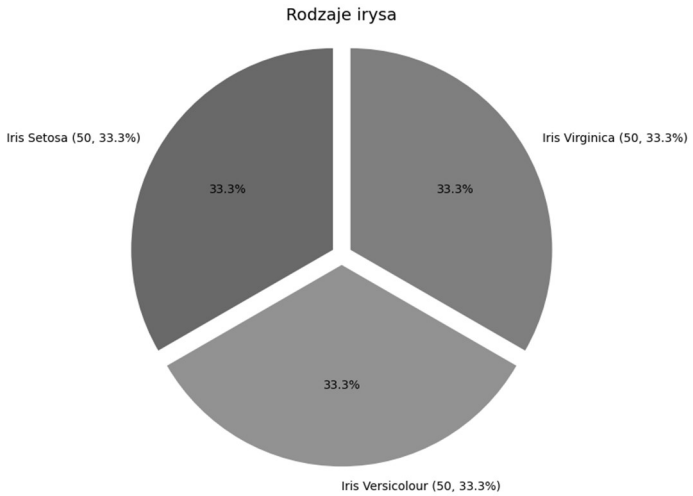
Analizując powyższe 13 608 kombinacji hiperparametrów na zbiorze uczącym składającym się zaledwie z 151 przypadków, należy mieć na uwadze ogromne

ryzyko przeuczenia modelu (overfittingu) pod wybrany zbiór walidacyjny. W przeprowadzonych badaniach brakuje wyodrębnionego, całkowicie niezależnego zbioru testowego, który nie brałby udziału w strojeniu parametrów. Z tego względu uzyskane wyniki i poszukiwania optymalnych parametrów mają charakter głównie dydaktyczno-weryfikacyjny, demonstrując wrażliwość i mechanikę działania klasycznego algorytmu LVQ.

5.2. Sprawdzenie poprawności klasyfikacyjnej modelu na uniwersalnym zestawie danych „Iris"

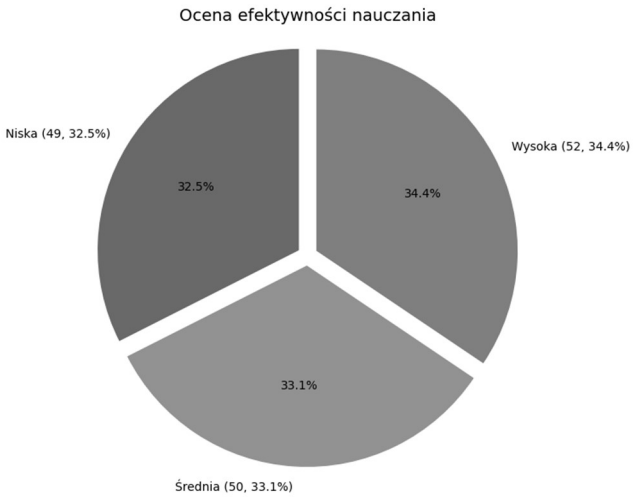
Uzyskana poprawność na poziomie 54% nie jest wynikiem satysfakcjonującym. Oznacza to, że tylko połowa danych wejściowych została poprawnie zaklasyfikowana. Aby zweryfikować, czy niska skuteczność wynika z błędów w implementacji algorytmu LVQ, czy z charakterystyki danych, przeprowadzono testy na standardowym zbiorze „Iris". Dla tego zbioru oczekiwana skuteczność algorytmu powinna oscylować w granicach 95%.

Rysunki 18. i 19. porównują rozkład klas w zbiorach danych „Iris" i „Teaching Assistant Evaluation", co ułatwia interpretację wyników testów modelu. Jak widać po wykresach, rozkłady klas w obu zbiorach danych są bardzo podobne. Jest podział na trzy klasy, które w przypadku „Iris" są równo podzielone. W drugim zestawie różnica między największą i najmniejszą klasą wynosi jedynie 3, więc też jest równo podzielone, natomiast z lekkimi różnicami. Ilość cech też jest w pewnym stopniu podobna.



Rys. 18. Rozkład klas w zbiorze danych „Iris"

Fig. 18. Class distribution in the "Iris" dataset

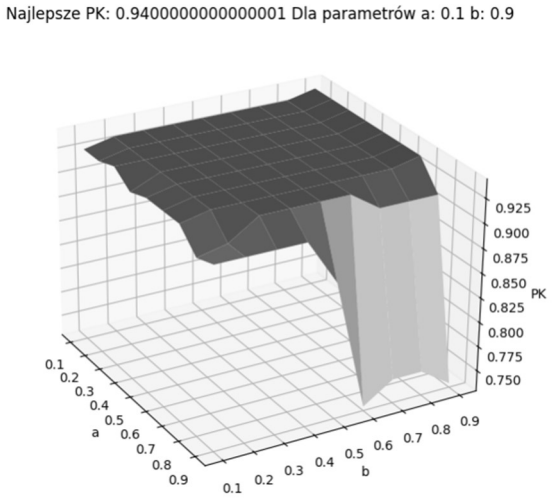


Rys. 19. Rozkład klas w zbiorze danych „Teaching Assistant Evaluation”
Fig. 19. Class distribution in the “Teaching Assistant Evaluation” dataset

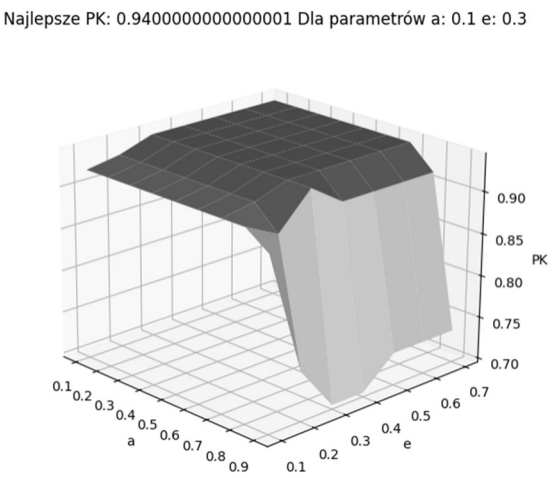
Zestaw danych „Teaching Assistant Evaluation” posiada 5 cech, na podstawie których dobiera się klasę. Zestaw danych „Iris” ma tylko jedną cechę mniej. Jego cechy składają się z: długości działki kielicha, szerokości działki kielicha, długości płatka kielicha oraz szerokości płatka kielicha. Zarówno rozkład klas, jak i ilość cech są bardzo podobne w obu zestawach danych, co sugeruje, że poprawność klasyfikacyjna będzie podobna.

Wpływ parametrów na poprawność klasyfikacji danych „Iris” przedstawiono na rys. 20.-23. Sprawdzanie poprawności zostało wykonane przy CVN=10. Zaskakująco dla parametrów: a=0.1, b=0.9, e=0.3, max_ep=50 oraz min_a=0.000015, poprawność klasyfikacyjna modelu wyniosła aż 94%, co jest wynikiem o wiele lepszym niż 54% przy zbiorze „Teaching Assistant Evaluation”. W tym przypadku prawie każde dane wejściowe są prawidłowo zaklasyfikowane.

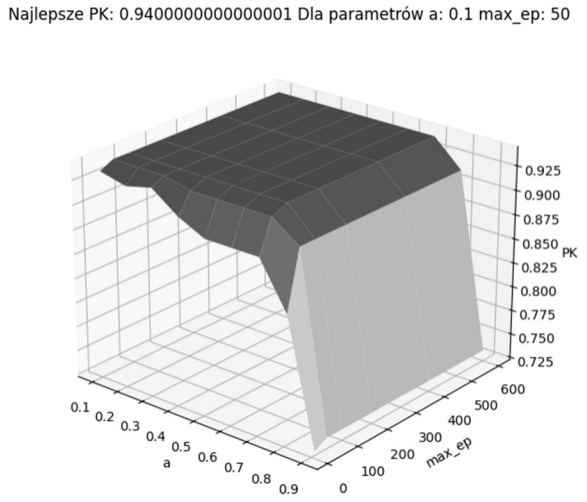
Mimo pozornego podobieństwa między zbiorem „Iris” i „Teaching Assistant Evaluation”, istnieje wiele możliwych czynników wpływających na różnicę w poprawności. Cechy w zbiorze „Iris” są precyzyjnymi pomiarami biologicznymi, które bezpośrednio określają morfologię kwiatów. Istnieje silna korelacja między tymi cechami a gatunkiem rośliny, co jest uwarunkowane genetycznie. w przypadku zbioru „Teaching Assistant Evaluation”, cechy takie jak ocena efektywności mogą być interpretowane subiektywnie przez różnych oceniających, wprowadzając szum do danych. Ponadto, niektóre istotne czynniki wpływające na ocenę asystenta mogą nie być uwzględnione w dostępnych cechach.



Rys. 20. Wykres przestawiający zależność parametrów a i b dla 10 foldów w zbiorze danych „Iris”
Fig. 20. Graph showing the dependence of parameters a and b for 10 folds using “Iris” dataset

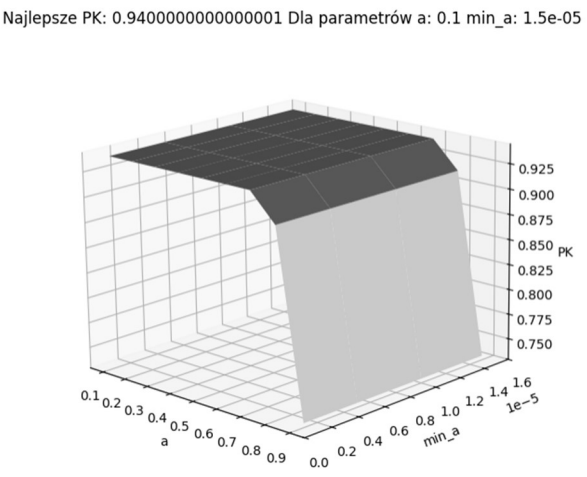


Rys. 21. Wykres przestawiający zależność parametrów a i e dla 10 foldów w zbiorze danych „Iris”
Fig. 21. Graph showing the dependence of parameters a and e for 10 folds using “Iris” dataset



Rys. 22. Wykres przestawiający zależność parametrów a i max_ep dla 10 foldów w zbiorze danych „Iris”

Fig. 22. Graph showing the dependence of parameters a and max_ep for 10 folds using “Iris” dataset



Rys. 23. Wykres przestawiający zależność parametrów a i min_a dla 10 foldów w zbiorze danych „Iris”

Fig. 23. Graph showing the dependence of parameters a and min_a for 10 folds using “Iris” dataset

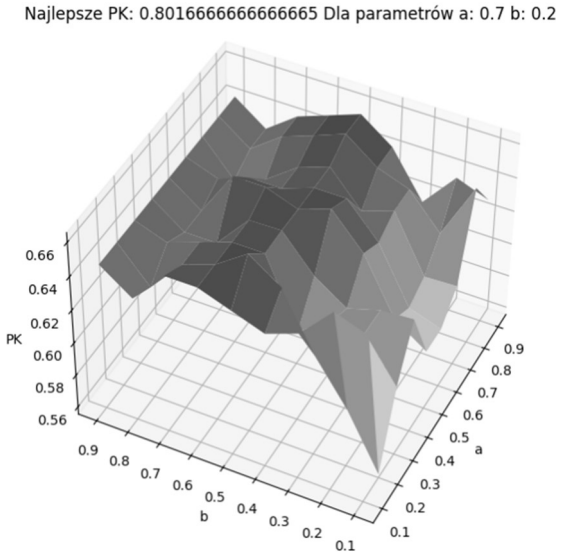
Algorytm LVQ, jako metoda oparta na prototypach, jest szczególnie skuteczny, gdy klasy tworzą wyraźne skupiska w przestrzeni cech. Oznacza to, że w przypadku zbioru „Iris”, LVQ może łatwo zlokalizować reprezentatywne prototypy dla każdej klasy. Dla zbioru „Teaching Assistant Evaluation” znalezienie

odpowiednich prototypów jest znacznie trudniejsze, ponieważ granice między klasami są rozmyte.

5.3. Wpływ zmiany cechy klasyfikującej na skuteczność

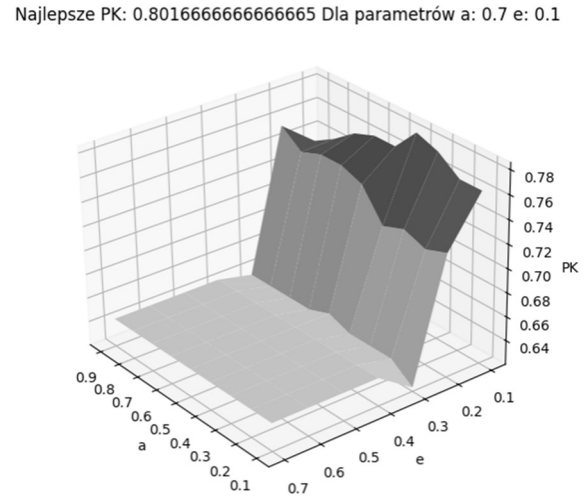
Można podjąć inne próby, aby zwiększyć poprawność modelu. Jeżeli wybranie ostatniej cechy jako klasyfikatora daje jedynie około 54% poprawności, to można spróbować ustawić inną cechę jako klasyfikator. w tym celu, zamiast jako klasy przyjąć ocenę efektywności nauczania, przyjmie się cechę pierwszą: czy nauczyciel jest rodzimym użytkownikiem języka angielskiego. Ten klasyfikator nie może być subiektywny, więc może dać lepszą poprawność klasyfikacyjną. Testujemy model przy podzieleniu danych na 10 foldów.

Rysunki 24.-27. prezentują wyniki eksperymentów po zmianie cechy klasyfikującej. Wartość poprawności klasyfikacyjnej modelu wynosi około 80.17%, dla parametrów: a=0.7, b=0.2, e=0.1, max_ep=10 oraz min_a=0.000015. Należy jednak zauważyć bardzo istotny problem analityczny. Dla tak niezbalansowanych danych (80,8% do 19,2%), skuteczność wyrażona klasyczną metryką dokładności (Accuracy) rzędu 80% oznacza w praktyce predykcję trywialną, ponieważ sieć mogła zoptymalizować się do ciągłego zgadywania klasy dominującej. Wskazuje to jednoznacznie, że sama dokładność nie jest odpowiednią metryką dla takich zbiorów. W przyszłych badaniach rozwinięciem tego problemu powinno być zastosowanie metryk takich jak Precision, Recall, F1-score oraz analiza macierzy pomyłek.



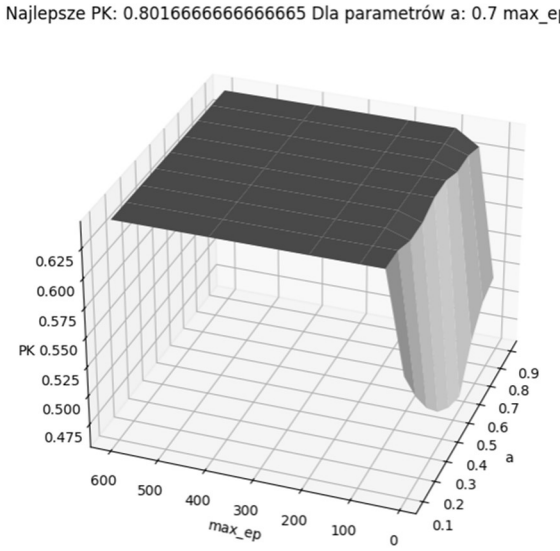
Rys. 24. Wykres przestawiający zależność parametrów a i b dla 10 foldów po zmianie cechy klasyfikującej

Fig. 24. Graph showing the dependence of parameters a and b for 10 folds after changing the classifying attribute



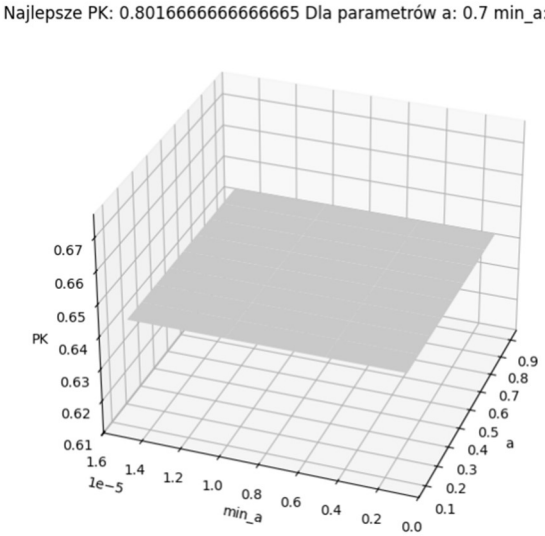
Rys. 25. Wykres przestawiający zależność parametrów a i e dla 10 foldów po zmianie cechy klasyfikującej

Fig. 25. Graph showing the dependence of parameters a and e for 10 folds after changing the classifying attribute



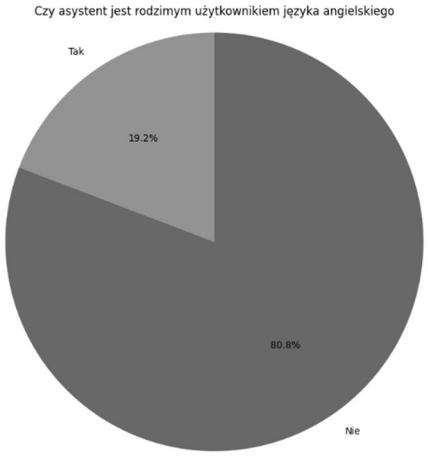
Rys. 26. Wykres przestawiający zależność parametrów a i max_ep dla 10 foldów po zmianie cechy klasyfikującej

Fig. 26. Graph showing the dependence of parameters a and max_ep for 10 folds after changing the classifying attribute



Rys. 27. Wykres przestawiający zależność parametrów a i min_a dla 10 foldów po zmianie cechy klasyfikującej

Fig. 27. Graph showing the dependence of parameters a and min_a for 10 folds after changing the classifying attribute



Rys. 28. Rozkład klas po zmianie klasyfikatora

Fig. 28. Class distribution after changing the classifier

Jak pokazuje rys. 28., rozkład klas w nowym wariancie jest silnie niezbalansowany, gdzie jedna wartość zajmuje nawet 80.8% danych, co wpływa na interpretację poprawności. Są to bardzo niefortunne wiadomości dla powodzenia tego

eksperymentu, ponieważ z dużą poprawnością liczącą 80% model LVQ mógł zawsze zgadywać, że odpowiedzią będzie klasa z większą ilością przypadków. Wybór takiego klasyfikatora, bez poprzedniej analizy danych, na których się pracuje, może wprowadzać w błąd, że model działa zgodnie z oczekiwaniami.

6. Podsumowanie i wnioski

Learning Vector Quantization (LVQ) stanowi klasę algorytmów uczenia nadzorowanego opartych na koncepcji sieci neuronowych. Główna idea LVQ polega na tworzeniu prototypów reprezentujących poszczególne klasy w przestrzeni cech, które następnie są iteracyjnie aktualizowane, aby wyznaczyć odpowiednie granice decyzyjne między klasami. Zaimplementowana w projekcie ewolucja algorytmu (warianty LVQ 1, LVQ 2, LVQ 2.1 oraz LVQ 3) obrazuje zróżnicowane podejścia do modyfikacji reguł aktualizacji tych prototypów.

Wykorzystanie klasycznego algorytmu LVQ w odniesieniu do standardowych zbiorów danych z repozytorium UCI (TAE, Iris) nadaje przeprowadzonym badaniom charakter aplikacyjny oraz dydaktyczno-weryfikacyjny. Celem projektu było przede wszystkim praktyczne studium oceniające zachowanie modelu w obliczu zróżnicowanych danych.

Przeprowadzone eksperymenty w pełni potwierdziły powszechnie znane i ugruntowane w literaturze zależności. Algorytm charakteryzuje się bardzo wysoką skutecznością (na poziomie 94%) w przypadku danych o wyraźnej separacji klastrów i obiektywnych miarach, co zademonstrowano na przykładzie zbioru „Iris". Z drugiej strony, modele oparte na wektorach prototypowych napotykają znaczne trudności w przypadku danych o zatartej granicy decyzyjnej, takich jak subiektywne ewaluacje studenckie w zbiorze „Teaching Assistant Evaluation" (skuteczność rzędu 54-56%). Projekt stanowi weryfikację tezy, że skuteczność działania metod klasyfikacyjnych zależy nie tylko od precyzyjnego dostrojenia hiperparametrów, ale przede wszystkim od charakterystyki, zbalansowania i obiektywnej jakości samych danych wejściowych.

Literatura

[1] Archive ICS UCI, Teaching Assistant Evaluation Data Set, https://archive.ics.uci.edu/dataset/100/teaching+assistant+evaluation

[2] Archive ICS UCI, Iris Data Set, https://archive.ics.uci.edu/dataset/53/iris

[3] DergiPark, Artykuł naukowy, https://dergipark.org.tr/tr/download/article-file/2844839

[4] Kohonen T., "Learning Vector Quantization", Neural Networks, Vol. 1, Supplement 1, 1988.

[5] Osowski S., "Sieci neuronowe do przetwarzania informacji", Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2006.

[6] Tadeusiewicz R., "Sieci neuronowe", Akademicka Oficyna Wydawnicza, Warszawa 1993.

USE OF LEARNING VECTOR QUANTIZATION FOR ANALYSIS AND CLASSIFICATION OF TEACHING ASSISTANT EVALUATION DATA

Summary

The aim is to present the process of designing and implementing a neural network using the Learning Vector Quantization (LVQ) algorithm in a data classification task. The subject of the research was the "Teaching Assistant Evaluation" dataset, containing data regarding the evaluation of teaching assistant effectiveness based on attributes such as native English speaker status, instructor, course, class size, or semester type. The theoretical part of the work introduces the structure of the McCulloch-Pitts neuron and characterizes in detail the operating mechanisms of the algorithm variants: LVQ 1, LVQ 2, LVQ 2.1, and LVQ 3, emphasizing differences in prototype vector update methods. The research part covers the implementation of the model in Python and conducting a series of validation experiments using the cross-validation technique. a key element of the research was a multi-faceted analysis of the impact of hyperparameter selection – including the learning rate, its decay rate, and the number of epochs – on the network's generalization abilities. It was shown that for the analyzed evaluation dataset, the maximum accuracy obtained was about 56%, which suggests difficulties related to the subjective nature of the input data. To verify the correctness of the implementation, comparative tests were conducted on the standard "Iris" dataset, achieving a result of 94%, which confirmed the algorithm's effectiveness for data with clear cluster separation. An attempt was also made to change the decision attribute, which, however, resulted in the problem of strongly imbalanced classes. Final conclusions highlight that the LVQ method is highly sensitive to data quality and structure and requires precise parameter tuning, and its effectiveness decreases in the case of data with blurred decision boundaries.

Keywords: neural networks, lvq, classification, supervised learning, data analysis.

Patryk Błoński, Maja Jaszowska, Sandra Kudła
Koło Naukowe Automatyków i Robotyków

dr inż. Tomasz Żabiński
opiekun Koła Naukowego

ANALIZA DZIAŁANIA I IMPLEMENTACJA PRAKTYCZNA TRANSFORMATORA TESLI

Streszczenie

Niniejszy rozdział poświęcony jest wieloaspektowej analizie teoretycznej oraz praktycznej realizacji modelu transformatora rezonansowego, znanego powszechnie jako cewka Tesli, z wykorzystaniem nowoczesnych elementów półprzewodnikowych. Podstawowym celem podjętych badań było zaprojektowanie, konstrukcja oraz weryfikacja parametrów pracy układu typu Slayer Exciter, który stanowi bezpieczną i wydajną alternatywę dla klasycznych konstrukcji iskiernikowych. W procesie badawczym wykorzystano metodę krytycznej analizy literatury przedmiotu, obejmującą zarówno klasyczne prace Nikoli Tesli, jak i współczesne opracowania z zakresu inżynierii wysokich częstotliwości, a także metodę eksperymentalną opartą na testach prototypu. Obiekt badań stanowił układ elektroniczny bazujący na tranzystorze średniej mocy, który współpracował z precyzyjnie wykonanym uzwojeniem o liczbie 270 zwojów. W toku przeprowadzonych doświadczeń wykazano, że kluczowym czynnikiem decydującym o sprawności urządzenia jest osiągnięcie stanu pełnego rezonansu elektromagnetycznego, co wymaga idealnego dopasowania częstotliwości drgań obwodu pierwotnego i wtórnego. Istotnym elementem projektu było zastosowanie elektrody górnej w postaci torusa (terminala górnego), co pozwoliło na optymalizację rozkładu pola elektrycznego i zwiększenie pojemności szczytowej układu. Wyniki eksperymentów potwierdziły wysoką skuteczność bezprzewodowego przesyłu energii elektrycznej, demonstrując możliwość wzbudzania gazów w lampach wyładowczych bez bezpośredniego połączenia galwanicznego. Przeprowadzona analiza dowodzi, że historyczne koncepcje Tesli znajdują dziś realne zastosowanie w systemach bezprzewodowego ładowania pojazdów elektrycznych (WPT) oraz w nowoczesnej diagnostyce medycznej. Rozdział konkluduje, że zrealizowany model fizyczny stanowi doskonałe narzędzie dydaktyczne, pozwalające na efektywną wizualizację złożonych zjawisk fizycznych zachodzących w obwodach wysokiego napięcia.

Słowa kluczowe: cewka rezonansowa, indukcja magnetyczna, technologia półprzewodnikowa, systemy WPT, jonizacja powietrza, tranzystor bipolarny.

1. Wprowadzenie

Transformator Tesli, znany również jako cewka Tesli, stanowi jedno z najbardziej fascynujących osiągnięć inżynierii elektrycznej przełomu XIX i XX wieku. Urządzenie to, opatentowane przez Nikolę Teslę w 1891 roku[8], opiera się na wykorzystaniu zjawiska rezonansu w obwodach prądu przemiennego o wysokiej częstotliwości. W przeciwieństwie do tradycyjnych transformatorów, które bazują na sprzężeniu magnetycznym realizowanym poprzez rdzeń ferromagnetyczny, konstrukcja Tesli wykorzystuje cewki powietrzne, co pozwala na generowanie ekstremalnie wysokich napięć, sięgających milionów woltów.

Głównym celem niniejszego rozdziału jest przedstawienie zasady działania dwuobwodowego transformatora rezonansowego oraz szczegółowy opis procesu budowy jego fizycznego modelu. W pracy podjęto próbę analizy parametrów wpływających na efektywność przesyłu energii elektrycznej pomiędzy obwodem pierwotnym a wtórnym. Obiektem badań jest transformator składający się z kondensatora wysokonapięciowego, iskiernika oraz precyzyjnie nawiniętych cewek. Kluczowym wyzwaniem inżynierskim w tym procesie jest doprowadzenie obu obwodów do stanu rezonansu, co wymaga precyzyjnego zsynchronizowania ich częstotliwości drgań własnych. Częstotliwość ta jest bezpośrednio uzależniona od wzajemnej relacji parametrów fizycznych komponentów, w szczególności Indukcyjności cewek wykonanych z zastosowanych uzwojeń oraz pojemności zastosowanych kondensatorów.

Metodyka badawcza obejmuje zarówno analizę literatury technicznej, w tym książek i recenzowanych artykułów naukowych, jak i badania własne przeprowadzone podczas testów gotowego urządzenia. W dobie rosnącego zainteresowania bezprzewodowym przesyłem energii oraz nowoczesną dydaktyką techniczną, analiza historycznych projektów Tesli pozwala na lepsze zrozumienie fundamentów współczesnej radiotechniki i systemów komunikacji bezprzewodowej . Zgodnie z wymaganiami redakcyjnymi, rozdział skupia się na rzetelnym przedstawieniu podbudowy teoretycznej przed opisem części praktycznej.

2. Wstęp teoretyczny

Fundamentem działania transformatora Tesli jest zjawisko rezonansu elektromagnetycznego, które zachodzi w układzie dwóch sprzężonych ze sobą obwodów drgających. W klasycznej elektrotechnice transformatory wykorzystują rdzenie ferromagnetyczne do maksymalizacji strumienia magnetycznego przy niskich częstotliwościach. Jednak w przypadku urządzenia Tesli, pracującego przy bardzo wysokich częstotliwościach, zastosowanie żelaznego rdzenia prowadziłoby do ogromnych strat energii wynikających z histerezy oraz prądów wirowych [3], [7]. Z tego powodu konstrukcja opiera się na cewkach powietrznych, gdzie sprzężenie między uzwojeniami zachodzi głównie przez zmienne pole magnetyczne, a powietrze pełni rolę ośrodka izolacyjnego i niemagnetycznego.

Urządzenie to zostało zaprojektowane jako transformator rezonansowy, w którym obwód pierwotny, zasilany wysokim napięciem, pobudza obwód wtórny do drgań. Warunkiem koniecznym do uzyskania maksymalnej sprawności urządzenia jest zsynchronizowanie obu obwodów tak, aby pracowały one na tej samej częstotliwości rezonansowej. W literaturze przedmiotu wskazuje się, że transformator ten miał służyć nie tylko do widowiskowych pokazów wyładowań, ale przede wszystkim jako kluczowy element systemów bezprzewodowego przesyłu energii elektrycznej oraz sygnałów radiowych na duże odległości.

Pierwotnym zamierzeniem twórcy było stworzenie globalnego systemu komunikacji i dystrybucji energii, który wykorzystywałby przewodnictwo górnych warstw atmosfery lub samej ziemi [1], [4]. Choć wizja ta nie została w pełni zrealizowana w skali przemysłowej, to zasady fizyczne opracowane przy budowie cewki były ważnym wkładem w rozwój nowoczesnej radiodyfuzji oraz technologii medycznych, takich jak diatermia. W ujęciu teoretycznym, transformator Tesli jest przykładem układu o bardzo wysokiej dobroci, co pozwala na gromadzenie energii w polu magnetycznym i elektrycznym, a następnie jej gwałtowne uwalnianie w postaci wyładowań koronowych i iskrowych. Zastosowanie m.in. dużej liczby zwojów w jednej warstwie, odpowiednich odstępów, izolacji, właściwego kształtu uzwojenia oraz gładkiego terminala górnego, który ogranicza lokalne natężenie pola elektrycznego, pozwala na uniknięcie przebić międzyzwojowych, co jest krytyczne przy operowaniu na potencjałach rzędu setek tysięcy woltów [11].

3. Historia Nikoli Tesli

Nikola Tesla urodził się w 1856 roku w miejscowości Smiljan [5] (na terenie dzisiejszej Chorwacji) w rodzinie serbskiego kapłana prawosławnego. Od najmłodszych lat wykazywał ponadprzeciętne zdolności matematyczne oraz niezwykłą wyobraźnię przestrzenną, która pozwalała mu konstruować i testować urządzenia w pamięci, bez konieczności sporządzania wstępnych szkiców. Po studiach na Politechnice w Grazu oraz Uniwersytecie w Pradze, Tesla rozpoczął pracę w Budapeszcie, a następnie w Paryżu, gdzie pracował dla Continental Edison Company. To właśnie tam zyskał pierwsze doświadczenie w dziedzinie praktycznej inżynierii elektrycznej, co ostatecznie doprowadziło go do decyzji o emigracji do Stanów Zjednoczonych w 1884 roku.

Przybycie Tesli do Nowego Jorku stało się początkiem jednego z najsłynniejszych konfliktów w historii nauki, znanego jako "wojna o prąd". Tesla, będący zwolennikiem prądu przemiennego (AC), szybko wszedł w konflikt z Thomasem Edisonem, który zainwestował cały swój kapitał i autorytet w promowanie prądu stałego (DC). Drogi obu wynalazców rozeszły się po sporze finansowym, co popchnęło Teslę do współpracy z George’em Westinghouse’em.

Kluczowe innowacje i wydarzenia z życia Tesli, które bezpośrednio poprzedziły wynalezienie transformatora rezonansowego, obejmują:

- opracowanie koncepcji wirującego pola magnetycznego,

- skonstruowanie pierwszego silnika indukcyjnego prądu przemiennego,
- oświetlenie wystawy światowej w Chicago (1893) [5] przy użyciu systemu AC,
- uruchomienie pierwszej elektrowni wodnej na wodospadzie Niagara.

Najbardziej owocny okres dla badań nad wysokimi częstotliwościami przypadł na lata 90. XIX wieku. W 1891 roku Tesla uzyskał obywatelstwo amerykańskie i w tym samym roku opatentował urządzenie, które dziś nazywamy cewką Tesli. Wynalazca przeniósł swoje laboratorium do Colorado Springs w 1899 roku, gdzie prowadził eksperymenty na niespotykaną dotąd skalę. Twierdził tam, że udało mu się przesłać energię elektryczną bez użycia przewodów na odległość kilkunastu kilometrów, zapalając żarówki wbite bezpośrednio w ziemię [8]. Choć zapiski z tych badań do dziś budzą kontrowersje, stały się one fundamentem dla jego najambitniejszego projektu – wieży Wardenclyffe.

Wieża na Long Island miała być sercem "Światowego Systemu" bezprzewodowej komunikacji i energii. Tesla marzył o darmowej energii dostępnej dla każdego mieszkańca planety, jednak brak stałego finansowania ze strony bankiera J.P. Morgana oraz sukcesy Marconiego w dziedzinie prostszej komunikacji radiowej doprowadziły projekt do upadku. Tesla zmarł w zapomnieniu w 1943 roku w hotelu New Yorker, zostawiając po sobie setki patentów, które dopiero po dekadach zostały w pełni docenione przez środowisko naukowe. Współcześnie postać Tesli jest symbolem genialnego wizjonera, którego prace nad transformatorem rezonansowym wyprzedziły swoją epokę o całe stulecie.

4. Jak działa transformator Tesli

Analiza teoretyczna działania transformatora Tesli wymaga odwołania się do zasad dynamiki obwodów rezonansowych o słabym sprzężeniu magnetycznym. W odróżnieniu od konwencjonalnych transformatorów, które pracują przy silnym sprzężeniu magnetycznym, urządzenie Tesli wykorzystuje sprzężenie luźne i bazuje na procesie akumulacji energii w czasie. Układ ten składa się z dwóch odrębnych obwodów drgających LC (indukcyjno-pojemnościowych), które muszą zostać precyzyjnie nastrojone do tej samej częstotliwości drgań własnych. W publikacji *The Ultimate Tesla Coil Design and Construction Guide* autor wskazuje, że kluczem do sprawności urządzenia jest wysoka dobroć (Q) obwodu wtórnego, która pozwala na minimalizację strat rezystancyjnych przy jednoczesnym maksymalizowaniu amplitudy napięcia wyjściowego.

Cykl pracy urządzenia rozpoczyna się od ładowania kondensatora pierwotnego przez transformator zasilający wysokiego napięcia. W momencie osiągnięcia potencjału przebicia na iskierniku, dochodzi do zamknięcia obwodu pierwotnego. W tym krótkim interwale czasowym kondensator rozładowuje się przez cewkę pierwotną, generując oscylacje o wysokiej częstotliwości. Zmienne pole magnetyczne wytworzone przez cewkę pierwotną indukuje w cewce wtórnej siłę elektromotoryczną. Zjawisko to jest potęgowane przez rezonans – Zjawisko to jest

potęgowane przez rezonans – w wyniku sprzężenia rezonansowego energia jest stopniowo przekazywana z obwodu pierwotnego do wtórnego w kolejnych półokresach oscylacji, co prowadzi do narastania amplitudy napięcia na szczycie urządzenia. Oryginalne zapiski Tesli podkreślają, że dla uzyskania optymalnych efektów, stosunek indukcyjności do pojemności w obu obwodach musi być starannie wyważony [8].

Istotnym zjawiskiem fizycznym determinującym konstrukcję urządzenia jest efekt naskórkowości. W wyniku tego zjawiska, występującego przy wysokich częstotliwościach rzędu setek kiloherców, gęstość prądu w przewodniku koncentruje się głównie w cienkiej warstwie przypowierzchniowej [3].Wymusza to stosowanie w obwodzie pierwotnym grubych rurek miedzianych lub taśm, które oferują dużą powierzchnię czynną, minimalizując tym samym nagrzewanie się elementów. Ponadto, istotną rolę odgrywa terminal górny (torus). Działa on jak kondensator obciążający obwód wtórny, co nie tylko obniża jego częstotliwość rezonansową, ale także dzięki sferycznemu kształtowi pozwala na nagromadzenie ogromnego ładunku elektrycznego przed momentem inicjacji wyładowania plazmowego w atmosferze.

5. Rzeczywiste zastosowania transformatora Tesli w nowoczesnym przemyśle i technologii

Współczesna inżynieria w dużej mierze opiera się na fundamentach teoretycznych, które Nikola Tesla sformułował ponad sto lat temu. Choć klasyczna cewka Tesli jest dziś postrzegana głównie jako urządzenie demonstracyjne, jej zasada działania – oparta na rezonansowym sprzężeniu magnetycznym i wysokich częstotliwościach – stała się sercem wielu nowoczesnych technologii. Przemysł 4.0, dążenie do pełnej bezprzewodowości oraz zaawansowana diagnostyka medyczna to obszary, w których koncepcje Tesli znajdują realne i komercyjne zastosowanie. Poniżej przedstawiono kluczowe podmioty i technologie, które implementują te rozwiązania w swojej działalności.

5.1. WiTricity Corporation (Bezprzewodowy przesył energii - WPT)

Jest to amerykańska firma technologiczna wywodząca się z Massachusetts Institute of Technology (MIT), która uchodzi za globalnego lidera w dziedzinie bezprzewodowego przesyłu energii na odległość. Technologia WiTricity opiera się bezpośrednio na zjawisku rezonansowego sprzężenia indukcyjnego (Highly Resonant Wireless Power Transfer), które jest tożsame z zasadą działania transformatora Tesli. W przeciwieństwie do standardowego ładowania indukcyjnego (znanego z telefonów), rozwiązanie to pozwala na efektywny przesył energii na większe odległości i z większą tolerancją na niewspółosiowość cewek. Firma skupia się obecnie na sektorze motoryzacyjnym, współpracując z gigantami takimi jak Hyundai czy BMW, aby umożliwić bezprzewodowe ładowanie pojazdów elektrycznych poprzez najechanie nad płytę ładującą umieszczoną w jezdni. Dzięki

zastosowaniu układów rezonansowych o wysokiej dobroci, systemy te osiągają sprawność przekraczającą 90%, co czyni wizję Tesli o bezprzewodowej energetyce realnym rozwiązaniem rynkowym.

5.2. ABB i Siemens Energy (Badania wysokich napięć)

Globalni liderzy w sektorze energetycznym, tacy jak szwajcarskie ABB oraz niemiecki Siemens Energy, wykorzystują zmodyfikowane układy transformatorów rezonansowych do testowania izolacji oraz wytrzymałości dielektrycznej komponentów sieci elektroenergetycznych. Systemy testowe oparte na rezonansie pozwalają na generowanie bardzo wysokich napięć przy stosunkowo niskim poborze mocy z sieci zasilającej. Jest to kluczowe przy badaniu kabli wysokiego napięcia, transformatorów mocy oraz izolatorów stacyjnych przed ich dopuszczeniem do eksploatacji. Wykorzystanie cewek powietrznych w układach rezonansowych eliminuje ryzyko nasycenia rdzenia i pozwala na precyzyjne symulowanie przepięć komutacyjnych oraz atmosferycznych, co bezpośrednio przekłada się na bezpieczeństwo i stabilność systemów energetycznych całych.

5.3. Medtronic i sektor Plasma Medicine

W branży medycznej, firmy takie jak Medtronic wykorzystują technologię prądów wysokiej częstotliwości w urządzeniach elektrochirurgicznych. Choć nie są to "cewki Tesli" w tradycyjnym sensie, ich zasada działania (generowanie wysokich napięć o wysokiej częstotliwości w celu cięcia tkanek i koagulacji) wywodzi się z eksperymentów Tesli z prądami w.cz. Ponadto, dynamicznie rozwija się dziedzina medycyny plazmowej (Plasma Medicine), gdzie urządzenia generujące tzw. "zimną plazmę" (Cold Atmospheric Plasma) wykorzystują transformatory rezonansowe do jonizacji gazu w temperaturze pokojowej. Technologia ta jest stosowana do sterylizacji ran przewlekłych oraz niszczenia komórek nowotworowych bez uszkadzania zdrowych tkanek, co stanowi milowy krok w onkologii i dermatologii.

5.4. Qualcomm (Technologia Halo)

Amerykański gigant technologiczny Qualcomm przez lata rozwijał system Qualcomm Halo, dedykowany bezprzewodowemu zasilaniu pojazdów w ruchu (tzw. Dynamic Electric Vehicle Charging). System ten zakłada umieszczenie cewek rezonansowych bezpośrednio pod nawierzchnią autostrad. Wykorzystując zasady rezonansu elektromagnetycznego, pojazd poruszający się z dużą prędkością mógłby być doładowywany w trakcie jazdy, co rozwiązałoby problem ograniczonego zasięgu aut elektrycznych. Mimo że Qualcomm sprzedał część tych technologii firmie WiTricity, patenty i rozwiązania techniczne wypracowane w tym procesie są fundamentem dla przyszłych "inteligentnych dróg".

Podsumowując tę sekcję, można zauważyć, że transformator Tesli ewoluował z widowiskowego przyrządu fizycznego do rangi kluczowego komponentu nowoczesnej gospodarki. Od bezprzewodowego zasilania maszyn w magazynach

Amazona (poprzez technologię WPT), po zaawansowane testy systemów energetycznych Siemens Energy – dziedzictwo Nikoli Tesli jest obecne w niemal każdym aspekcie współczesnej inżynierii elektrycznej. Integracja tych rozwiązań z technologiami cyfrowymi i sztuczną inteligencją pozwala na jeszcze większą optymalizację przesyłu energii, co wpisuje się w ideę Przemysłu 4.0.

6. Opis projektu

Głównym przedmiotem części praktycznej niniejszej pracy było zaprojektowanie oraz wykonanie niskonapięciowego, półprzewodnikowego transformatora Tesli działającego w topologii samowzbudnej, znanej w literaturze jako *Slayer Exciter* [7]. Wybór tej konkretnej architektury podyktowany był chęcią uzyskania stabilnej pracy urządzenia przy zasilaniu prądem stałym o niskim napięciu, co czyni model bezpiecznym narzędziem dydaktycznym. W przeciwieństwie do klasycznych konstrukcji iskiernikowych, opisany model wykorzystuje szybkie przełączanie tranzystora mocy do indukowania drgań o wysokiej częstotliwości w obwodzie pierwotnym.

Sercem układu elektronicznego jest tranzystor bipolarny NPN typu **D2394** (2SD2394). Jest to komponent mocy charakteryzujący się wysokim prądem kolektora oraz odpowiednią szybkością przełączania, co jest niezbędne do pracy w zakresie częstotliwości radiowych generowanych przez cewkę. Tranzystor ten pełni rolę klucza elektronicznego, który jest sterowany prądem sprzężenia zwrotnego płynącym z dolnego wyprowadzenia cewki wtórnej. Aby zapewnić poprawne warunki pracy bazy tranzystora oraz zabezpieczyć go przed uszkodzeniem, w układzie zastosowano rezystor o wartości **22 kΩ**, który ustala punkt pracy urządzenia, oraz diodę elektroluminescencyjną (LED) w kolorze czerwonym. Dioda ta w tej konkretnej aplikacji pełni podwójną rolę: służy jako wskaźnik pracy urządzenia oraz pełni funkcję elementu ustalającego napięcie progowe bazy, chroniąc tranzystor przed ujemnymi szpilkami napięciowymi powracającymi z obwodu wtórnego.

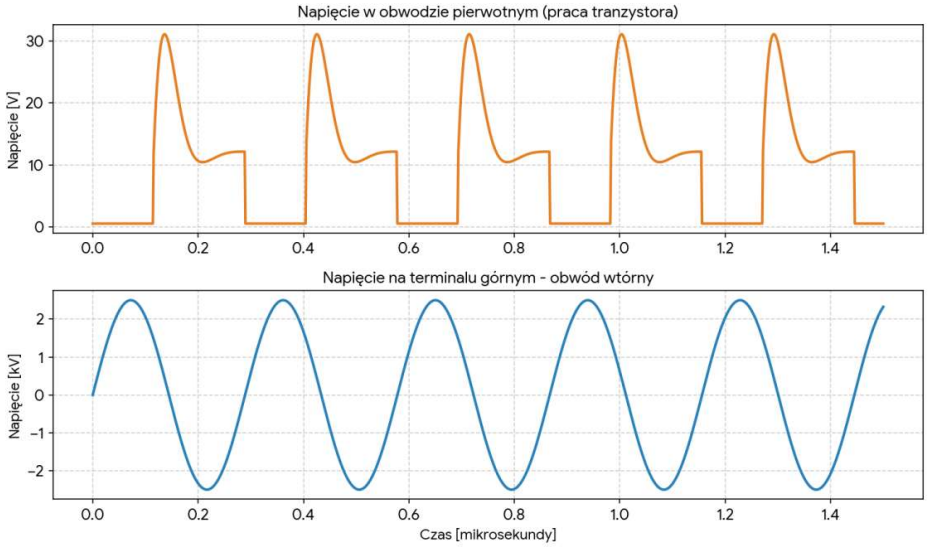
Konstrukcja mechaniczna transformatora obejmuje dwa uzwojenia nawinięte na dielektrycznym karkasie. Cewka wtórna, będąca najważniejszym elementem rezonansowym, składa się z około **270 zwojów** cienkiego drutu miedzianego o średnicy 0,5 mm w emalii. Zastosowanie gęstego nawoju jednowarstwowego pozwoliło na uzyskanie indukcyjności na poziomie około 492 mikrohenrów przy zachowaniu minimalnej pojemności własnej uzwojenia. Dolna część cewki została połączona z bazą tranzystora, tworząc pętlę sprzężenia zwrotnego, która automatycznie dostosowuje częstotliwość pracy klucza do częstotliwości rezonansowej uzwojenia. Cewka pierwotna, wykonana z grubszego przewodu, składa się zaledwie z kilku zwojów nawiniętych u podstawy cewki wtórnej, co zapewnia odpowiednie sprzężenie magnetyczne między obwodami (Źródło: opracowanie własne).



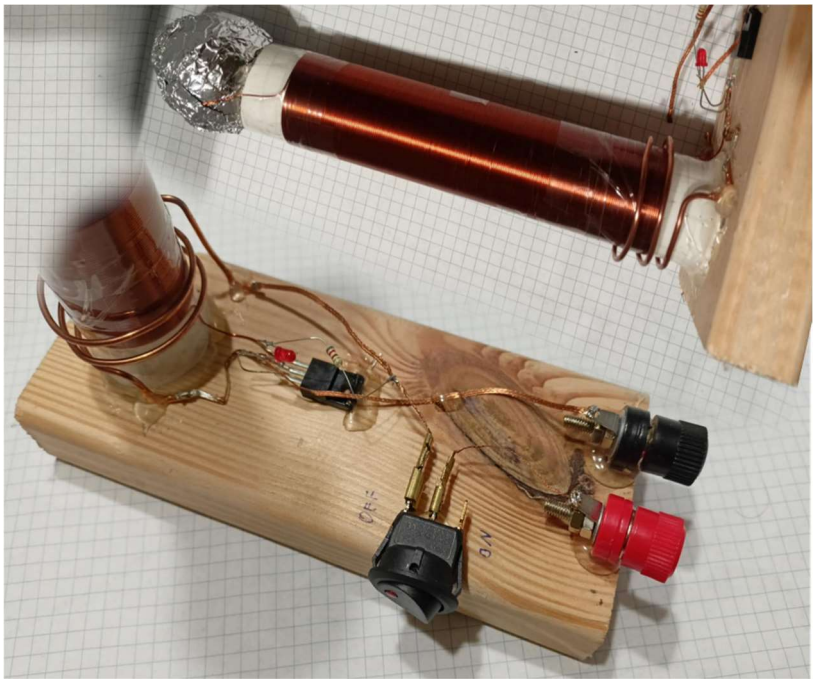
Rys. 1. Widok ogólny zmontowanego modelu transformatora Tesli z zamontowanym torusem
Fig. 1. General view of the assembled Tesla transformer model with the mounted torus

Interfejs użytkownika oraz system zasilania zostały zintegrowane w sposób umożliwiający łatwą diagnostykę i testy. Na panelu sterowania umieszczono mechaniczny przełącznik oraz standardowe gniazda na wtyki bananowe, co pozwala na szybkie podłączenie zewnętrznego zasilacza laboratoryjnego. Dzięki takiemu rozwiązaniu możliwe jest płynne regulowanie napięcia wejściowego i obserwacja wpływu mocy zasilania na intensywność generowanego pola elektromagnetycznego. Model w obecnej fazie pozwala na bezprzewodowe wzbudzanie gazów w lampach neonowych i energooszczędnych znajdujących się w pobliżu urządzenia.

Zwieńczeniem konstrukcji jest **torus (terminal górny)**, który został zamontowany na szczycie uzwojenia wtórnego. Wykonany z materiału przewodzącego o sferycznym kształcie, torus pełni kluczową rolę w zwiększeniu pojemności szczytowej układu. Jego obecność spowodowała pożądane obniżenie częstotliwości rezonansowej, co pozwoliło na gromadzenie większej gęstości ładunku elektrycznego przed momentem wystąpienia wyładowania. Z punktu widzenia fizyki wysokich napięć, torus optymalizuje rozkład pola elektrycznego wokół cewki [8], zapobiegając ulotowi energii bezpośrednio z ostatnich zwojów drutu i skupiając energię w wyznaczonym punkcie emisyjnym. Dzięki temu finalny model wykazuje wysoką sprawność w bezprzewodowym wzbudzaniu lamp gazowych znajdujących się w jego polu oddziaływania.



Rys. 2. przebiegi czasowe napięć
Fig. 2. Voltage waveforms



Rys. 3. Szczegóły układu sterującego Slayer Exciter oraz sposób nawinięcia cewki pierwotnej
Fig. 3. Details of the Slayer Exciter control circuit and the winding method of the primary coil

7. Podsumowanie

Niniejszy rozdział stanowi próbę kompleksowego ujęcia tematyki transformatora rezonansowego – od jego historycznych korzeni i wizjonerskich koncepcji Nikoli Tesli, po współczesne implementacje w dobie Przemysłu 4.0 oraz autorską realizację modelu fizycznego. Przeprowadzona analiza dowiodła, że mimo upływu ponad stu lat od opatentowania pierwszej cewki rezonansowej, jej fundamentalne zasady fizyczne pozostają fundamentem nowoczesnej inżynierii wysokich częstotliwości. Kluczowym wnioskiem płynącym z części teoretycznej jest fakt, że rezonans elektromagnetyczny pozwala na przesył energii z efektywnością niemożliwą do osiągnięcia w klasycznych układach indukcyjnych o stałym rdzeniu, co otwiera drogę do dalszego rozwoju technologii bezprzewodowych.

Praktyczna część projektu, oparta na budowie układu typu *Slayer Exciter*, pozwoliła na empiryczną weryfikację założeń teoretycznych. Wykorzystanie tranzystora bipolarnego **NPN D2394** oraz precyzyjnie nawiniętego uzwojenia wtórnego (270 zwojów) udowodniło, że współczesna elektronika półprzewodnikowa pozwala na miniaturyzację i zwiększenie bezpieczeństwa urządzeń, które pierwotnie wymagały ogromnych i niebezpiecznych iskierników oraz transformatorów zasilających o mocy wielu kilowatów. Eksperyment wykazał, że nawet w skali mikro, zjawisko sprzężenia rezonansowego jest niezwykle czułe na parametry komponentów, takie jak wartość rezystora bazowego (22 kΩ) czy obecność diody LED pełniącej funkcję ochronną i stabilizacyjną.

Szczególnie istotnym etapem prac było pomyślne zamontowanie i przetestowanie terminala górnego w postaci **torusa**. Realizacja tego elementu potwierdziła założenia teoretyczne dotyczące modyfikacji pojemności szczytowej układu. Dzięki torusowi udało się skutecznie obniżyć częstotliwość rezonansową całego systemu oraz zoptymalizować rozkład pola elektrycznego wokół uzwojenia wtórnego. W efekcie uzyskano stabilniejsze wyładowania oraz zwiększony zasięg bezprzewodowego przesyłu energii, co zamanifestowało się wyraźniejszą jonizacją gazów w lampach testowych. Potwierdza to, że precyzyjna geometria terminala wyjściowego jest równie ważna, co parametry elektryczne samego obwodu LC.

Analiza zastosowań przemysłowych pokazała, że wizja Tesli dotycząca bezprzewodowego przesyłu mocy przestała być jedynie domeną projektów teoretycznych. Firmy takie jak WiTricity czy Siemens Energy, implementując rezonansowe sprzężenie magnetyczne w ładowarkach do pojazdów elektrycznych i systemach diagnostycznych, udowadniają, że optymalizacja przesyłu energii bez użycia przewodów jest jednym z kluczowych kierunków rozwoju ekologicznego transportu i nowoczesnej energetyki. Transformator Tesli, ewoluując z widowiskowego generatora wyładowań do precyzyjnego narzędzia inżynieryjnego, potwierdza swoją rolę jako jeden z najważniejszych wynalazków w historii elektrotechniki.

Podsumowując, realizacja projektu w ramach koła naukowego pozwoliła nie tylko na zgłębienie wiedzy z zakresu fizyki i elektroniki, ale także na nabycie

umiejętności praktycznych w budowie i nastrojeniu układów wysokiej częstotliwości. Ukończony model fizyczny wraz z torusem stanowi kompletną bazę do dalszych badań nad optymalizacją cewek pierwotnych oraz systemami modulacji sygnału. Transformator Tesli pozostaje zatem nie tylko pomnikiem techniki, ale żywym, inspirującym tematem badawczym dla kolejnych pokoleń inżynierów.

Bibliografia

[1] Tesla Universe: Nikola Tesla Information Resource, https://teslauniverse.com (dostęp: 12 lutego 2026 r.).
[2] IEEE: Institute of Electrical and Electronics Engineers, https://www.ieee.org (dostęp: 12 lutego 2026 r.)
[3] Elert G.: The Physics Hypertextbook, https://physics.info (dostęp: 12 lutego 2026 r.)
[4] PBS: Tesla - Master of Lightning: Life and Legacy, https://www.pbs.org/tesla/ (dostęp: 12 lutego 2026 r.)
[5] ETHW: Nikola Tesla – Engineering and Technology History Wiki, https://ethw.org/Nikola_Tesla (dostęp: 12 lutego 2026 r.)
[6] WiTricity Corporation: How Wireless EV Charging Works, https://witricity.com (dostęp: 12 lutego 2026 r.).
[7] Tilbury M.: The Ultimate Tesla Coil Design and Construction Guide, McGraw-Hill Education, New York 2007.
[8] Martin T.C.: The Inventions, Researches and Writings of Nikola Tesla, The Electrical Engineer, New York 1894.
[9] Marinčić A.: Nikola Tesla and the Wireless Transmission of Energy.

Akty normatywne

[10] PN-EN 62305-1:2011: Ochrona odgromowa – Część 1: Zasady ogólne.
[11] PN-EN 60060-1:2011: Technika badań wysokich napięciowych – Część 1: Ogólne definicje i wymagania probiercze.

ANALYSIS AND PRACTICAL IMPLEMENTATION OF THE TESLA TRANSFORMER

Summary

This article is devoted to a multi-faceted theoretical analysis and practical implementation of a resonant transformer, commonly known as the Tesla coil, in a modern solid-state architecture. The primary objective of the conducted research was the design, construction, and verification of the operating parameters of a Slayer Exciter circuit, which represents a safe and efficient alternative to classical spark gap designs. In the research process, a method of critical analysis of the subject literature was used, covering both the classic works of Nikola Tesla and contemporary studies in the field of high-frequency engineering, as well as an experimental method based on prototype tests. The object of research was an electronic circuit based on the NPN D2394 power transistor, complemented by precision-made windings with 270 turns. In the course of the conducted experiments, it was shown that the key factor determining the efficiency of the device is the achievement of full electromagnetic resonance, which requires ideal

matching of the vibration frequency of the primary and secondary circuits. An essential element of the project was the use of a top terminal in the form of a torus, which allowed for the optimization of the electric field distribution and an increase in the peak capacity of the system. The results of the experiments confirmed the high efficiency of wireless energy transfer, demonstrating the possibility of exciting gases in discharge lamps without a direct galvanic connection. The conducted analysis proves that Tesla's historical concepts find real application today in wireless electric vehicle charging systems (WPT) and in modern medical diagnostics. The article concludes that the realized physical model constitutes an excellent educational tool, allowing for effective visualization of complex physical phenomena occurring in high-voltage circuits.

Keywords: resonant coil, magnetic induction, semiconductor technology, WPT systems, air ionization, bipolar transistor.

Martyna Bielenda, Karolina Kisała, Anna Turek
Koło Naukowe Systemów Złożonych

AUTOMATYZACJA ANALIZY STOMATOLOGICZNYCH OBRAZÓW RTG PRZY UŻYCIU ALGORYTMÓW GŁĘBOKIEGO UCZENIA

Streszczenie

Diagnostyka obrazowa stanowi fundamentalny element współczesnej praktyki stomatologicznej. Interpretacja pantomogramów jest jednak czasochłonna i podatna na błędy wynikające z nakładania się struktur oraz obecności szumów. Celem pracy była analiza skuteczności i optymalizacja modelu detekcji obiektów opartego na sieciach konwolucyjnych w zadaniu rozpoznawania struktur anatomicznych i patologii na zdjęciach RTG. W badaniach wykorzystano architekturę YOLOv11 w wariancie zorientowanych ramek ograniczających (Oriented Bounding Boxes – OBB) w wersjach Nano oraz Medium. Model trenowano na publicznym zbiorze Dental X-Ray Panoramic Dataset, na którym oznaczono kilkanaście klas obiektów, w tym zęby zatrzymane, ubytki próchnicowe, wypełnienia i implanty. Eksperymenty obejmowały badanie wpływu różnych strategii augmentacji (m.in. mozaikowania), skalowania modelu do większej pojemności oraz weryfikację wpływu rozkładu klas w zbiorze. Wykazano, że obecność marginalnych klas mniejszościowych znacząco obniża ogólne metryki skuteczności. Usunięcie klas incydentalnych z danych treningowych pozwoliło na poprawę średniej precyzji (mAP50) i poziomu 0,53 do 0,64. Ponadto wykazano, że testowanie modelu na obrazach o natywnych proporcjach panoramicznych, zamiast formatu kwadratowego z wypełnieniem, skutkuje skokowym wzrostem mAP50 do wartości 0,72. Uzyskane wyniki potwierdzają, że odpowiedni balans klas i dostosowanie rozdzielczości są kluczowe dla skutecznego wdrażania modeli detekcji w zautomatyzowanej diagnostyce stomatologicznej.

Słowa kluczowe: detekcja obiektów, sieci konwolucyjne, YOLOv11-OBB, diagnostyka stomatologiczna, uczenie głębokie.

1. Wprowadzenie

Diagnostyka obrazowa stanowi fundamentalny element współczesnej praktyki stomatologicznej, umożliwiając szybką i bezinwazyjną ocenę struktur anatomicznych [6]. Pantomogram dostarcza informacji niezbędnych do oceny stanu przyzębia, planowania leczenia oraz wykrywania patologii, takich jak ubytki

próchnicowe czy torbiele. Analiza obrazów pantomograficznych wiąże się jednak z wieloma wyzwaniami: obecnością szumów, nakładaniem się struktur anatomicznych oraz zróżnicowaniem rozmiarów diagnozowanych obiektów.

W ostatnich latach obserwuje się rosnące zainteresowanie wykorzystaniem głębokich sieci neuronowych do automatyzacji tego procesu, w tym architektur ukierunkowanych na detekcję obiektów w czasie rzeczywistym [7, 9]. Większość proponowanych w literaturze rozwiązań jest jednak ewaluowana na wysoce wyselekcjonowanych i zbalansowanych zbiorach danych. Próba przeniesienia tych modeli na surowe dane kliniczne często ujawnia ich podatność na zjawisko długiego ogona (long-tail distribution), co stanowi obecnie jedną z głównych barier aplikacyjnych w informatyce medycznej.

Wprowadzenie do oferty klinik zautomatyzowanej analizy zdjęć RTG za pomocą sieci konwolucyjnych mogłoby znacznie usprawnić proces diagnostyczny. Celem niniejszej pracy jest przeprowadzenie testów skuteczności modeli z rodziny YOLOv11-OBB [4] w zadaniu detekcji i klasyfikacji 17 różnych kategorii struktur stomatologicznych na obrazach panoramicznych.

2. Materiały i metody badawcze

W pracy wykorzystano publiczny zbiór danych Dental X-Ray Panoramic Dataset udostępniony na platformie Kaggle [5]. Baza zawiera 13932 unikalne obrazy dla formatu YOLO. Ze względu na naturalną krzywiznę łuku zębowego i zróżnicowane kąty nachylenia struktur, zastosowano zorientowane ramki ograniczające (YOLO-OBB), które eliminują problem nakładania się na siebie standardowych, prostokątnych ramek i poprawiają dokładność procesu detekcji. Przykłady obrazów w formacie kwadratowym oraz prostokątnym przedstawiono na rys. 1. oraz 2.

Charakterystykę zbioru przedstawiono na rys. 3. Wykres wprost uwidacznia silną asymetrię rozkładu – dominującą część zbioru stanowią zdjęcia tła oraz najczęstsze struktury anatomiczne, podczas gdy przypadki nietypowe posiadają znikomą reprezentację ilościową. Mając na uwadze, że pozostawienie niedostatecznie reprezentowanych klas w zbiorze treningowym niesie zagrożenie dla stabilności modelu oraz prowadzi do ryzyka przeuczenia (overfittingu) [3], podjęto decyzję o odrzuceniu marginalnych klas, których reprezentacja nie przekraczała 0,05% całkowitej liczby instancji.

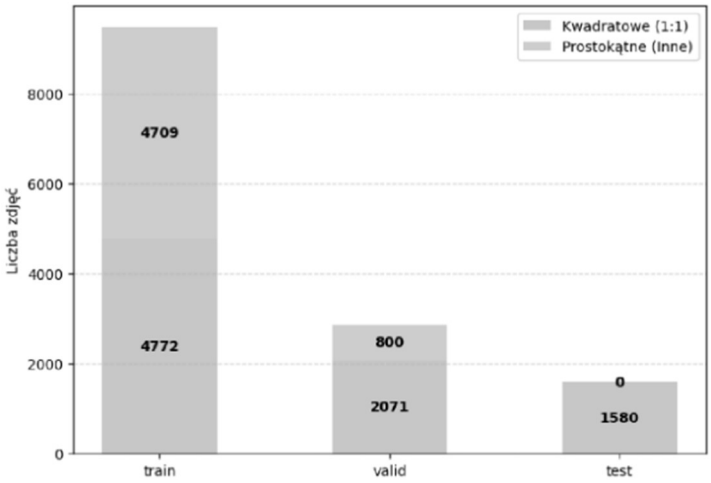
Jako architekturę wybrano model sieci YOLO11-OBB [4] w dwóch wariantach pojemnościowych: Nano (2,7 mln parametrów) oraz Medium (20,9 mln parametrów). Do trenowania modeli wykorzystano m.in. środowisko obliczeniowe Google Colaboratory [1].



Rys. 1. Zdjęcie w formacie kwadratowym (640×640 px)
Fig. 1. Square format image (640×640 px)



Rys. 2. Zdjęcie w formacie panoramicznym (prostokątnym) (1615×840 px)
Fig. 2. Image with oriented bounding boxes (OBB)



Rys. 3. Wykres rozkładu kształtów obrazów w podzbiorach (train/valid/test)
Fig. 3. Class distribution chart

3. Przebieg eksperymentów i wyniki

W celu wyłonienia optymalnej strategii detekcji, przeprowadzono ewaluację modelu przy różnych konfiguracjach hiperparametrów, wariantach augmentacji oraz pojemnościach architektury. Zestawienie głównych scenariuszy testowych zaprezentowano w tabeli 1.

Tabela 1. Zestawienie badanych konfiguracji eksperymentalnych

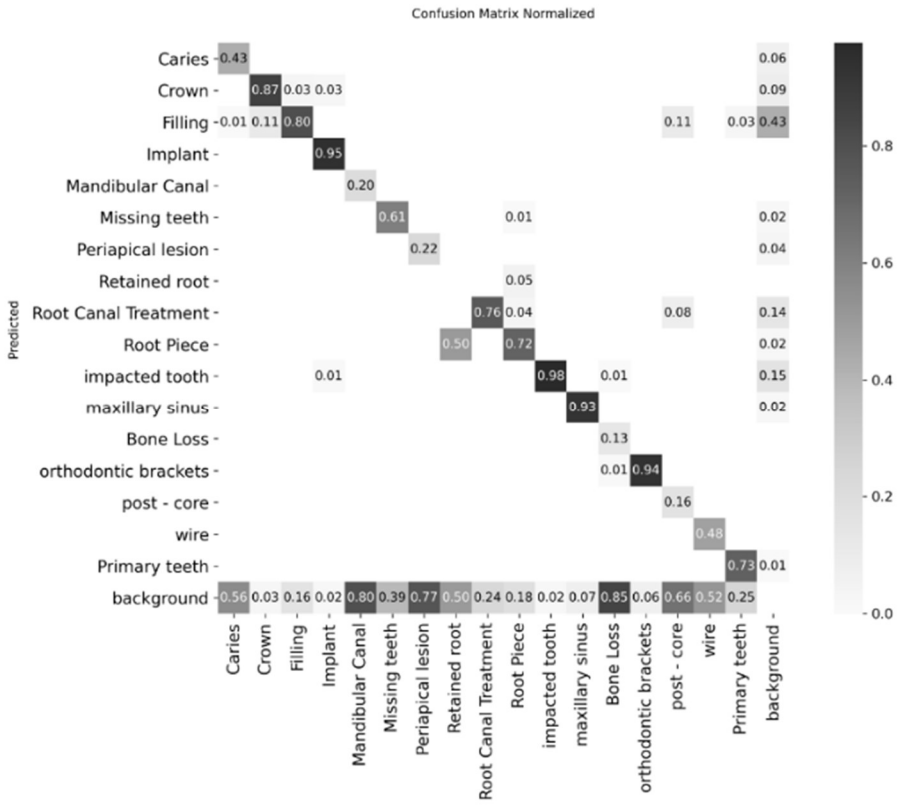
Table 1. Summary of experimental configurations tested

Wariant testowy	**Model**	**Liczba parametrów**	**Zastosowana augmentacja danych**
Baza	YOLO11-OBB Nano	2,7 mln	Silna (mozaikowanie, zmienność barwy, obrót)
Wariant A	YOLO11-OBB Nano	2,7 mln	Ograniczona (bez zniekształceń geometrycznych)
Wariant Zoptymalizowany	YOLO11-OBB Nano	2,7 mln	Zbalansowana (hybrydowa)
Skalowanie	YOLO11-OBB Medium	20,9 mln	Zbalansowana

Odpowiednio zaplanowany proces uczenia wymagał iteracyjnego podejścia. W początkowej fazie badano model Nano. Trening bazowy ze stosunkowo silną augmentacją (m.in. mozaikowanie, zmienność barwy, obrót) dawał dobre rezultaty dla wyraźnych struktur, ale gorzej radził sobie z mniejszymi obiektami.

Ograniczenie augmentacji poprawiło precyzję, lecz drastycznie obniżyło czułość (zdolność wykrywania). Połączenie obu podejść i zastosowanie zbalansowanej augmentacji pozwoliło na osiągnięcie najlepszych wyników w grupie modeli Nano, co przełożyło się na wskaźnik mAP50 wynoszący 0,532. Szczegółowe problemy modelu z rozpoznawaniem poszczególnych klas obrazuje macierz błędów (rys. 4.).

Kolejnym krokiem było skalowanie zoptymalizowanej konfiguracji do modelu o większej pojemności (Medium). Wymagało to odpowiedniego dostosowania parametrów partii w celu eliminacji zjawiska eksplozji gradientów prowadzącego do błędów numerycznych.



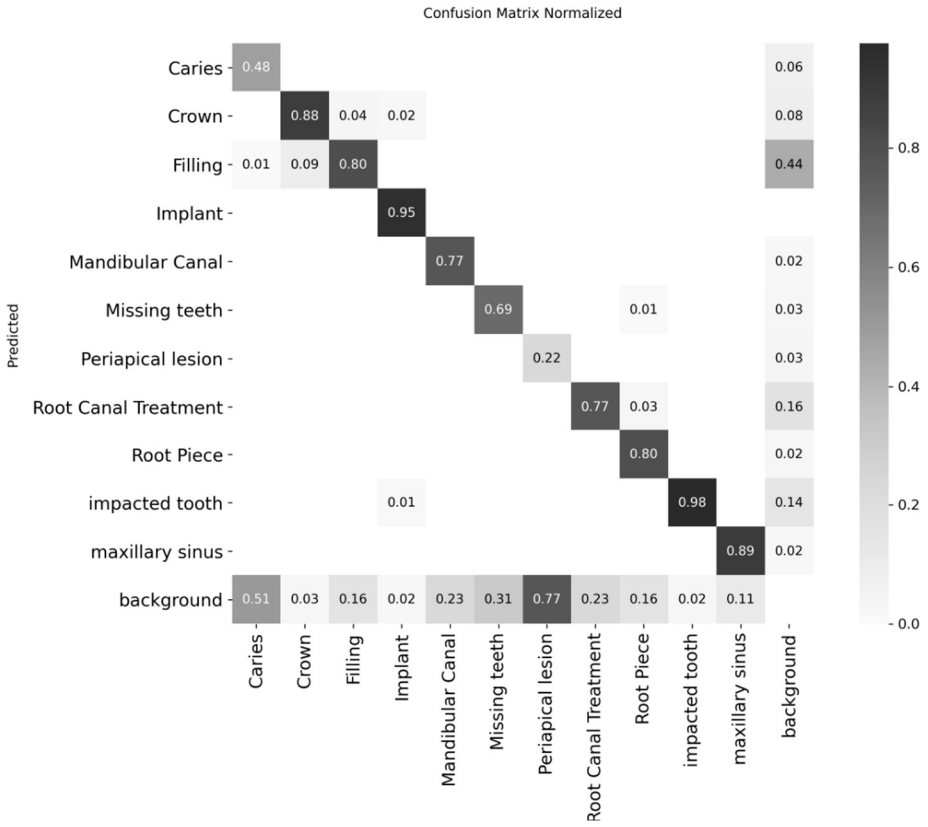
Rys. 4. Model YOLOv11 Nano (najlepsza konfiguracja augmentacji) – macierz błędów (znormalizowana)

Fig. 4. YOLOv11 Nano model (best augmentation configuration) – normalized confusion matrix

W analizie wyników dostrzeżono, że na ogólne metryki wydajności wciąż negatywnie wpływają rzadkie klasy mniejszościowe [11]. Przeprowadzono Trening weryfikacyjny 1, usuwając ze zbioru klasy o krytycznie małej ilości reprezentantów. Skutkowało to skokową poprawą i uzyskaniem przez model mAP50 na poziomie 0,642. Zredukowanie szumu informacyjnego wpłynęło na znaczne polepszenie precyzji dla pozostałych klas, co obrazuje znormalizowana macierz błędów (rys. 5.).

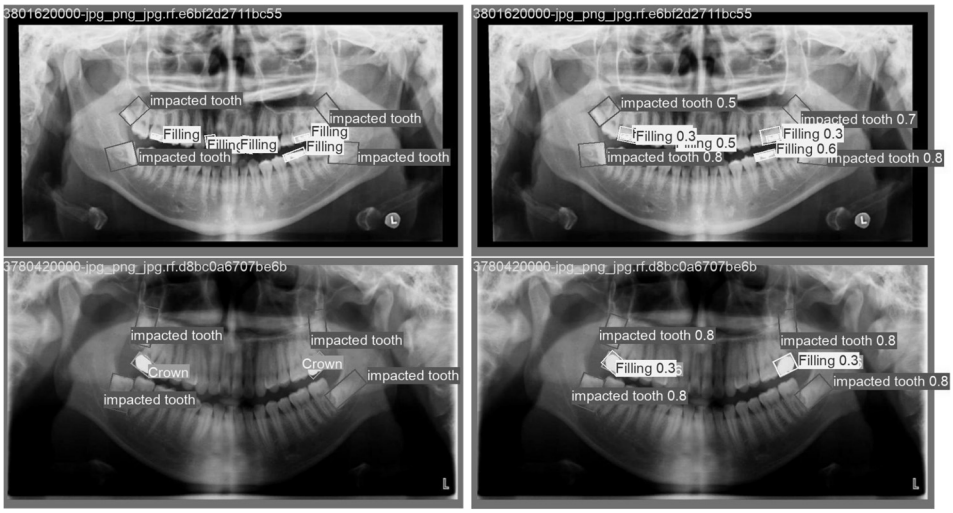
Analiza porównawcza zaprezentowanych macierzy uwidacznia w jak dużym stopniu eliminacja klas incydentalnych zredukowała odsetek fałszywych detekcji. Sieć przestała marnować swoją pojemność na identyfikację struktur o śladowej reprezentacji w zbiorze uczącym, co przełożyło się na wyższą pewność predykcji dla kluczowych elementów, takich jak ubytki próchnicowe, wypełnienia czy implanty. Potwierdziło to tezę, że w medycznych zbiorach danych charakteryzujących się skrajnym niezbalansowaniem i asymetrią rozkładu klas, rygorystyczna

selekcja i balansowanie klas są równie istotne jak dobór samej architektury. Niemniej jednak, szczegółowa analiza wizualna przewidywań wskazywała, że potencjał modelu wariantu OBB nadal jest tłumiony przez nienaturalną zmianę proporcji i utratę skali, wynikającą z kompresji oryginalnych, szerokich pantomogramów do standardowego dla algorytmów YOLO formatu kwadratu.



Rys. 5. Trening Weryfikacyjny 1 - macierz błędów (znormalizowana)
Fig. 5. Verification Training 1 - normalized confusion matrix

Dodatkowo sprawdzono wpływ formatu zdjęć. Zamiast testować model na sztucznie przeskalowanych obrazach w formacie kwadratu, Trening weryfikacyjny 2 przeprowadzono na obrazach zachowujących natywne, prostokątne proporcje pantomograficzne. Przyniosło to znaczną poprawę detekcji. Przykładowe rezultaty działania zoptymalizowanego modelu na zdjęciach panoramicznych przedstawiono na rys. 6.



Rys. 6. Przykładowe wyniki detekcji za pomocą zoptymalizowanego modelu YOLOv11-OBB: a) poprawne etykiety na zbiorze walidacyjnym, b) przewidywania modelu na zbiorze walidacyjnym

Fig. 6. Example detection results using the optimized YOLOv11-OBB model: a) Verification Training 1 - ground truth labels on the validation set, b) Verification Training 1 - model predictions on the validation set

Zestawienie wyników dla obu badanych wariantów rozdzielczości przedstawiono w tabeli 2. Jak wynika z zaprezentowanych danych, zachowanie naturalnych proporcji obrazu (1.92:1) pozwoliło na wyraźny wzrost wskaźnika mAP50 z poziomu 0,6423 do 0,7263, a także zauważalny wzrost precyzji (Precision). Mimo nieznacznego spadku czułości (Recall), ogólna zdolność modelu do prawidłowej klasyfikacji i dokładnego pozycjonowania ramek (mAP50-95) uległa znacznej poprawie, co ostatecznie potwierdza słuszność rezygnacji ze sztucznego dopełniania obrazów do formatu kwadratu.

Tabela 2. Porównanie wyników walidacji standardowej oraz testu weryfikacyjnego na obrazach o natywnych proporcjach panoramicznych

Table 2. Comparison of standard validation and verification test results on native panoramic aspect ratio images

Metryka	**Walidacja Standardowa (Zbiór Kwadratowy 1:1)**	**Test Weryfikacyjny (Zbiór Panoramiczny 1.92:1)**
Rozdzielczość (imgsz)	1024x1024 (Padding)	1615x840 (Native)
Precision	0,6202	0,7193
Recall	0,7655	0,7012
mAP50	0,6423	0,7263
mAP50-95	0,4228	0,5127

4. Podsumowanie

Niniejszy projekt ukazał, jak ogromne znaczenie dla wydajności klasyfikatorów medycznych ma odpowiednia reprezentacja danych. Najważniejszym odkryciem było zidentyfikowanie, że obecność marginalnych klas mniejszościowych w zbiorze znacząco zaburza metryki skuteczności całego modelu. Po ich wyeliminowaniu model osiągnął zdecydowanie lepsze wyniki.

Co więcej, ostatecznie wykazano, że dostosowanie rozdzielczości obrazu wejściowego i analiza obrazów zachowujących natywne proporcje zdjęć pantomograficznych stanowi ważny element dalszej optymalizacji skuteczności. Sieci z rodziny YOLOv11 wsparte ramkami zorientowanymi wykazują znakomity potencjał aplikacyjny jako zautomatyzowane asystenty diagnostyki stomatologicznej.

Kierunkiem dalszych prac badawczych mogłaby być weryfikacja zaproponowanego podejścia na niezależnych, wieloośrodkowych zbiorach danych klinicznych, co pozwoliłoby na ocenę odporności modelu na różnice w parametrach ekspozycji czy specyfice sprzętu rentgenowskiego różnych producentów. Dodatkowo, docelowe zintegrowanie tak zoptymalizowanego algorytmu z oprogramowaniem gabinetowym stworzyłoby środowisko „drugiego czytania", znacząco redukując czas analizy pantomogramów i minimalizując ryzyko błędu ludzkiego przy wykrywaniu wczesnych zmian patologicznych.

Literatura

[1] Bisong E.: Google Colaboratory. W: Building Machine Learning and Deep Learning Models on Google Cloud Platform. Apress, 2019, s. 59-64.

[2] Bonfanti-Gris M., Herrera A. i in.: Deep learning for tooth detection and segmentation in panoramic radiographs: a systematic review and meta-analysis. W: BMC Oral Health, 2025.

[3] Buda M., Maki A., Mazurowski M. A.: A systematic study of the class imbalance problem in convolutional neural networks. W: Neural Networks 106, 2018, s. 249-259.

[4] Jocher G., Qiu J.: Ultralytics YOLO11. Wersja 11.0.0. 2024.

[5] Raj A.: Dental X-Ray Panoramic Dataset. Kaggle, 2022.

[6] Różyło-Kalinowska I., Różyło T. K.: Współczesna radiologia stomatologiczna. Wydanie 3. Wydawnictwo Czelej.

[7] Singh N. K., Raza K.: Progress in deep learning-based dental and maxillofacial image analysis: A systematic review. W: Expert Systems with Applications, 2022.

[8] Tarighatnia A., Amanzadeh M., Hamedan M. i in.: Deep learning-based evaluation of panoramic radiographs for osteoporosis screening: a systematic review and meta-analysis. W: BMC Medical Imaging, 2025.

[9] Turosz N., Chęcińska K., Chęciński M. i in.: Applications of artificial intelligence in the analysis of dental panoramic radiographs: an overview of systematic reviews. W: Dentomaxillofacial Radiology, 2023.

[10] Umer F., Habib S., Adnan N.: Application of deep learning in teeth identification tasks on panoramic radiographs. W: Dentomaxillofacial Radiology, 2022.

[11] Wu Z. i in.: Medical long-tailed learning for imbalanced data: Bibliometric analysis. W: Computer Methods and Programs in Biomedicine 247, 2024, s. 108106.

AUTOMATION OF DENTAL X-RAY IMAGE ANALYSIS USING DEEP LEARNING ALGORITHMS

Summary

Diagnostic imaging is a fundamental element of modern dental practice. However, interpreting pantomograms is time-consuming and prone to errors resulting from overlapping structures and noise. The aim of this study was to analyze the effectiveness and optimize an object detection model based on convolutional neural networks for the task of recognizing anatomical structures and pathologies in X-ray images. The YOLOv11 architecture in the Oriented Bounding Boxes (OBB) variant was utilized, testing both Nano and Medium versions. The model was trained on the public Dental X-Ray Panoramic Dataset, which included several object classes such as impacted teeth, caries, fillings, and implants. The experiments involved testing various data augmentation strategies, scaling the model to a higher capacity, and verifying the impact of class balance. It was shown that the presence of marginal minority classes significantly lowers overall performance metrics. Removing incidental classes from the training data allowed for an improvement in mean Average Precision (mAP50) from 0.53 to 0.64. Furthermore, testing the model on images with native panoramic proportions, rather than a padded square format, resulted in a sharp increase in mAP50 to 0.72. The obtained results confirm that proper class balance and aspect ratio adjustment are crucial for the effective deployment of detection models in automated dental diagnostics.

Keywords: object detection, convolutional neural networks, YOLOv11-OBB, dental diagnostics, deep learning.

Anna Turek, Karolina Kisała, Martyna Bielenda
Koło Naukowe Systemów Złożonych

PREDYKCJA PRODUKCJI BIOGAZU Z UWZGLĘDNIENIEM CZYNNIKÓW METEOROLOGICZNYCH W PROCESIE FERMENTACJI METANOWEJ

Streszczenie

W pracy przedstawiono model predykcyjny produkcji biogazu w procesie fermentacji metanowej, uwzględniający wpływ czynników meteorologicznych. W przeciwieństwie do klasycznych podejść opartych wyłącznie na parametrach wsadu, zaproponowane rozwiązanie rozszerza wektor wejściowy o zmienne egzogeniczne, takie jak temperatura otoczenia, suma opadów oraz wilgotność względna, wraz z ich opóźnieniami czasowymi wynikającymi z inercyjnej natury procesu biologicznego. Ze względu na ograniczoną liczebność dostępnych danych eksploatacyjnych zastosowano technikę dezagregacji przestrzennej, polegającą na przekształceniu danych z czterech równolegle pracujących komór fermentacyjnych w zbiór danych panelowych. Pozwoliło to na zwiększenie liczby obserwacji uczących oraz poprawę zdolności generalizacji modelu. Do estymacji wykorzystano algorytm Random Forest, który umożliwia odwzorowanie nieliniowych zależności pomiędzy parametrami procesowymi a produkcją biogazu. W rygorystycznej procedurze walidacji Leave-One-Day-Out (LODO) uzyskano współczynnik determinacji na poziomie $R^2 = 0{,}756$ oraz średni bezwzględny błąd procentowy $MAPE = 7{,}4\%$, co potwierdza wysoką dokładność prognoz w warunkach przemysłowych. Uzyskane wyniki wskazują na istotne znaczenie czynników pogodowych w modelowaniu procesu fermentacji oraz na możliwość wykorzystania opracowanego rozwiązania w systemach sterowania predykcyjnego i wspomagania decyzji w oczyszczalniach ścieków.

Słowa kluczowe: efektywność energetyczna, modelowanie procesów biologicznych, uczenie maszynowe, Random Forest, zmienne egzogeniczne, dezagregacja danych.

1. Wstęp

Sprawnie funkcjonująca gospodarka wodno-ściekowa stanowi istotny element ochrony środowiska, wpływając zarówno na jakość ekosystemów wodnych, jak i na bilans energetyczny oczyszczalni. Odpowiednie prowadzenie procesów

technologicznych może zwiększać efektywność energetyczną obiektu oraz ograniczać jego oddziaływanie na środowisko [13]. Skład i ilość dopływających ścieków determinują przebieg fermentacji metanowej, która stanowi podstawę produkcji biogazu i dążenia do samowystarczalności energetycznej [16, 17].

Dotychczasowe badania autorów koncentrowały się na predykcji wydajności produkcji biogazu w oparciu o parametry wsadu [16, 17]. W niniejszej pracy rozszerzono analizę o wpływ czynników meteorologicznych, wykorzystując model Random Forest do prognozowania produkcji biogazu w warunkach zmiennych środowiskowo. Takie podejście pozwala na wcześniejszą identyfikację potencjalnych spadków wydajności procesu i może stanowić wsparcie decyzyjne w eksploatacji instalacji.

1.1. Kontekst energetyczny i specyfika procesu

Fermentacja metanowa jest obecnie postrzegana jako jeden z kluczowych elementów strategii prowadzących do zwiększenia samowystarczalności energetycznej oczyszczalni ścieków. W warunkach krajowych sprawnie funkcjonujące układy kogeneracyjne mogą pokrywać od około 40% do nawet 70–80% zapotrzebowania obiektu na energię elektryczną, co potwierdzają analizy dotyczące energochłonności instalacji w regionie południowo-wschodnim [11]. Oznacza to, że produkcja biogazu stanowi istotny składnik bilansu energetycznego oczyszczalni, a nie jedynie produkt uboczny procesu.

Wykorzystanie własnych zasobów osadowych ma znaczenie zarówno ekonomiczne, jak i strategiczne, zwiększając bezpieczeństwo energetyczne obiektu oraz ograniczając zależność od zewnętrznych dostaw energii [19]. Jednocześnie, mimo przeprowadzonych modernizacji infrastruktury [12], proces fermentacji pozostaje układem nieliniowym i wrażliwym na zmienność parametrów wsadu. Złożoność ta uzasadnia potrzebę stosowania zaawansowanych systemów sterowania wspierających stabilizację produkcji metanu w warunkach zmiennego dopływu [4].

1.2. Luka badawcza: brak systemów preskrypcyjnych w średnich obiektach

Współczesna inżynieria środowiska coraz silniej akcentuje cyfryzację procesów technologicznych oraz wykorzystanie narzędzi opartych na analizie danych. Integracja modelowania matematycznego z systemami sterowania i automatyki (ICA) jest wskazywana jako istotny czynnik poprawy efektywności energetycznej oraz redukcji śladu węglowego oczyszczalni [4]. Jednocześnie wdrażanie takich rozwiązań napotyka ograniczenia wynikające z nieliniowego charakteru procesów biologicznych oraz niskiej interpretowalności modeli typu „czarna skrzynka", co utrudnia ich wykorzystanie w systemach decyzyjnych wykraczających poza samą predykcję [2].

Rozwój technologii omicznych oraz zaawansowanych metod sztucznej inteligencji ukierunkowany jest obecnie na kompleksową bioremediację i odzysk zasobów [18]. Rozwiązania te wymagają jednak znaczących nakładów finansowych

i dostępu do rozbudowanych zbiorów danych, co ogranicza ich praktyczne zastosowanie w średnich obiektach komunalnych. W konsekwencji pojawia się luka między potencjałem technologii a możliwościami wdrożeniowymi, a także niedobór systemów proaktywnych, które integrowałyby zmienne egzogeniczne – w tym dane meteorologiczne – w celu generowania optymalnych strategii działania w czasie rzeczywistym [14, 7].

1.3. Cel pracy

Celem niniejszej pracy jest opracowanie oraz walidacja modelu predykcyjnego produkcji biogazu w procesie fermentacji metanowej z uwzględnieniem czynników meteorologicznych. Proponowane podejście rozszerza klasyczne modelowanie o zmienne egzogeniczne, takie jak temperatura otoczenia i opady atmosferyczne, wpływające na dynamikę procesu.

W pracy zastosowano technikę dezagregacji przestrzennej (Spatial Disaggregation) w celu zwiększenia efektywnej liczebności zbioru uczącego oraz poprawy zdolności generalizacji modelu. Uzyskane rozwiązanie stanowi próbę praktycznego wykorzystania metod uczenia maszynowego w warunkach ograniczonej dostępności danych przemysłowych.

2. Obiekt badań i inżynieria danych

Układ technologiczny miejskiej oczyszczalni ścieków w Rzeszowie był sukcesywnie modernizowany w kilku etapach inwestycyjnych realizowanych w latach 1992–1998, 2001, 2003–2004 oraz 2013–2014. Wprowadzone zmiany infrastrukturalne wpłynęły na charakter pracy instalacji, co uzasadnia konieczność okresowej analizy i weryfikacji aktualnych parametrów eksploatacyjnych obiektu. Oczyszczalnia zlokalizowana jest w północnej części aglomeracji rzeszowskiej, przy ul. Ciepłowniczej 2, wzdłuż prawego brzegu rzeki Wisłok [10].

3. Metodyka badań

3.1. Charakterystyka zbioru danych i imputacja

Badania przeprowadzono na podstawie danych eksploatacyjnych z okresu od 1 stycznia do 31 sierpnia 2024 roku (244 dni operacyjne). Analiza ciągłości wykazała, że 11,1% obserwacji (27 dni) zawierało braki danych. Ze względu na biologiczną bezwładność procesu fermentacji przyjęto hybrydową strategię imputacji.

Dla krótkich luk (do 2 dni) zastosowano interpolację liniową, natomiast dla dłuższych przerw wykorzystano uśrednianie danych z siedmiodniowych okresów poprzedzających i następujących po incydencie. Podejście to pozwoliło zachować charakterystykę statystyczną procesu bez wprowadzania sztucznych trendów.

Imputację opisano zależnością (1):

$$y_{imp}(t) = \begin{cases} y(t_{start}) + (t - t_{start}) \frac{y(t_{end}) - y(t_{start})}{t_{end} - t_{start}} & dla\, \Delta t \leq 2\ dni \\ \frac{1}{2}\left(\frac{1}{N_{pre}} \sum_{i=1}^{N_{pre}} y\,(t_{start} - i) + \frac{1}{N_{post}} \sum_{j=1}^{N_{post}} y\,(t_{end} + j) \right) & dla\, \Delta t > 2\ dni \end{cases} \quad (1)$$

gdzie Δt oznacza długość luki pomiarowej, a $N_{pre} = N_{post} = 7$.

Zgodnie ze wzorem (1) proces imputacji braków danych realizowano z wykorzystaniem interpolacji liniowej dla krótkich luk oraz uśredniania dla dłuższych przerw pomiarowych. Do danych procesowych dołączono informacje meteorologiczne pozyskane z serwisu Open-Meteo [20]. Uwzględniono temperaturę otoczenia, sumę opadów oraz wilgotność względną. Dane godzinowe zagregowano do wartości dobowych, aby zapewnić zgodność z rozdzielczością danych technologicznych. Średnia całkowita produkcja biogazu dla instalacji (suma czterech komór) wynosiła około 6600 m³/d, natomiast modelowanie prowadzono na poziomie pojedynczych reaktorów.

Imputacja krótkoterminowa (luka 1-dniowa).

Dla pojedynczych braków zastosowano interpolację liniową między wartościami brzegowymi. Metoda ta jest uznawana za wystarczającą w przypadku krótkich luk, przy zachowanej ciągłości dynamiki procesu [1]. Jej skuteczność dla niewielkich braków danych potwierdzono również w innych badaniach [3, 15]

Imputacja średnio- i długoterminowa (luki 4–17 dni).

W przypadku dłuższych przerw odstąpiono od interpolacji liniowej, która nie odzwierciedla nieliniowego charakteru procesu i może prowadzić do zniekształceń statystycznych [9, 5]. Zastosowano uśrednianie danych z siedmiodniowych okresów referencyjnych przed i po incydencie, co pozwoliło zachować charakterystykę procesu bez wprowadzania sztucznych trendów.

Decyzję o imputacji, zamiast usuwania obserwacji, podjęto w celu zachowania ciągłości szeregów czasowych niezbędnej dla algorytmów uczących oraz ze względu na stabilność biologiczną procesu, ograniczającą ryzyko gwałtownych fluktuacji. W analizowanym okresie zidentyfikowano cztery incydenty utraty danych obejmujące łącznie 27 dni (11,1% zbioru): jedną lukę jednodniową, dwie luki średnioterminowe (4–5 dni) oraz jedną lukę 17-dniową. Dla krótkiej przerwy zastosowano interpolację liniową, natomiast w przypadku dłuższych luk wykorzystano uśrednianie danych z siedmiodniowych okresów referencyjnych przed i po incydencie.

3.2. Augmentacja danych poprzez dezagregację przestrzenną

Jednym z głównych ograniczeń modelowania procesu fermentacji z wykorzystaniem metod uczenia maszynowego jest niewielka liczba dostępnych obserwacji. Jak w swojej pracy wskazują Cruz i inni, zbyt mały zbiór uczący zwiększa ryzyko przeuczenia i obniża zdolność generalizacji modelu [2]. W analizowanym okresie

zgromadzono $N = 244$ dobowe wektory danych dla całego obiektu, co stanowiło ograniczenie przy zastosowaniu algorytmów takich jak Random Forest.

W celu zwiększenia liczebności zbioru zastosowano technikę przestrzennej dezagregacji danych (Spatial Disaggregation), wykorzystującą równoległą pracę czterech komór fermentacyjnych. Podejście to, zgodne z koncepcją modelowania globalnego dla jednostek równoległych [8] pozwoliło na zwiększenie efektywnej liczby obserwacji oraz poprawę zdolności generalizacji modelu. Statystyczną charakterystykę zmiennych dla zbioru po augmentacji przedstawiono w tabeli 1.

Oczyszczalnia eksploatuje cztery równoległe komory fermentacyjne (WKF A–D) o zbliżonej konstrukcji i parametrach pracy. Pozwoliło to przekształcić dane zagregowane dla całego obiektu w zbiór danych panelowych poprzez rozdzielenie każdej doby t na cztery obserwacje odpowiadające poszczególnym reaktorom.

$$D_{augmented} = D \ \bigcup_{k\in\{A,B,C,D\}}\{(x_{t,k}, y_{t,k}) |\ t \in \{1, \dots, T\}\} \tag{2}$$

gdzie $D_{augmented}$ oznacza zbiór danych po augmentacji, D zbiór danych pierwotnych, $k \in \{A, B, C, D\}$ identyfikator komory fermentacyjnej, $x_{(t,k)}$ wektor zmiennych wejściowych dla doby ti komory k, $y_{(t,k)}$ wartość zmiennej wyjściowej dla doby t i komory k, natomiast T oznacza liczbę analizowanych dób.

Zgodnie ze wzorem (2) przeprowadzono augmentację danych poprzez dezagregację przestrzenną, umożliwiającą zwiększenie liczebności zbioru uczącego. W rezultacie liczebność zbioru uczącego wzrosła z $N = 244$ do $N = 976$, co ograniczyło ryzyko przeuczenia i poprawiło zdolność generalizacji modelu [8]. Statystyki opisowe zmiennych po augmentacji przedstawiono w tabeli 1.

Tabela 1. Statystyki opisowe zmiennych procesowych i pogodowych po augmentacji (N = 976)

Table 1. Descriptive statistics of process and meteorological variables after data augmentation (N = 976)

Zmienna	Jednostka	Średnia	Odch. std.	Min	Max
Produkcja biogazu (Qbio)	m³/d	1652.31	324.84	962.00	3005.00
Osad wstępny (QOW)	m³/d	118.98	540.00	35.10	7639.00
Osad nadmierny (QON)	m³/d	45.32	7.11	25.50	83.20
Stosunek OW/ON (ROW/ON)	–	2.57	10.51	0.74	148.91
Obciążenie ogólne (LA)	kg s.m./m³	2.96	8.27	0.00	117.90
Obciążenie organiczne (LA')	kg s.m.o./m³	2.19	6.28	0.00	89.50
Temperatura otoczenia (Tamb)	°C	12.88	8.45	-14.70	26.90
Suma opadów (Psum)	mm	2.08	3.79	0.00	31.80
Wilgotność powietrza (Havg)	%	74.31	10.37	47.00	95.00

Podejście to, określane jako Global Modeling for Parallel Units, zakłada wspólną charakterystykę kinetyczną równoległych komór fermentacyjnych. Model uczony na tak przygotowanym zbiorze jest bardziej odporny na lokalne zakłócenia i lepiej generalizuje zależność między obciążeniem a produkcją metanu.

3.3. Algorytm optymalizacji sterowania

W niniejszej pracy skoncentrowano się na budowie i walidacji modelu predykcyjnego produkcji biogazu. Model Random Forest traktowany jest jako narzędzie prognostyczne wspierające ocenę wpływu parametrów wsadu i warunków meteorologicznych na wydajność procesu. Na podstawie prognoz możliwa jest identyfikacja okresów podwyższonego ryzyka spadków produkcji, co stanowi punkt wyjścia do dalszych prac nad modułem optymalizacyjnym.

3.4. Ocena jakości modelu i walidacji

Ze względu na strukturę zbioru augmentowanego, w którym każda doba reprezentowana jest przez cztery skorelowane próbki, zastosowanie standardowej walidacji krzyżowej mogłoby prowadzić do wycieku informacji (data leakage). W celu rzetelnej oceny zdolności generalizacji modelu wdrożono dwustopniową procedurę walidacyjną obejmującą LOO oraz LODO, której charakterystykę zestawiono w tabeli 2.

W procedurze LODO zbiór treningowy tworzony jest poprzez usunięcie wszystkich próbek należących do testowanej doby t. Eliminuje to wyciek informacji wynikający z korelacji przestrzennej między komorami i odzwierciedla rzeczywiste warunki prognozowania.

Tabela 2. Zastosowane procedury walidacyjne

Table 2. Applied validation procedures

Procedura	Zakres walidacji	Charakter oceny
LOO (Leave-One-Out)	Model trenowany na całym zbiorze z wyłączeniem jednej próbki	Ocena spójności wewnętrznej; możliwe współdzielenie informacji z tej samej doby
LODO (Leave-One-Day-Out)	Usunięcie wszystkich próbek należących do testowanej doby t	Rygorystyczna ocena zdolności predykcyjnej bez wycieku informacji

Zgodnie ze wzorem (3) zbiór treningowy w procedurze walidacyjnej konstruowano poprzez wykluczenie wszystkich obserwacji należących do analizowanej doby.

$$D_{train}^{t} = D \setminus \{(x_{t,k}, y_{t,k})\},\ k = 1, \ldots, 4 \tag{3}$$

gdzie D_t oznacza zbiór wszystkich próbek należących do doby t, a k identyfikator komory. Do oceny dokładności predykcji zastosowano błąd średniokwadratowy

RMSE zdefiniowany wzorem (4). Kolejną wykorzystaną miarą jakości modelu był średni bezwzględny błąd procentowy MAPE opisany zależnością (5). Uzupełnieniem analizy był współczynnik determinacji R^2, przedstawiony wzorem (6), pozwalający ocenić stopień dopasowania modelu do danych rzeczywistych

$$RMSE = \sqrt{\frac{1}{N_R} * \sum_{i \in R}(Q_{bio,i} - \hat{Q}_{bio,i})^2} \tag{4}$$

$$MAPE = \frac{1}{N_R} * \sum_{i \in R} \left| \frac{(Q_{bio,i} - \hat{Q}_{bio,i})}{Q_{bio,i}} \right| * 100\% \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i \in R}(Q_{bio,i} - \hat{Q}_{bio,i})^2}{\sum_{i \in R}(Q_{bio,i} - \bar{Q}_{bio,i})^2} \tag{6}$$

gdzie: $Q_{bio,i}$ – oznacza rzeczywistą produkcję biogazu,
$\hat{Q}_{bio,i}$ – wartość prognozowaną przez model,
$\bar{Q}_{bio}$ – średnią produkcję w zbiorze walidacyjnym,
N_R – liczbę próbek walidacyjnych.

4. Wyniki i dyskusja

4.1. Analiza korelacji i wydajność modelu predykcyjnego

W celu identyfikacji zależności między zmiennymi procesowymi, pogodowymi a produkcją biogazu przeprowadzono analizę korelacji z wykorzystaniem współczynników Pearsona (r) oraz Spearmana (ρ), a jej wyniki zestawiono w tabeli 3. [6].

Najsilniejszy związek odnotowano dla temperatury otoczenia z jednodniowym opóźnieniem ($T^{-1}amb$), co potwierdza istotny wpływ warunków termicznych na dynamikę fermentacji. Istotną zależność wykazano również dla wilgotności względnej z opóźnieniem pięciodniowym. W przypadku opadów wpływ miał charakter krótkoterminowy i słabszy.

Zastosowanie współczynnika Spearmana potwierdziło nieliniowy, monotoniczny charakter zależności, co uzasadnia uwzględnienie opóźnionych zmiennych meteorologicznych w modelu predykcyjnym.

Tabela 3. Korelacja zmiennych procesowych i wyselekcjonowanych wskaźników pogodowych z produkcją biogazu

Table 3. Correlation between process and meteorological variables and biogas production

Zmienna wejściowa	Jednostka	Pearson (r)	Spearman (ρ)
Osad wstępny (QOW)	m³/d	-0.019	-0.137
Osad nadmierny (QON)	m³/d	0.049	0.038
Stosunek OW/ON (ROW/ON)	-	-0.022	-0.123

Tabela 3 (cd.). Korelacja zmiennych procesowych i wyselekcjonowanych wskaźników pogodowych z produkcją biogazu

Table 3 (cont.). Correlation between process and meteorological variables and biogas production

Zmienna wejściowa	Jednostka	Pearson (r)	Spearman (ρ)
Obciążenie ogólne (LA)	kg s.m./m³	-0.026	-0.134
Obciążenie organiczne (LA')	kg s.m.o./m³	-0.022	-0.053
Identyfikator komory (IDkomory)	-	0.484	0.524
Objętość czynna WKF (Vczynna)	m³	0.558	0.595
Temperatura (Lag 1) (T^{-1}amb)	°C	-0.557	-0.483
Suma opadów (Lag 2) (P^{-2}sum)	mm	0.022	0.084
Wilgotność (Lag 5) (H^{-5}avg)	%	0.395	0.404

Analiza agregacji tygodniowej, której wyniki przedstawiono w tabeli 4., wykazała utrzymującą się korelację temperatury i wilgotności przez okres 28 dni, co potwierdza inercyjny charakter procesu fermentacji.

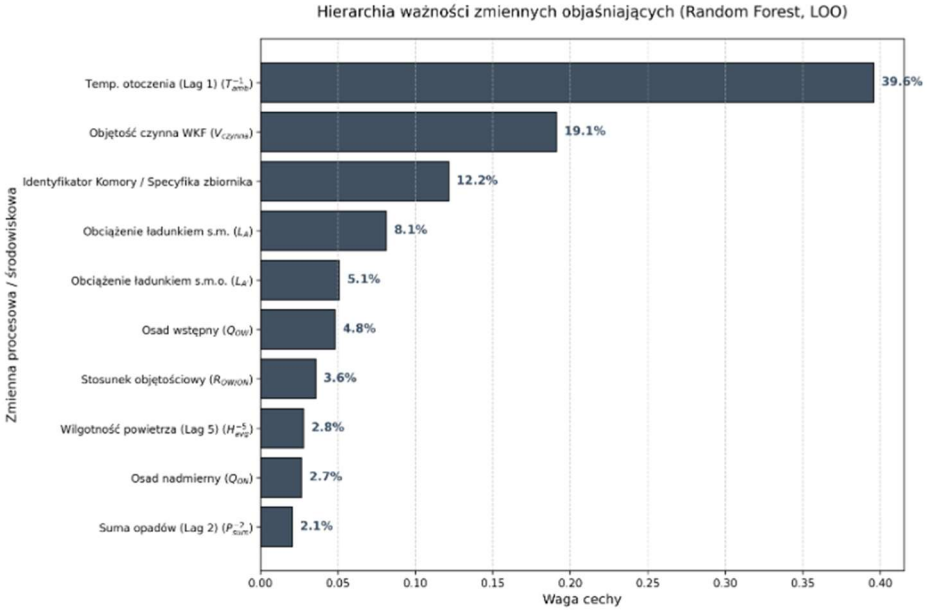
Tabela 4. Analiza długoterminowego wpływu warunków pogodowych

Table 4. Long-term impact of meteorological variables on biogas production

Okres (opóźnienie)	Jednostka	Pearson (r)	Spearman (ρ)
Średnia temp. (Tydzień 1)	°C	-0.559	-0.497
Średnia temp. (Tydzień 2)	°C	-0.511	-0.482
Średnia temp. (Tydzień 3)	°C	-0.496	-0.466
Średnia temp. (Tydzień 4)	°C	-0.497	-0.481
Suma opadów (Tydzień 1)	mm	0.029	0.059
Suma opadów (Tydzień 2)	mm	0.025	0.039
Suma opadów (Tydzień 3)	mm	-0.036	-0.009
Suma opadów (Tydzień 4)	mm	-0.037	-0.010
Średnia wilgotność (Tydzień 1)	%	0.506	0.522
Średnia wilgotność (Tydzień 2)	%	0.451	0.461
Średnia wilgotność (Tydzień 3)	%	0.401	0.421
Średnia wilgotność (Tydzień 4)	%	0.424	0.445

4.2. Wybór modelu predykcyjnego

Algorytm Random Forest wybrano na podstawie wcześniejszych analiz porównawczych przeprowadzonych na tym samym obiekcie [16], gdzie wykazał najwyższą odporność na nieliniowość i szum procesowy. Ranking ważności cech potwierdził dominujący wpływ temperatury (rys. 1.) oraz specyfiki komory fermentacyjnej. Włączenie zmiennych meteorologicznych zwiększyło stabilność modelu w warunkach rzeczywistych.



Rys. 1. Ranking ważności zmiennych objaśniających w modelu Random Forest

Fig. 1. Feature importance ranking in the Random Forest model

Wpływ zmiennych objaśniających na jakość predykcji przedstawiono na rys. 1. Największe znaczenie wykazuje temperatura otoczenia z opóźnieniem jednodniowym $\left(T^{-1}_{amb}\right)$, której udział w modelu wynosi ok. 0.40. Potwierdza to wysoką wrażliwość procesu fermentacji na zmiany warunków termicznych.

Istotną rolę odgrywa również identyfikator komory $\left(ID_{\{komory\}}\right)$, co wskazuje na różnice konstrukcyjne i eksploatacyjne między reaktorami.

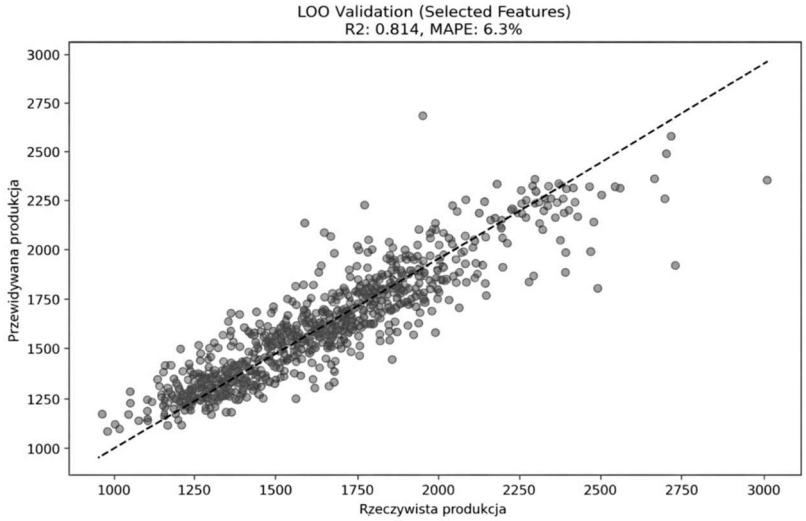
W związku z wcześniejszymi analizami porównawczymi [16] zrezygnowano z ponownej walidacji modeli liniowych, wykorzystując Random Forest jako sprawdzony model predykcyjny. jest rozszerzenie wektora wejściowego o zmienne meteorologiczne oraz ich opóźnienia czasowe, co poprawia stabilność predykcji w warunkach rzeczywistych, co potwierdzają wskaźniki jakości dopasowania (tabela 5.).

Tabela 5. Wyniki walidacji modelu predykcyjnego (Random Forest)

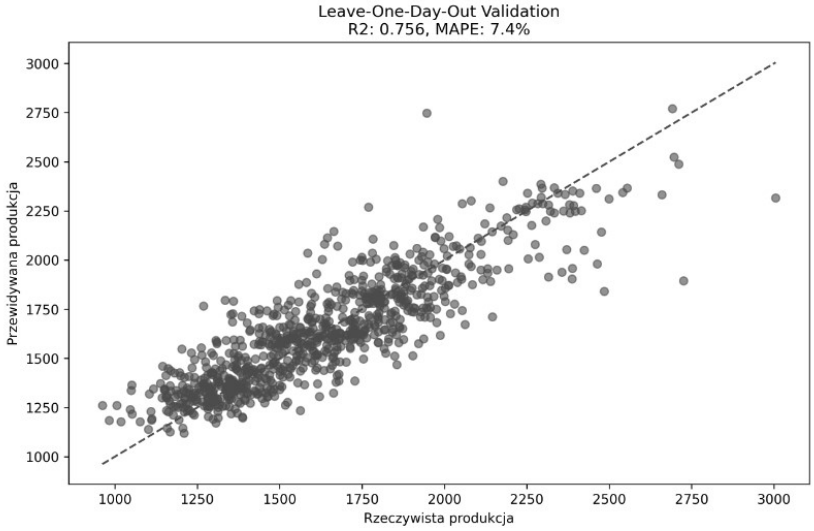
Table 5. Validation results of the Random Forest predictive model

Wskaźnik jakości	**Jednostka**	**Wartość**
Współczynnik determinacji (R^2)	-	0.8403
Błąd średniokwadratowy (RMSE)	m^3/d	127.84
Średni błąd procentowy (MAPE)	%	5.81

Wysoka koncentracja punktów wokół linii idealnego dopasowania na wykresie walidacji LOO (rys. 2.) potwierdza dobrą aproksymację zależności nieliniowych. Rozkład jest symetryczny, co wskazuje na brak systematycznego obciążenia modelu. Niewielkie odchylenia dotyczą jedynie wartości skrajnych, co jest typowe dla danych przemysłowych.



Rys. 2. Wykres dopasowania modelu (Walidacja LOO)
Fig. 2. Model fit plot (LOO validation)



Rys. 3. Wykres dopasowania modelu (Walidacja LODO)
Fig. 3. Model fit plot (LODO validation)

Walidacja LODO zilustrowana na rys. 3. uzyskała wskaźnik $R^2 = 0.756$, podczas gdy dla LOO $R^2 = 0.84$. Różnica wynika z eliminacji korelacji przestrzennej między komorami. Wynik LODO jest bardziej reprezentatywny dla rzeczywistych warunków sterowania, gdzie brak informacji o przyszłych parametrach całego obiektu.

Mimo spadku dokładności, $MAPE = 7.4\%$ potwierdza przydatność modelu w systemach sterowania.

Uzyskane wyniki stanowią punkt wyjścia do dalszych badań nad integracją modeli predykcyjnych z metodami optymalizacji ewolucyjnej, które w literaturze wskazywane są jako obiecujące narzędzie poprawy jakości danych i sterowania procesami środowiskowymi [9].

5. Podsumowanie

W rozdziale zaproponowano model predykcyjny produkcji biogazu oparty na algorytmie Random Forest, uwzględniający parametry wsadu oraz opóźniony wpływ czynników meteorologicznych. Model skutecznie odwzorowuje nieliniową dynamikę procesu fermentacji metanowej, co potwierdzają uzyskane wartości wskaźników jakości dopasowania. Analiza ważności cech wykazała istotny wpływ temperatury otoczenia oraz parametrów eksploatacyjnych komór fermentacyjnych.

Wyniki walidacji wskazują na dobrą zdolność generalizacji modelu, przy czym procedura LODO dostarcza bardziej realistycznej oceny skuteczności niż LOO, eliminując efekt wycieku informacji między równoległymi reaktorami. Potwierdzono, że dezagregacja przestrzenna zwiększa liczebność zbioru uczącego, jednak wymaga zastosowania walidacji grupowej w celu uniknięcia zawyżenia wyników.

Uzyskane rezultaty wskazują na potencjał wykorzystania modelu w systemach sterowania predykcyjnego oczyszczalni ścieków. Dalsze badania powinny obejmować weryfikację metody w warunkach rzeczywistych oraz rozwój mechanizmów adaptacyjnych uwzględniających zmienność składu osadu.

Literatura

[1] Chaoui A., Rebija K., Chkaiti K., et al.: Applications of missing data imputation methods in wastewater treatment plants, *International Journal of Advanced Computer Science and Applications* 14(10) (2023).

[2] Cruz I.A., Chuenchart W., Long F., et al.: Application of machine learning in anaerobic digestion: Perspectives and challenges, *Bioresource Technology* 345, 126433 (2022).

[3] De Mulder C., Flameling T., Weijers S., et al.: An open software package for data reconciliation and gap filling in preparation of water and resource recovery facility modeling, *Environmental Modelling Software* 107, 186–198 (2018).

[4] Drewnowski J., Szeląg B., Sabba F., et al.: Innovations in wastewater treatment: Harnessing mathematical modeling and computer simulations with cutting-edge technologies and advanced control systems, *Journal of Ecological Engineering* 24(12), 208–222 (2023).

[5] Han H., Sun M., Wu X., et al.: Double-cycle weighted imputation method for wastewater treatment process data with multiple missing patterns, *Science China Technological Sciences* 65(12), 2967–2978 (2022).

[6] Hauke J., Kossowski T.: Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data, *Quaestiones Geographicae* 30, 87–93 (2011).

[7] Irati A., Chang N.-B.: Advances in control technologies for wastewater treatment processes: status, challenges, and perspectives, *IEEE/CAA Journal of Automatica Sinica* 6(2), 337–363 (2019).

[8] Kadlec P., Gabrys B., Strandt S.: Data-driven soft sensors in the process industry, *Computers & Chemical Engineering* 33(4), 795–814 (2009).

[9] Kazadi Mbamba C., Keymer P., Alvi M., et al.: Enhancing data quality in wastewater processes: Missing data imputation with deep variational autoencoders and genetic algorithms, *Computers & Chemical Engineering* 199, 109123 (2025).

[10] Maj K., Masłoń A.: Wieloparametryczna ocena efektywności oczyszczalni ścieków w Rzeszowie przed modernizacją, *Czasopismo Inżynierii Lądowej, Środowiska i Architektury* (62, nr 1), 299–315 (2015).

[11] Masłoń A.: Analysis of energy consumption at the Rzeszów wastewater treatment plant, *E3S Web Conf.* 22, 00115 (2017).

[12] Masłoń A., Tendera K.: Gospodarka osadami ściekowymi w oczyszczalni ścieków Rzeszów, *Forum Eksploatatora* 1(88), 38–43 (2017).

[13] Masłoń A., Tomaszek J.A.: Ocena efektywności oczyszczalni ścieków w Lubaczowie, *Czasopismo Inżynierii Lądowej, Środowiska i Architektury* 30(60), 3 (2013).

[14] Nicole A., Korodi A., Silea J.: Weather-based prediction strategy inside the proactive historian with application in wastewater treatment plants, *Applied Sciences* 10(9) (2020).

[15] Oliveira-Esquerre K.P., Seborg D.E., Bruns R.E., et al.: Application of steady-state and dynamic modeling for the prediction of the BOD of an aerated lagoon at a pulp and paper mill: Part I. Linear approaches, *Chemical Engineering Journal* 104(1), 73–81 (2004).

[16] Organiściak P., Masłoń A., Kowal B., et al.: Machine learning-based prediction of biogas production from sludge characteristics in four anaerobic digesters: Development of the ad2biogas prediction tool, *Advances in Science and Technology Research Journal* 18(8), 1–15 (2024).

[17] Organiściak P., Turek A., Stęchły A., et al.: Prediction of biogas yield from anaerobic digesters using meta-regression, *Zeszyty Naukowe SGSP/Szkoła Główna Służby Pożarniczej* (2025).

[18] Renganathan P., Gaysina L.A.: Next-generation wastewater treatment: Omics and AI-driven microbial strategies for xenobiotic bioremediation and circular resource recovery, *Processes* 13(10) (2025).

[19] Szczyrba P., Masłoń A., Czarnota J., et al.: Analiza gospodarki osadowej i biogazowo-energetycznej w oczyszczalni ścieków w Opolu, *Ecological Engineering & Environmental Technology* 21(2), 26–34 (2020).

[20] Zippenfenig P.: Open-Meteo.com weather API, *Zenodo* (2023).

PREDICTION OF BIOGAS PRODUCTION WITH CONSIDERATION OF METEOROLOGICAL FACTORS IN THE ANAEROBIC DIGESTION PROCESS

This study presents a predictive model of biogas production in the anaerobic digestion process, incorporating the influence of meteorological factors. Unlike conventional approaches based solely on substrate characteristics, the proposed framework extends the input feature set to include exogenous variables such as ambient temperature, precipitation sum, and relative humidity, together with their time lags reflecting the inertial nature of biological processes. Due to the limited size of the available operational dataset, a Spatial Disaggregation technique was applied. This approach transforms data from four parallel anaerobic digesters into a panel dataset, effectively increasing the number of training observations and improving the model's generalization capability. Biogas production was estimated using the Random Forest algorithm, which enables accurate modeling of nonlinear relationships between process parameters and methane yield. In a rigorous Leave-One-Day-Out (LODO) validation procedure, the model achieved a coefficient of determination of $R^2 = 0.756$ and a mean absolute percentage error of $MAPE = 7.4\%$, confirming reliable predictive performance under real industrial conditions. The results highlight the significant role of meteorological variables in modeling anaerobic digestion performance and demonstrate the potential applicability of the proposed approach in predictive control systems and decision support frameworks for wastewater treatment plants.

Keywords: energy efficiency, biological process modeling, machine learning, Random Forest, exogenous variables, data disaggregation.

Karolina Kisała, Anna Turek, Martyna Bielenda
Koło Naukowe Systemow Złożonych

ADAPTACJA PLATFORMY LMS MOODLE DLA NIEZALEŻNYCH OŚRODKÓW EDUKACYJNYCH Z WYKORZYSTANIEM ZEWNĘTRZNYCH SYSTEMÓW ZARZĄDZANIA

Streszczenie

Celem rozdziału jest analiza możliwości adaptacji platformy LMS Moodle do obsługi wielu niezależnych podmiotów edukacyjnych (m.in. szkół) w ramach jednej instancji systemowej, ze szczególnym uwzględnieniem problemu izolacji uprawnień administracyjnych. Obiektem badań jest zintegrowane środowisko informatyczne składające się z platformy LMS Moodle oraz autorskiego systemu zarządzania opartego na frameworku Django, komunikujących się za pośrednictwem interfejsu API. W pracy zastosowano metodę studium przypadku (case study) opartego na implementacji technicznej zewnętrznej warstwy logicznej. Przeanalizowano ograniczenia natywnych mechanizmów Moodle w zakresie separacji danych użytkowników pomiędzy różnymi jednostkami organizacyjnymi i zaprojektowano architekturę, w której zewnętrzny system zarządzania przejmuje rolę zarządcy tożsamości. Wyniki wskazują, że wprowadzenie logiki biznesowej (tworzenie kont, zapisy na kursy) do zewnętrznej aplikacji Django pozwala na skuteczną izolację grup użytkowników. Umożliwia to dyrektorom placówek edukacyjnych zarządzanie wyłącznie własnymi kursantami bez wglądu w dane innych podmiotów, co w standardowej konfiguracji Moodle jest utrudnione. Zaprezentowane rozwiązanie zwiększa bezpieczeństwo danych, skalowalność wdrożeń w modelu SaaS i może służyć jako wsparcie dla podobnych wdrożeń.

Słowa kluczowe: Kontrola dostępu, Izolacja danych, Django, Multi-tenancy, API Integration.

1. Wprowadzenie

Dynamiczny rozwój technologii cyfrowych oraz rosnąca dostępność Internetu w istotny sposób zmieniają sposób funkcjonowania współczesnej edukacji. Procesy te obejmują nie tylko szkolnictwo formalne, lecz w coraz większym stopniu dotyczą także edukacji pozaformalnej i nieformalnej, realizowanej przez niezależne ośrodki edukacyjne, organizacje szkoleniowe oraz podmioty sektora EdTech

[7]. W literaturze przedmiotu podkreśla się, że e-learning i *blended learning* przestały być rozwiązaniami pomocniczymi, a stały się trwałym elementem ekosystemu edukacyjnego, wspierającym personalizację kształcenia, elastyczność organizacyjną oraz dostęp do wiedzy w modelu uczenia się przez całe życie [10].

Jednym z najczęściej wykorzystywanych rozwiązań technologicznych w tym obszarze jest platforma Moodle - otwartoźródłowy system zarządzania nauczaniem (LMS) [2], który dzięki swojej modułowej architekturze i szerokim możliwościom konfiguracji znajduje zastosowanie zarówno w szkolnictwie wyższym, jak i w instytucjach prowadzących edukację pozaformalną. Badania wskazują, że platformy LMS pełnią dziś podwójną rolę: z jednej strony wspierają proces dydaktyczny, z drugiej - stają się narzędziem zarządzania, administracji oraz monitorowania jakości kształcenia.

Niezależne ośrodki edukacyjne coraz częściej funkcjonują w złożonym środowisku technologicznym, w którym sama platforma LMS nie jest wystarczająca do obsługi wszystkich procesów organizacyjnych. W praktyce rośnie znaczenie architektur opartych na integracji systemów, gdzie Moodle pełni rolę platformy dydaktycznej, natomiast systemy zewnętrzne realizują funkcje zarządzania użytkownikami, zapisami na kursy, obsługą procesów administracyjnych oraz raportowaniem.

Tego typu podejście umożliwia automatyzację procesów, takich jak rejestracja uczestników szkoleń, synchronizacja danych użytkowników pomiędzy systemami oraz generowanie raportów dotyczących postępów i aktywności uczestników. Integracja poprzez interfejsy API pozwala systemowi nadrzędnemu na tworzenie kont w Moodle, przypisywanie użytkowników do kursów oraz zarządzanie procesem edukacyjnym.

Według raportów branżowych, technologie edukacyjne rozwijają się w tempie kilkunastu-kilkudziesięciu procent rocznie, a cyfryzacja procesów edukacyjnych stanowi jeden z czynników zwiększających efektywność organizacyjną instytucji szkoleniowych [6]. Tworzenie ekosystemów informatycznych opartych na współpracujących platformach edukacyjnych umożliwia automatyzację procesów administracyjnych, skalowanie działalności szkoleniowej oraz centralne zarządzanie danymi uczestników [4].

Celem niniejszego rozdziału jest analiza możliwości adaptacji platformy LMS Moodle do zarządzania wieloetapowym środowiskiem edukacyjnym, ze szczególnym uwzględnieniem integracji z zewnętrznym systemem zarządzania. Analiza obejmuje wykorzystanie mechanizmów integracyjnych umożliwiających automatyzację procesów tworzenia kont użytkowników, zarządzania zapisami na kursy oraz agregacji danych raportowych. Rozważania te stanowią element szerszej analizy roli nowoczesnych technologii edukacyjnych w cyfryzacji procesów administracyjnych oraz budowie zintegrowanych środowisk e-learningowych umożliwiających skalowalne zarządzanie ofertą szkoleniową.

2. Charakterystyka platformy LMS Moodle i rola zewnętrznego systemu zarządzania

Learning Management System (LMS) to system informatyczny służący do zarządzania procesem kształcenia realizowanym z wykorzystaniem technologii cyfrowych. Umożliwia on tworzenie, udostępnianie i organizowanie treści edukacyjnych w formie kursów online. LMS wspiera administrację procesów dydaktycznych, takich jak zapisy uczestników, harmonogramy zajęć czy kontrola dostępu do materiałów. System ten pozwala także na monitorowanie postępów uczących się, ocenianie ich aktywności oraz analizę wyników nauczania. LMS zapewnia narzędzia komunikacji między uczestnikami a prowadzącymi, m.in. czaty i wiadomości. Może być wykorzystywany zarówno w edukacji formalnej, pozaformalnej, jak i w szkoleniach biznesowych. Wiele systemów LMS umożliwia integrację z innymi narzędziami i systemami informatycznymi. Dzięki temu LMS stanowi centralny element cyfrowego środowiska uczenia się [1].

2.1. Charakterystyka platformy Moodle

Platforma Moodle oferuje rozbudowany mechanizm zarządzania użytkownikami, obejmujący rejestrację, zapisy do kursów oraz przypisywanie ról. System wykorzystuje elastyczny model ról i uprawnień, który pozwala precyzyjnie określać zakres działań poszczególnych użytkowników, takich jak administratorzy, nauczyciele, studenci czy goście. Role mogą być modyfikowane lub definiowane od podstaw, co umożliwia dopasowanie struktury uprawnień do specyfiki danej organizacji [9].

Moodle posiada wbudowany dziennik ocen umożliwiający gromadzenie i przetwarzanie wyników uzyskiwanych przez uczestników kursów. Prowadzący mogą stosować różne metody oceniania, w tym ocenianie punktowe, skale opisowe oraz rubryki. System automatycznie rejestruje aktywność użytkowników, co pozwala generować raporty dotyczące postępów w nauce, frekwencji oraz zaangażowania uczestników. Dostępne raporty wspierają analizę procesu kształcenia oraz podejmowanie decyzji dydaktycznych.

Moodle stanowi atrakcyjne rozwiązanie dla małych i średnich organizacji edukacyjnych, szkoleniowych oraz instytucji prowadzących działania rozwojowe, głównie ze względu na swoją elastyczność, dostępność oraz brak kosztów licencyjnych wynikający z modelu open source. Pozwala to przeznaczyć zasoby finansowe na rozwój treści dydaktycznych oraz infrastruktury technicznej zamiast na zakup komercyjnych systemów zamkniętych.

Istotną zaletą platformy jest możliwość skalowania wraz z rozwojem organizacji. System sprawdza się zarówno w niewielkich zespołach szkoleniowych, jak i w instytucjach obsługujących dużą liczbę uczestników. Modularna architektura umożliwia wdrażanie wyłącznie tych funkcjonalności, które są faktycznie potrzebne, co upraszcza administrację i ogranicza złożoność środowiska.

Dodatkowym atutem Moodle jest rozbudowana społeczność użytkowników oraz dostępność dokumentacji, forów dyskusyjnych i rozszerzeń. Ułatwia to rozwój platformy oraz rozwiązywanie problemów bez konieczności korzystania z kosztownych usług zewnętrznych. System wspiera również różne modele kształcenia, w tym e-learning oraz *blended learning*, co umożliwia dostosowanie oferty edukacyjnej do potrzeb odbiorców.

2.2. LMS w kontekście zarządzania niezależnymi podmiotami naukowymi

Istotnym wyzwaniem w środowiskach współdzielonych jest zapewnienie prywatności danych pomiędzy niezależnymi instytucjami korzystającymi z tej samej infrastruktury. Wiele niezależnych szkół korzysta z jednej wspólnej platformy Moodle, ale ich dane muszą być od siebie całkowicie odseparowane. Nauczyciele nie mogą widzieć kursów ani uczniów nie swojej placówki.

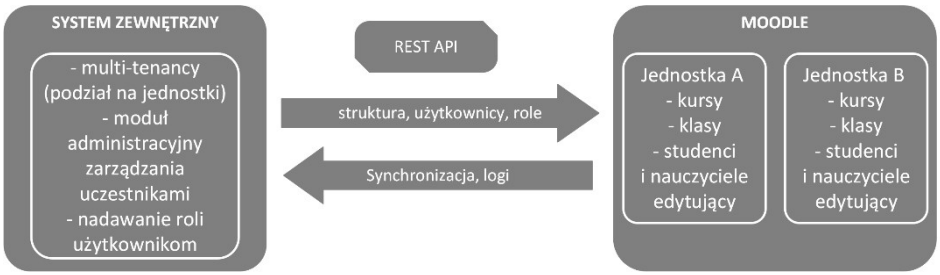
Separacja danych w przedstawionym rozwiązaniu realizowana jest na poziomie zewnętrznego systemu zarządzania, gdzie wprowadzono podział na instytucje edukacyjne jako podstawowe jednostki organizacyjne. Każdy użytkownik – dyrektor, nauczyciel, uczeń jest przypisywany do konkretnej jednostki, do której ma dostęp. Role determinują uprawnienia użytkowników. Do Moodle przesyłane są ich dane oraz informacje dotyczące struktury jednostki edukacyjnej, co minimalizuje ryzyko przenikania danych między instytucjami. Dodatkowo, istnieje separacja po stronie Moodle poprzez odpowiednią organizację kursów z podziałem na kategorie oraz z możliwością wykorzystania grup, które pozwalają dzielić uczestników na klasy zachowując porządek i izolację danych.

Kluczowym elementem działania systemu jest wielodostępowość zapewniona poprzez logiczny podział użytkowników i kursów na odrębne jednostki dydaktyczne, funkcjonujące w ramach jednej infrastruktury. Przynależność do instytucji determinuje zakres widocznych danych, dostępnych kursów oraz możliwych działań w systemie. Każda szkoła posiada własną, izolowaną przestrzeń organizacyjną, przy zachowaniu współdzielonego środowiska technicznego.

Istotnym elementem zapewniającym separację i bezpieczeństwo jest kontrola dostępu oparta o przypisywanie ról. Każda rola w systemie zawiera określony pakiet uprawnień, obejmujący dostęp do kursu, możliwość edycji treści, czy udział w aktywnościach. W omawianym rozwiązaniu role takie jak uczeń czy nauczyciel nie są nadawane bezpośrednio w Moodle, lecz są definiowane i zarządzane centralnie w zewnętrznym systemie i synchronizowane z Moodle. Takie rozwiązanie eliminuje potrzebę ręcznej synchronizacji danych w Moodle oraz zapobiega powstawaniu niespójności takich jak różnica ról pomiędzy systemami. Dodatkowo, centralne zarządzanie uprawnieniami zwiększa elastyczność systemu poprzez możliwość masowej zmiany uprawnień dla grupy użytkowników oraz zwiększa przejrzystość struktury organizacyjnej.

Zarządzanie niezależnymi podmiotami edukacyjnymi w ramach jednej platformy LMS wymaga połączenia platformy Moodle z logiczną separacją danych

i uprawnień na poziomie zewnętrznego systemu. W przedstawionym rozwiązaniu rolę nadrzędną pełni Django, który odwzorowuje strukturę organizacyjną szkoły oraz centralnie zarządza użytkownikami i rolami. Moodle funkcjonuje jako system wykonawczy, odzwierciedlający strukturę warstwy nadrzędnej. Zastosowany model multi-tenancy umożliwia obsługę wielu jednostek dydaktycznych w ramach jednego systemu, przy jednoczesnym zachowaniu pełnej izolacji danych. Separacja realizowana jest poprzez przypisanie użytkowników do konkretnych instytucji oraz kontrolę dostępu opartą na rolach zsynchronizowanych z zewnętrznym systemem zarządzania. Dzięki temu nauczyciele i uczniowie mają dostęp wyłącznie do zasobów własnej jednostki organizacyjnej, co zapewnia prywatność oraz bezpieczeństwo informacji. Opisywany model został przedstawiony na poniższym rysunku (rys. 1.).



Rys. 1. Schemat architektury logicznej separacji danych
Fig. 1. Logical data separation architecture diagram

Przyjęta architektura łączy utrzymanie wysokiej kontroli z przejrzystością zarządzania. Pozwala na łatwe skalowanie, zapewnia utrzymanie spójności danych oraz spełnienie założeń związanych z izolacją danych.

2.3. Spotkania online w kontekście dydaktyki i napotkane ograniczenia

Istotnym elementem aktualnej dydaktyki i samych kursów e-learningowych są spotkania zdalne. Moodle natywnie nie pozwala na taką formę zajęć, ale wspiera narzędzia zewnętrzne, takie jak BigBlueButton [8]. Jest to system do wideokonferencji z możliwością pełnej integracji z Moodle. System Moodle wymaga dostępu do serwera oraz zainstalowania i skonfigurowania powiązanej wtyczki, aby w pełni korzystać z możliwości zajęć zdalnych.

Wtyczka BigBlueButton umożliwia tworzenie spotkań online w obrębie danego kursu wraz z możliwością nagrywania. System sprzyja kontroli uczestnictwa przez rejestrowanie obecności i aktywności użytkowników. Po uruchomieniu widoczna jest opcja pobrania listy uczestników oraz po zakończeniu wyświetlany jest pulpit analizy spotkania, zawierający szczegółowy raport aktywności. Ponadto, rozwiązanie jest odpowiednie do organizowania konferencji. Domyślnie tylko moderator ma dostęp do wirtualnej tablicy, ale po wyznaczeniu prelegenta opcja jest

rozszerzana o danego uczestnika. Zapewniona jest wystarczająca kontrola dostępu, co wpływa na bezpieczeństwo spotkań.

W kontekście integracji z systemami zewnętrznymi, Moodle nie udostępnia natywnie funkcji, która pozwala w pełni tworzyć i konfigurować aktywność BBB poprzez API. Analiza przypadku wykazała, że sposobem rozwiązania tego problemu jest stosowanie zewnętrznych wtyczek lub kopiowanie całych kursów, w których są już skonfigurowane spotkania BBB.

W omawianej implementacji zastosowano wtyczkę BBB bez integracji z API systemu Moodle. Zarządzanie kursem zostało częściowo podzielone na dwa systemy - spotkania online są wyłącznie konfigurowane w Moodle.

2.4. Rola zewnętrznego systemu zarządzania w architekturze rozwiązania

Zewnętrzny system zarządzania ułatwia centralizację procesów administracyjnych poza LMS, co umożliwia zachowanie spójności danych oraz logiczną izolację środowisk dydaktycznych. Wszystkie operacje związane z zarządzaniem, takie jak tworzenie struktur organizacyjnych, nadawanie ról użytkownikom, przypisywanie ich do odpowiednich instytucji oraz edycja kursów realizowane są w jednym, zewnętrznym systemie. Podejście pozwala na rozdzielenie środowisk poszczególnych jednostek edukacyjnych w ramach jednej infrastruktury LMS.

Takie rozwiązania okazują się skuteczne w środowisku zintegrowanym z zewnętrznym systemem zarządzania opartym na frameworku Django funkcjonalności Moodle związane z zarządzaniem użytkownikami mogą być realizowane w sposób scentralizowany [3], gdzie system nadrzędny odpowiada za proces rejestracji użytkownika, zarządzanie jego danymi oraz inicjowanie operacji zapisu do kursów poprzez wykorzystanie interfejsów integracyjnych Moodle.

3. Analiza strategii mapowania struktury organizacyjnej

Podczas projektowania architektury integracyjnej przeanalizowano dwa warianty odwzorowania struktur klasowych w systemie Moodle. Każda strategia wpływa odmiennie na sposób działania interfejsu API odpowiedzialnego za synchronizację danych, jak i organizację pracy wewnątrz platformy przez osoby docelowe. Wybór odpowiedniego modelu determinuje strukturę kursów w Moodle, sposób izolacji danych pomiędzy jednostkami i podjednostkami wewnętrznymi oraz poziom redundancji samych materiałów e-learningowych.

3.1. Model atomizacji kursów (jeden kurs - jedna klasa)

Model zakłada, że każda klasa funkcjonuje jako odrębny, niezależny kurs w systemie Moodle. Oznacza to, że nawet jeśli kilka klas w szkole realizuje dokładnie ten sam program, tworzona jest osobna replika materiałów dydaktycznych. Zapewnia pełną separowalność poprzez tworzenie dla każdej klasy osobnej przestrzeni kursowej.

W praktyce oznacza to, że każdy kurs posiada odseparowanych użytkowników oraz oddzielne zasoby. Uczniowie przypisani do jednej klasy nie mają możliwości zobaczenia uczestników innej klasy lub szkoły, ponieważ funkcjonują na odrębnych strukturach. Dyrektorzy widzą jedynie panel zarządzania z poziomu aplikacji zewnętrznej (nauczyciele i uczniowie), a nauczyciele wyłącznie uczniów przydzielonych do kursu. Zestawienie wad i zalet modelu atomizacji kursów przedstawiono w poniższej analizie (tabela 1.).

Tabela 1. Analiza zalet i wad modelu atomizacji kursów

Table 1. Analysis of the advantages and disadvantages of the rate atomization model

Zalety	Wady
Pełna, naturalna izolacja danych i użytkowników bez konieczności konfiguracji dodatkowych restrykcji wewnątrz kursu.	Wysoka redundancja treści dydaktycznych. Nauczyciel lub trener prowadzący ten sam kurs w kilku grupach zmuszony jest do powielania zasobów i aktywności, co utrudnia zarządzanie materiałem i aktualizację treści.

3.2. Model rocznikowy (jeden kurs – wiele wewnętrznych jednostek)

Model rocznikowy z separacją grupową zakłada utworzenie jednego kursu na dany poziom edukacyjny i separację klas poprzez podział na grupy. Nauczyciel pracuje na jednej wersji kursu odseparowanej pomiędzy klasami. Podział na klasy odbywa się na poziomie Moodle poprzez opcję przypisywania do grup. Zapewniony jest rozdział pomiędzy oddziałami szkolnymi, obejmujący dostosowanie dostępu do kursu w obrębie utworzonych grup. Jednocześnie każdy trener lub nauczyciel prowadzący zajęcia w obrębie jednego oddziału (na przykład rocznika klasy w szkole) ma możliwość podglądu uczestników przypisanych do tego kursu. Tabela 2. przedstawia zalety i wady rozwiązania z użyciem modelu rocznikowego z separacją grup.

Tabela 2. Analiza zalet i wad modelu rocznikowego

Table 2. Analysis of the advantages and disadvantages of the year model

Zalety	Wady
Centralizacja zasobów edukacyjnych – nauczyciel lub trener danej jednostki edukacyjnej zarządza jedną wersją materiałów dla wszystkich klas danego roku	Wymaga precyzyjnego zarządzania przynależnością do grup z poziomu zewnętrznego systemu. API musi nie tylko zapisać użytkownika do kursu, ale jednocześnie przypisać go do właściwej grupy, aby zapewnić, że kursanci z różnych oddziałów (na przykład szkół lub uczelni) nie wchodzili ze sobą w interakcje

3.3. Dopasowanie modelu do badanego problemu

W opisywanym rozwiązaniu zastosowano model atomizacji, przenosząc ciężar zarządzania polityką dostępu na warstwę logiczną aplikacji zarządzającej.

Pozwala to dyrektorom szkół na wirtualne "zarządzanie klasami", podczas gdy system w tle tłumaczy te operacje na przypisania do odpowiednich grup w Moodle, zachowując separację danych między uczniami.

4. Architektura systemu i integracja wielopodmiotowa

4.1. Działanie zintegrowanego systemu Moodle-Django

W przyjętym modelu aplikacja Django pełni rolę zewnętrznego systemu zarządzania, w której definiowana jest struktura szkoły oraz przypisywane są role użytkowników, takie jak student czy nauczyciel. Po dodaniu użytkownika do odpowiedniej jednostki organizacyjnej system Django wykonuje synchronizację z platformą Moodle, tworząc konto użytkownika (jeśli nie istnieje) oraz zapisując go do wskazanego kursu z właściwą rolą.

Moodle nadal wykorzystuje własny system ról i uprawnień, jednak ich tworzenie i przypisywanie może być inicjowane z poziomu systemu zewnętrznego. Pozwala to na utrzymanie centralnego zarządzania użytkownikami przy jednoczesnym zachowaniu pełnej funkcjonalności LMS w zakresie realizacji procesu dydaktycznego.

Dane dotyczące aktywności użytkowników, wyników oraz postępów w nauce mogą być synchronizowane z systemem zewnętrznym w celu generowania raportów administracyjnych i analitycznych. Informacje udostępniane poprzez interfejs API Moodle są agregowane w systemie nadrzędnym, co umożliwia tworzenie zestawień dotyczących efektywności kształcenia, frekwencji oraz zaangażowania uczestników.

Rozdzielenie warstwy dydaktycznej i administracyjnej pozwala zwiększyć przejrzystość procesów organizacyjnych oraz zautomatyzować wiele czynności, takich jak rejestracja użytkowników, zapisy na kursy czy generowanie raportów. Integracja Moodle z systemem Django sprzyja budowie skalowalnego środowiska edukacyjnego, w którym LMS realizuje funkcje dydaktyczne, natomiast zewnętrzny system zarządzaina obsługuje procesy zarządzania i organizacji szkoleń.

Rysunek 2. (schemat sekwencji) obrazuje przebieg integracji pomiędzy systemem zarządzania a platformą LMS. Przedstawia komunikację pomiędzy administratorem, aplikacją Django, interfejsem Moodle Web Services API oraz systemem Moodle. Proces rozpoczyna się od działania użytkownika administracyjnego, który w systemie Django tworzy użytkownika oraz przypisuje mu odpowiednią rolę w strukturze szkoły, np. student lub nauczyciel. Następnie aplikacja Django, pełniąca funkcję zewnętrznego systemu zarządzania, inicjuje komunikację z platformą Moodle poprzez wywołanie odpowiednich usług API. W przypadku braku konta użytkownika w Moodle wykonywana jest operacja jego utworzenia, po czym system realizuje zapis uczestnika do wybranego kursu wraz z przypisaniem właściwej roli. Moodle zwraca informację o powodzeniu operacji, która jest reje-

strowana w systemie Django w celu zachowania spójności danych pomiędzy systemami. Dzięki takiemu mechanizmowi możliwe jest centralne zarządzanie strukturą użytkowników i dostępem do kursów z poziomu aplikacji zarządzającej, przy jednoczesnym wykorzystaniu Moodle jako środowiska realizacji procesu dydaktycznego i rejestrowania aktywności uczestników.



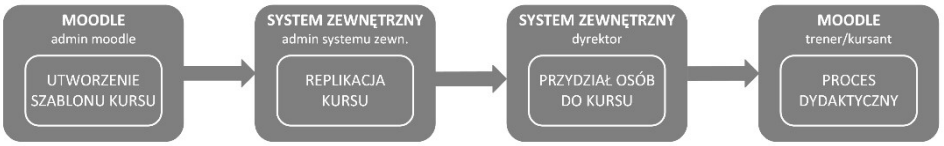
Rys. 2. Schemat sekwencji

Fig. 2. Sequence diagram

4.2. Architektura systemu

Architektura analizowanego rozwiązania opiera się na integracji platformy LMS Moodle z zewnętrznym systemem zarządzania zrealizowanym w technologii Django. Przyjęty model zakłada rozdzielenie warstwy dydaktycznej i administracyjnej, w którym Moodle odpowiada za realizację procesu kształcenia, udostępnianie treści oraz rejestrowanie aktywności uczestników, natomiast aplikacja Django pełni funkcję zewnętrznego systemu zarządzania strukturą organizacyjną, użytkownikami oraz procesem przypisywania ról. W systemie zarządzania definiowane są jednostki organizacyjne, takie jak szkoły, do których przypisywani są użytkownicy z rolami student oraz nauczyciel, a następnie synchronizowani z platformą Moodle poprzez interfejsy integracyjne. Takie podejście umożliwia centralne zarządzanie dostępem do kursów oraz zachowanie spójności danych pomiędzy systemami, jednocześnie wykorzystując Moodle jako środowisko realizacji działań dydaktycznych.

Działanie systemu opiera się na replikacji wzorcowego kursu utworzonego w środowisku Moodle. Administrator Moodle tworzy szablon kursu, definiując jego strukturę i materiały dydaktyczne. Następnie kurs stanowi podstawę do tworzenia kolejnych poprzez kopiowanie i przypisywanie do odpowiedniej jednostki, co wykonuje administrator systemu zewnętrznego. W kolejnym kroku dyrektor instytucji przydziela osoby do kursu (kursantów oraz trenerów) z użyciem zewnętrznego systemu zarządzania. Po zakończeniu etapu organizacyjnego, właściwy proces dydaktyczny odbywa się za pośrednictwem platformy Moodle. Całość procesu prezentuje schemat przedstawiony na rys. 3.



Rys. 3. Sekwencja procesu systemu wraz z wykonawcą

Fig. 3. System process sequence with contractor

Zewnętrzny system zarządzania odpowiada za zarządzanie strukturą, a Moodle stanowi przestrzeń do realizacji treści dydaktycznych. Zaprezentowane podejście replikacji automatyzuje proces dodawania kursów do jednostek edukacyjnych, eliminując potrzebę ręcznej konfiguracji każdego kursu w Moodle. Rozwiązanie minimalizuje ryzyko powstania błędów administracyjnych powstałych podczas dodawania nowych instancji. Z kolei, rozdzielenie zarządzania do systemu zewnętrznego wraz z synchronizacją danych zapewnia spójność oraz bezpieczeństwo.

5. Implementacja mechanizmów synchronizacji

Implementacja stanowi kluczowy etap realizacji zaprojektowanej architektury systemu. Celem projektu było stworzenie bezpiecznego mechanizmu wymiany danych pomiędzy zewnętrznym systemem zarządzania a Moodle, przy zachowaniu spójności struktury organizacyjnej oraz kontroli dostępu użytkowników. Podejście opiera się na centralnym zarządzaniu konfiguracją oraz odwzorowaniu struktur systemowych w Moodle, co umożliwia automatyzację procesów administracyjnych.

5.1. Mapowanie struktur z użyciem klas

W celu zapewnienia spójności pomiędzy systemami zaimplementowano zestaw klas odpowiedzialnych za mapowanie struktur organizacyjnych, obsługę kursów, konfigurację integracji oraz monitorowanie przebiegu synchronizacji. Dzięki takiemu podejściu architektura zachowuje modularność, a poszczególne elementy mogą być niezależnie rozwijane lub modyfikowane. Zestawienie tych elementów wraz z opisem ich funkcji w systemie przedstawiono w Tabeli 3, która porządkuje zaimplementowane klasy oraz wskazuje ich rolę w procesie integracji pomiędzy systemami.

Tabela 3. Opis zaimplementowanych klas

Table 3. Description of implemented classes

Model systemowy	Nazwa klasy	Opis
model klasy	moodleCategory	Odpowiada za mapowanie klasy z zewnętrznego systemu zarządzania na kategorię w Moodle, zapewnia przestrzeń organizacyjną dla każdej klasy z zachowaniem separacji danych.
model kursu	moodleCourse	Odwzorowuje konkretny kurs w Moodle oraz zapewnia jego powiązanie z odpowiednią jednostką organizacyjną w zewnętrznym systemem zarządzania, gwarantuje spójność danych.
szablon kursu	courseTemplate	Wskazuje i odwzorowuje kurs wzorcowy w Moodle, który służy do repliki kursów.

Tabela 3 (cd.). Opis zaimplementowanych klas

Table 3 (cont.). Description of implemented classes

Model systemowy	Nazwa klasy	Opis
integracja Moodle z zewnętrznym systemem zarządzania	moodleToken	Odpowiada za konfigurację i zabezpieczenie komunikacji pomiędzy aplikacją Django a platformą Moodle poprzez interfejs REST API.
ustawienia	settings	Konfiguracja pomiędzy zewnętrznym systemem zarządzania a Moodle.
zapis logów	moodleSyncLog	Rejestruje przebieg synchronizacji pomiędzy zewnętrznym systemem zarządzania a LMS.

5.2. Asynchroniczne przetwarzanie operacji

Ze względu na ograniczenia wydajnościowe interfejsów sieciowych oraz czasochłonność operacji takich jak masowe duplikowanie kursów czy rejestracja całych grup użytkowników, w warstwie implementacyjnej zastosowano mechanizm asynchronicznego kolejkowania zadań z wykorzystaniem biblioteki Celery. Pakiet pozwala to na zlecanie długotrwałych procesów integracyjnych w tle, co zapobiega blokowaniu interfejsu użytkownika w systemie nadrzędnym i gwarantuje wysoką responsywność aplikacji. Dodatkowo wdrożono dedykowany model logowania zdarzeń, który archiwizuje przebieg komunikacji, umożliwiając administratorom monitorowanie poprawności synchronizacji.

5.3. Komunikacja za pośrednictwem Moodle Web Services API

Funkcje dostępne w platformie Moodle stanowią podstawę realizacji procesów administracyjnych oraz dydaktycznych w systemie e-learningowym. Odpowiadają one za wykonywanie konkretnych operacji, związanych z zarządzaniem strukturą kursów, użytkownikami, rolami oraz mechanizmami zapisu na kursy. Każdej funkcji przypisane są konkretne uprawnienia systemowe Moodle.

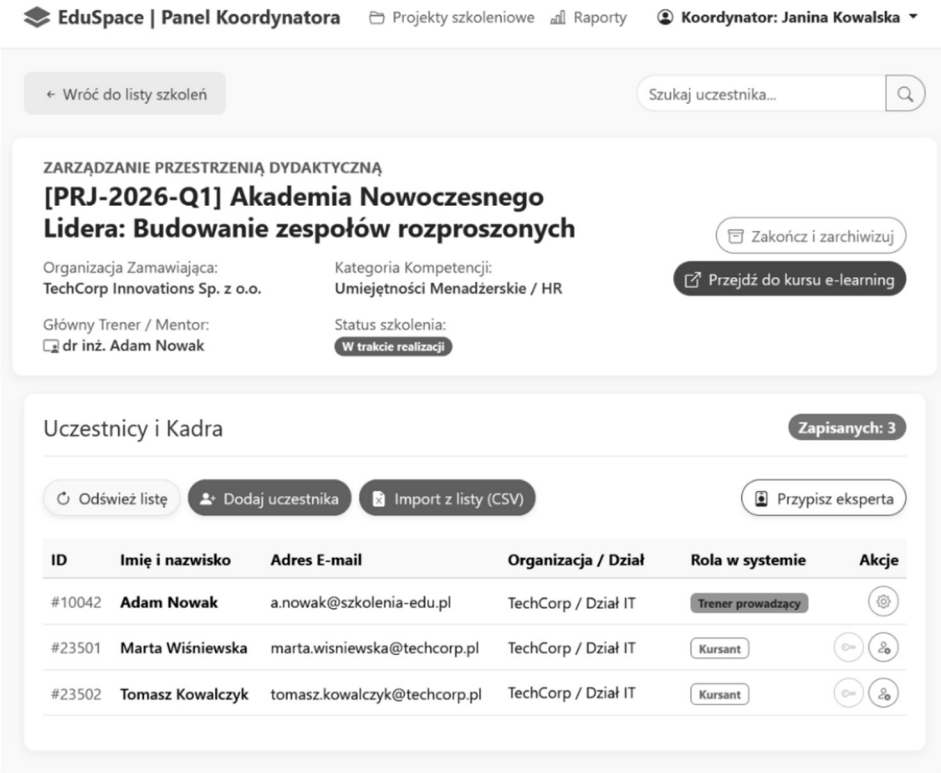
Funkcje są udostępniane w postaci usług sieciowych (Web Services), co umożliwia ich wywołanie przez API. Dzięki temu Moodle może realizować określone działania automatycznie, zgodnie z logiką systemu nadrzędnego, przy jednoczesnym zachowaniu kontroli nad uprawnieniami. Zewnętrzny system zarządzania może wywoływać jedynie te funkcje, które zostały dopisane w panelu administracyjnym moodle, co wpływa na zwiększenie kontroli i bezpieczeństwa. Moodle zawsze odzwierciedla strukturę z systemu nadrzędnego, co zapewnia spójność danych. Każda zmiana powoduje wywołanie odpowiedniej funkcji w Web Service w Moodle. W efekcie struktura Moodle jest zawsze synchronizowana z aktualnym stanem.

Każda funkcja API posiada określone uprawnienia dostępu Moodle. Oznacza to, że do przeprowadzenia operacji konieczne jest posiadanie odpowiednich uprawnień. Taki model dostępu zwiększa bezpieczeństwo środowiska i ogranicza możliwość wprowadzenia nieautoryzowanych zmian.

Integracja została zaprojektowana w sposób umożliwiający skalowalność oraz łatwe utrzymanie systemu. Logika związana z tworzeniem i modyfikacją struktury organizacyjnej znajduje się wyłącznie po stronie zewnętrznego systemu zarządzania, co eliminuje konieczność wprowadzania zmian bezpośrednio w Moodle.

5.4. Interfejs zewnętrznego systemu zarządzania

Interfejs aplikacji zarządzającej (rys. 4.) umożliwia koordynatorom centralne administrowanie procesem dydaktycznym. Z poziomu jednego widoku możliwe jest zarządzanie cyklem życia szkolenia, w tym przypisywanie ekspertów oraz masowy import kursantów. Wszelkie operacje wykonywane w tym panelu są automatycznie, w czasie rzeczywistym, tłumaczone na żądania API i odzwierciedlane w strukturze platformy Moodle, bez konieczności bezpośredniego logowania się administratorów do systemu LMS.



Rys. 4. Ekran aplikacji zarządzającej ze sztucznymi danymi dopasowanymi tematycznie do działalności IT

Fig. 4. Management application screen with artificial data thematically tailored to IT activities

6. Podsumowanie

Celem przeprowadzonych badań była analiza i praktyczna weryfikacja możliwości adaptacji platformy LMS Moodle do obsługi wielu niezależnych podmiotów edukacyjnych w architekturze współdzielonej. Wyniki prac wdrożeniowych potwierdzają, że przeniesienie logiki biznesowej oraz zarządzania tożsamością do zewnętrznej aplikacji nadrzędnej (zrealizowanej w technologii Django) pozwala na skuteczną, logiczną izolację środowisk dydaktycznych w obrębie pojedynczej instancji Moodle.

Wdrożenie architektury opartej na prezentowanym modelu z centralnym systemem zarządzającym przynosi znaczne korzyści, w tym: znaczące zwiększenie bezpieczeństwa i prywatności danych, możliwość łatwego skalowania infrastruktury (szczególnie istotne w modelu SaaS) oraz daleko idącą automatyzację procesów administracyjnych. Dzięki komunikacji przez interfejs API wyeliminowano konieczność ręcznej obsługi struktury kursów i uprawnień użytkowników.

Podczas realizacji projektu zidentyfikowano również obszary wymagające dalszych prac. Należą do nich przede wszystkim wyzwania związane z integracją narzędzi zewnętrznych, takich jak moduł wideokonferencji BigBlueButton, który w standardowej konfiguracji Moodle API nie oferuje pełnej elastyczności przy zarządzaniu spotkaniami z poziomu systemu nadrzędnego. Dalsze kierunki badań powinny koncentrować się na opracowaniu alternatywnych architektur komunikacyjnych rozwiązujących ten problem, szczególnie w obszarze zachowania standardów bezpieczeństwa i prywatności danych przy wymianie informacji między systemami [5] jak również na optymalizacji procesów asynchronicznej wymiany dużych wolumenów danych pomiędzy zintegrowanymi systemami. Warto też przeprowadzić analizę wzrostu kosztów związanych z utrzymaniem i skalowaniem rozszerzonej o system nadzorczy infrastruktury.

Literatura

[1] Berking P., Gallagher S.: Choosing a Learning Management System Advanced Distributed Learning (ADL) Initiative.

[2] Colman H.: LMS (learning management system): Wszystko, co trzeba wiedzieć, https://www.ispring.pl/blog/co-to-jest-lms (dostęp: 20.02.2026).

[3] Dupate A., Dicholkar M., Dsouza R., Sangit P.: "K.C Moodle", International Journal of Creative Research Thoughts (IJCRT), Volume 13, Issue 4, pp. h955–h958, April 2025.

[4] Ninoriya S., Chawan P. M., Meshram B. B.: CMS, LMS and LCMS for elearning. International Journal of Computer Science Issues (IJCSI), 8(2), 644.

[5] Samuel N., Oyenuga A. O.: Education Systems Interoperability: Implications for Privacy and Security in Educational Management Information Systems. University of Ibadan Journal of Science and Logics in ICT Research, 13(1), 153–163.

[6] Timchenko V. V., Trapitsin S. Y., Apevalova Z. V.: Educational technology market analysis. In 2020 International Conference Quality Management, Transport and Information Security (pp. 612–617). IEEE, DOI:10.1109/ITQMIS51053.2020.9322982.

[7] Zielonka M.: Sektor EdTech w Polsce, RAPORT EKSPERCKI.
[8] Dokumentacja BigBlueButton: https://docs.bigbluebutton.org/ (dostęp: 20.02.2026).
[9] Dokumentacja Moodle: https://docs.moodle.org (dostęp: 20.02.2026).
[10] PCG Akademia, Blended learning – co to i dlaczego zyskuje na popularności? https://pcgacademia.pl/blended-learning/ (dostęp: 17.01.2026).

ADAPTATION OF THE LMS MOODLE PLATFORM FOR INDEPENDENT EDUCATIONAL INSTITUTIONS USING EXTERNAL MANAGEMENT SYSTEMS

Summary

The aim of this article is to analyze the possibilities of adapting the LMS Moodle platform to support multiple independent educational entities (including schools) within a single system instance, with particular emphasis on the problem of administrative permission isolation.

The subject of the study is an integrated IT environment consisting of the LMS Moodle platform and a proprietary management system based on the Django framework, communicating via an API interface. The research applies a case study method based on the technical implementation of an external logical layer.

The limitations of Moodle's native mechanisms in terms of user data separation between different organizational units were analyzed, and an architecture was designed in which the external system assumes the role of identity manager. The results indicate that transferring business logic (account creation, course enrollment) to an external Django application enables effective isolation of user groups.

This approach allows school principals to manage only their own students without access to data belonging to other entities, which is difficult to achieve in the standard Moodle configuration. The proposed solution enhances data security and scalability in a SaaS deployment model and may serve as support for similar implementations.

Keywords: Access control, Data isolation, Django, Multi-tenancy, API Integration.

Kajetan Żarski, Mirosław Sobaszek
Koło Podstaw Elektroniki

MODEL SYMULACYJNY GŁOŚNIKA W OBUDOWIE ZAMKNIĘTEJ

Streszczenie

Celem rozdziału jest weryfikacja poprawności modelu matematycznego głośnika niskotonowego zaimplementowanego w środowisku MATLAB. Metoda badawcza obejmowała wyznaczenie teoretycznej charakterystyki ciśnienia akustycznego (SPL) na podstawie parametrów Thiele-Smalla oraz jej walidację za pomocą pomiarów rzeczywistych w warunkach quasi-bezechowych. Analiza porównawcza wykazała wysoką zgodność modelu z pomiarem w zakresie niskich częstotliwości, co potwierdza rzetelność zastosowanych algorytmów numerycznych opartych na parametrach skupionych. Zidentyfikowano również ograniczenia modelu wynikające z efektu *baffle step* oraz zjawiska *cone breakup* powyżej 1 kHz, których uproszczony model matematyczny nie uwzględnia. Uzyskane wyniki potwierdzają użyteczność opracowanego skryptu jako rzetelnego narzędzia inżynierskiego wspomagającego proces projektowania zestawów głośnikowych typu *closed-box*.

Słowa kluczowe: thiele-small, model matematyczny, charakterystyka SPL, pomiary akustyczne.

1. Wprowadzenie

Współcześnie projektowanie zestawów głośnikowych w dużej mierze opiera się na precyzyjnym przewidywaniu zachowania przetworników jeszcze przed etapem budowy fizycznego prototypu. Modelowanie matematyczne pozwala nie tylko zaoszczędzić czas i surowce, ale przede wszystkim umożliwia optymalizację parametrów akustycznych w kontrolowanym środowisku.

Fundamentem nowoczesnej elektroakustyki stały się prace Neville'a Thiele'a (1961) oraz Richarda Smalla (1972) [5]. Zdefiniowali oni zestaw parametrów, znanych dziś powszechnie jako parametry Thiele-Smalla (T-S), które opisują głośnik jako układ elektromechaniczny. Dzięki ich badaniom możliwe stało się matematyczne opisanie relacji między głośnikiem a objętością powietrza w obudowie, co zrewolucjonizowało podejście do projektowania zestawów głośnikowych.

Celem niniejszego rozdziału jest weryfikacja poprawności modelu matematycznego zaimplementowanego w środowisku MATLAB. Praca skupia się na porównaniu teoretycznie wyliczonej charakterystyki ciśnienia akustycznego (SPL)

z rzeczywistym pomiarem mikrofonowym głośnika, umieszczonego w obudowie zamkniętej. Analiza ta pozwoli ocenić użyteczność stworzonego narzędzia w praktycznym projektowaniu systemów audio.

Struktura rozdziału obejmuje szczegółowy opis parametrów TS, wyprowadzenie zależności logicznych modelu, prezentację implementacji skryptu w MATLABie oraz analizę wyników pomiarowych wraz z dyskusją dotyczącą napotkanych wyzwań pomiarowych.

2. Parametry Thiele-Smalla

Zrozumienie działania głośnika sprowadza się do przełożenia jego fizycznej budowy na równania matematyczne. Parametry Thielea-Smalla (TS) pozwalają modelować głośnik jako układ o parametrach skupionych, czyli uproszczony do kilku podstawowych elementów: masy (ciężar), sprężyny (zawieszenie) i opornika (cewka) [4].

Model ten najlepiej sprawdza się w zakresie małych sygnałów. Gdy głośnik gra z umiarkowaną głośnością, jego ruchy są przewidywalne i liniowe. Przy wyższych amplitudach znaczenie zaczyna mieć opór, zawieszenie pracuje nieliniowo, a ciepło wydzielane przez cewkę wpływa na jej rezystancję, co skutecznie utrudnia matematyczny opis.

Na potrzeby modelu w programie MATLAB, parametry badanego przetwornika podzieliliśmy na trzy grupy:

2.1. Parametry elektryczne

Opisują one właściwości cewki:

R_e (Rezystancja cewki): Jest to opór, jaki stawia drut w cewce, gdy płynie przez niego prąd stały;

L_e (Indukcyjność cewki): Cewka głośnika wytwarza pole magnetyczne. Indukcyjność sprawia, że przy wyższych częstotliwościach dźwięku, prąd napotyka wyższy opór. W efekcie głośnik naturalnie przestaje grać w wyższym paśmie częstotliwości.

2.2. Parametry mechaniczne

Dotyczą one fizycznego ruchu membrany oraz właściwości jej zawieszenia:

F_s (Częstotliwość rezonansowa): To częstotliwość głośnika, przy której najłatwiej wpada on w drgania. Im niższa jest wartość F_s, tym niższą częstotliwość potrafi wygenerować głośnik;

M_{ms} (Masa układu drgającego): To łączny ciężar ruchomych elementów głośnika: membrany, cewki oraz tzw. masy współdrgającej powietrza;

C_{ms} (Podatność zawieszeń): Mówi o tym, z jaką łatwością porusza się membrana głośnika. Przykładowo: jeśli C_{ms} wynosi 1 $\frac{mm}{N}$, to znaczy, że po naciśnięciu membrany z siłą 1 Newtona, wychyli się ona o 1 milimetr. Im wyższa wartość, tym głośnik łatwiej się porusza.

2.3. Parametry elektromechaniczne i akustyczne

Łączą one obie dziedziny i są kluczowe dla doboru obudowy.

V_{as} (Objętość ekwiwalentna): Objętość powietrza, która posiada taką samą podatność (sztywność) jak zawieszenia głośnika. Większe V_{as} sugeruje, że głośnik będzie wymagał większej obudowy;

Q_{ts} (Dobroć całkowita głośnika): To parametr bezwymiarowy (nieposiadający jednostki). Opisuje on stopień tłumienia układu, czyli jak szybko głośnik przestanie drgać po odcięciu sygnału. Przy niskim Q_{ts} (poniżej 0,4) głośnik jest sztywno kontrolowany przez magnes (szybko hamuje). Wysokie Q_{ts} (powyżej 0,6) zwiększa tendencję głośnika do swobodnego wibrowania. Basu jest więcej, lecz jest mniej precyzyjny;

Q_{tc} (Dobroć całkowita układu): Parametr opisujący głośnik po zamontowaniu w obudowie. Mówi o tym, jak zachowuje się membrana w zamkniętej przestrzeni;

Bl (Współczynnik siły): Współczynnik zależny od parametrów magnesu i cewki głośnika. Im wyższy Bl, tym mocniej magnes „potrafi" poruszyć cewką;

S_d (Powierzchnia czynna membrany): Efektywna powierzchnia, którą głośnik tłoczy powietrze. Znaczenie ma tu sama membrana z górną połową zawieszenia.

3. Model głośnika niskotonowego w obudowie zamkniętej

Obudowa zamknięta jest uznawana za najprostszy typ konstrukcji głośnikowej pod względem obliczeń i wykonania. Jej nadrzędnym celem jest całkowite wytłumienie fali dźwiękowej generowanej przez tylną stronę membrany.

3.1. Oddziaływanie objętości powietrza

Kluczowym zjawiskiem w tym modelu jest potraktowanie powietrza uwięzionego w obudowie zamkniętej (szczelnej) jako elementu sprężystego, współpracującego z zawieszeniem głośnika. To prowadzi do dwóch kluczowych zmian parametrów układu:

- podwyższenie częstotliwości rezonansowej z wartości F_s (dla głośnika bez obudowy) do wartości F_c (dla układu w obudowie);
- wzrost (pod warunkiem braku wypełnienia obudowy np. watą) dobroci wypadkowej do poziomu Q_{tc};

W modelu matematycznym zjawisko to odwzorowuje się poprzez wyznaczenie nowej, wypadkowej podatności zawieszenia C_{ms_box}. Parametr ten określa, o ile sztywniejszy stał się układ po dodaniu oporu powietrza w obudowie o konkretnej objętości V_b:

$$C_{ms\_box} = \frac{C_{ms}}{1+\frac{V_{as}}{V_b}} \quad (1)$$

gdzie: C_{ms_box} – wypadkowa podatność układu w obudowie,
C_{ms} – katalogowa podatność zawieszeń głośnika,
V_{as} – objętość ekwiwalentna głośnika,
V_b – objętość netto obudowy.

Z powyższej zależności wynika, że im mniejsza jest objętość obudowy V_b, tym mniejsza staje się wypadkowa podatność, co oznacza wzrost sztywności całego układu drgającego. Zmiana ta jest bezpośrednią przyczyną wzrostu dobroci całkowitej do poziomu Q_{tc}, co w modelu matematycznym opisuje zależność:

$$Q_{tc} = Q_{ts} \cdot \sqrt{\frac{C_{ms}}{C_{ms\_box}}} \quad (2)$$

gdzie: Q_{tc} – dobroć całkowita głośnika w obudowie,
Q_{ts} – dobroć całkowita głośnika w wolnej przestrzeni.

3.2. Model zależności ciśnienia akustycznego (SPL) w funkcji częstotliwości

Aby wyznaczyć teoretyczny poziom ciśnienia akustycznego SPL (ang. Sound Pressure Level) w funkcji częstotliwości, model realizuje następujący ciąg zależności:

Przetworzenie sygnału elektrycznego na siłę mechaniczną: Napięcie dostarczone ze wzmacniacza U generuje przepływ prądu przez cewkę. Dzięki obecności w szczelinie magnesu stałego, na cewkę działa siła elektrodynamiczna F, która jest wprost proporcjonalna do natężenia prądu I oraz współczynnika siły Bl. Trafne jest tu porównanie do silnika elektrycznego, gdzie energia elektryczna zamieniana jest na ruch mechaniczny.

$$F = Bl \cdot I \quad (3)$$

gdzie: F – siła elektrodynamiczna,
Bl – współczynnik siły (iloczyn indukcji magnetycznej i długości drutu cewki),
I – natężenie prądu płynącego przez cewkę.

Wyznaczenie impedancji mechanicznej Z_m: Ruch membrany nie odbywa się bez oporów. Model matematyczny uwzględnia wspomniane wcześniej siły bezwładności (wynikające z masy M_{ms}), tarcie wewnętrzne zawieszeń R_{ms} oraz sztywność wynikającą z połączenia zawieszenia i powietrza w obudowie C_{ms_box}. Suma tych oporów tworzy impedancję mechaniczną Z_m [1].

Obliczanie prędkości membrany v i wpływ siły przeciwelektromotorycznej: Prędkość, z jaką porusza się membrana, zależy od siły napędowej oraz

całkowitego oporu układu. Kluczowym elementem jest tutaj uwzględnienie zjawiska sprzężenia elektromechanicznego $(Bl)^2$ [1]. Podczas ruchu cewki w polu magnetycznym generuje ona własne napięcie, które przeciwdziała napięciu ze wzmacniacza. Im silniejszy jest układ magnetyczny, tym lepiej głośnik jest kontrolowany i szybciej wytłumia niepożądane drgania.

$$v(s) = \frac{Bl \cdot U(s)}{Z_e(s) \cdot Z_m(s) + (Bl)^2} \tag{4}$$

gdzie: $v(s)$ – prędkość membrany w dziedzinie częstotliwości,
s – operator Laplace'a ($s = j2\pi f$),
$U(s)$ – napięcie wejściowe zasilające cewkę,
$Z_e(s)$ – impedancja elektryczna cewki,
$Z_m(s)$ – impedancja mechaniczna układu drgającego,
$(Bl)^2$ – współczynnik sprzężenia elektromechanicznego.

Powyższa zależność pokazuje, że prędkość membrany v ściśle zależy od częstotliwości sygnału wejściowego. Operator s (reprezentujący częstotliwość) jest integralną częścią impedancji elektrycznej Z_e oraz mechanicznej Z_m.

Promieniowanie ciśnienia akustycznego: Drgająca membrana o określonej powierzchni S_d porusza się z wyliczoną prędkością v, co tworzy tzw. wydajność objętościową:

$$Q_{vol} = S_d \cdot v \tag{5}$$

gdzie: Q_{vol} – strumień objętościowy (wydajność objętościowa),
S_d – powierzchnia czynna membrany.

Ta wielkość, przy uwzględnieniu gęstości powietrza i odległości od głośnika, pozwala, na podstawie wzoru (4), wyliczyć ciśnienie akustyczne w polu dalekim p_{far}, wyrażone w Paskalach [Pa]:

$$p_{far}(s) = \frac{s \cdot \rho_0 \cdot e^{-\frac{s}{c}r}}{4\pi r} \cdot S_d \cdot \left[\frac{Bl \cdot U(s)}{(R_e + sL_e) \cdot \left(M_{ms}s + R_{ms} + \frac{1}{sC_{ms\_box}} \right) + (Bl)^2} \right] \tag{6}$$

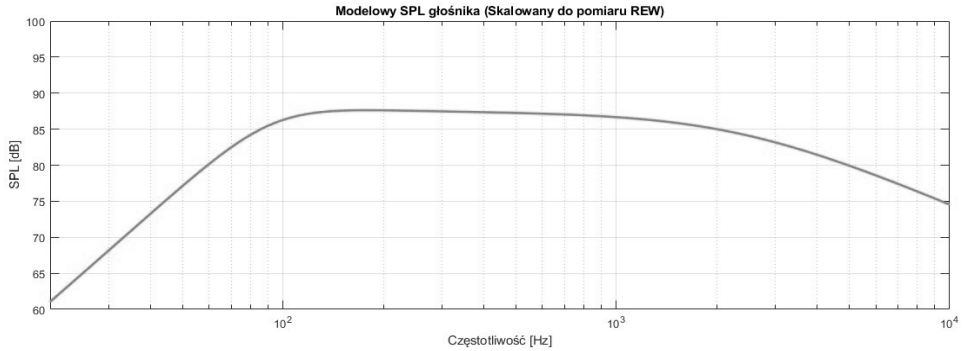
gdzie: $p_{far}(s)$ – ciśnienie akustyczne w polu dalekim,
ρ_0 – gęstość powietrza (ok. $1{,}18\,\frac{kg}{m^3}$),
c – prędkość dźwięku w powietrzu,
r – odległość mikrofonu od źródła,
R_e, L_e – rezystancja i indukcyjność cewki,
M_{ms}, R_{ms} – masa układu drgającego i mechaniczna rezystancja strat.

3.3. Teoretyczna charakterystyka przetwarzania SPL głośnika

Poziom SPL wyrażonego w decybelach (dB):

$$\mathrm{SPL}(s) = 20\log_{10}\frac{|p_{far}(s)|}{p_{ref}} + Offset \tag{7}$$

gdzie: $\mathrm{SPL}(s)$ – poziom ciśnienia akustycznego,
p_{ref} – ciśnienie odniesienia (próg słyszalności) $(20 \cdot 10^{-6}\ Pa)$,
$Offset$ – parametr kalibracyjny, uwzględnienie rzeczywistej efektywności akustycznej głośnika.



Rys. 1. Modelowy SPL głośnika
Fig. 1. Theoretical SPL of the loudspeaker

Jak przedstawiono na rysunku 1., dla obudowy zamkniętej model przewiduje, że poniżej częstotliwości rezonansowej układu F_c następuje spadek efektywności o stałym nachyleniu 12 dB na oktawę.

Spadek charakterystyki SPL powyżej 1 kHz w modelu matematycznym wynika z indukcyjności cewki L_e. Wraz ze wzrostem częstotliwości, impedancja elektryczna Z_e wzrasta, co ogranicza prąd płynący przez głośnik, a w konsekwencji redukuje ciśnienie akustyczne.

4. Implementacja modelu w środowisku MATLAB

Celem etapu implementacyjnego było stworzenie narzędzia, które na podstawie parametrów Thiele-Smalla głośnika oraz danych obudowy, wygeneruje charakterystykę częstotliwościową SPL. Wybór środowiska MATLAB podyktowany był jego natywną obsługą operacji macierzowych oraz liczb zespolonych, co jest konieczne przy analizie układów elektroakustycznych w dziedzinie częstotliwości.

4.1. Struktura algorytmu

Opracowany skrypt realizuje obliczenia w sposób wektorowy. Oznacza to, że zamiast stosowania pętli iteracyjnych dla każdej częstotliwości z osobna, zdefiniowano wektor częstotliwości f, a wszystkie operacje matematyczne (wyznaczanie impedancji, prędkości membrany, ciśnienia) wykonywane są jednocześnie dla całego pasma akustycznego. Kod źródłowy programu został zamieszczony w pracy [9].

Algorytm podzielono na cztery logiczne bloki:

1. Wprowadzenie danych katalogowych głośnika, parametrów fizycznych powietrza oraz objętości obudowy.
2. Przeliczenie podatności zawieszenia C_{ms} na wypadkową podatność układu w obudowie C_{ms_box} co uwzględnia sztywność poduszki powietrznej wewnątrz obudowy.
3. Wyznaczenie impedancji elektrycznej i mechanicznej w dziedzinie zespolonej $s = j\omega$, a następnie obliczenie prędkości wychylenia membrany v na podstawie przyłożonego napięcia.
4. Przeliczenie prędkości membrany na strumień objętościowy i finalnie na poziom ciśnienia akustycznego w dB.

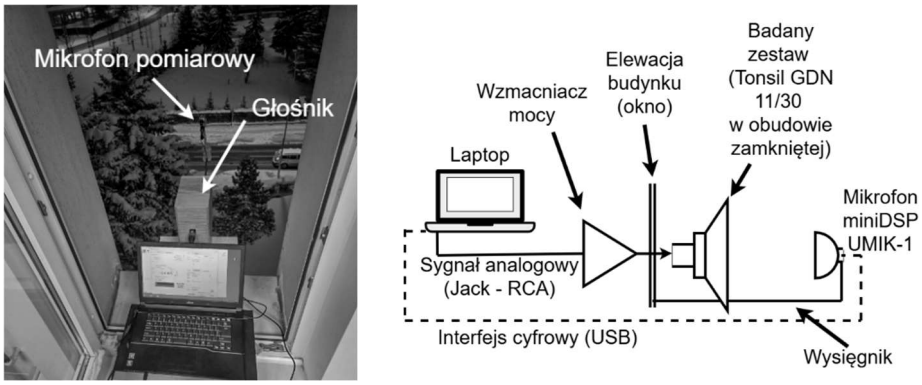
5. Pomiary rzeczywistego układu głośnika w obudowie zamkniętej

Nadrzędnym celem przeprowadzonych badań eksperymentalnych jest walidacja opracowanego modelu matematycznego oraz ocena rzetelności symulacji wykonanej w środowisku MATLAB. Porównanie wyników teoretycznych z rzeczywistą odpowiedzią częstotliwościową zestawu pozwala określić granice stosowalności parametrów (T-S) w projektowaniu systemów audio oraz zidentyfikować nieliniowości przetwornika, których model skupiony nie jest w stanie przewidzieć.

5.1. Stanowisko pomiarowe

Ze względu na brak dostępu do profesjonalnej komory bezechowej, zdecydowano się na realizację pomiaru w warunkach quasi-bezechowych. Metoda ta polega na umieszczeniu badanego głośnika w przestrzeni wolnej od przeszkód odbijających dźwięk.

Zestaw głośnikowy został wystawiony na dedykowanym stelażu przez okno na trzecim piętrze budynku, co pozwoliło na umieszczenie membrany głośnika w znacznej odległości od podłoża oraz ścian elewacji. Mikrofon MiniDSP UMIK-1 umieszczono na osi głośnika (on-axis) w odległości około 70 cm na sztywnym wysięgniku. Takie ustawienie pozwoliło zminimalizować wpływ wczesnych odbić dźwięku, które w typowym pomieszczeniu uniemożliwiłyby rzetelną analizę pasma poniżej 300 Hz.



Rys. 2. Stanowisko pomiarowe: a) widok rzeczywisty, b) schemat układu

Fig. 2. Measurement setup: a) real view, b) system diagram

Zastosowano tor pomiarowy składający się z laptopa z oprogramowaniem Room EQ Wizard, wzmacniacza oraz cyfrowego mikrofonu pomiarowego. Poglądowy układ przedstawiono na rysunku 2a, zaś szczegółową architekturę połączeń na rysunku 2b.

Do pomiaru wykorzystano głośnik Tonsil GDN11/30 8 Ω, o parametrach przedstawionych w tabeli 1.

Tabela 1. Zestawienie parametrów T-S głośnika Tonsil GDN 11/30 8 Ω

Table 1. T-S parameters of the Tonsil GDN 11/30 8 Ω loudspeaker

Parametr T-S	**Wartość parametru w badanym głośniku**
R_e	6,75 Ω
L_e	0,44 mH
Fs	55 Hz
M_{ms}	7,9 g
C_{ms}	$1{,}1 \cdot 10^{-3}\,\frac{\mathrm{m}}{\mathrm{N}}$
V_{as}	6 dm^3
Q_{ts}	0,5
Bl	5,6 Tm
S_d	62,4 cm^2

Źródło: sklep internetowy Tonsil Polska.

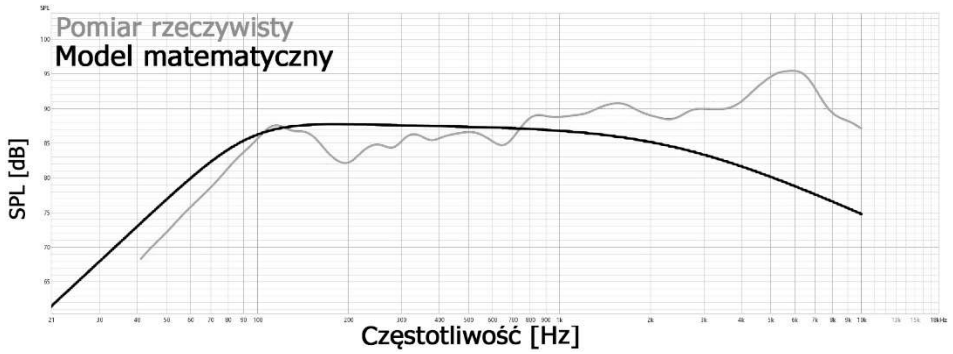
Source: Tonsil Poland online store.

Objętość netto wykorzystanej obudowy wynosi 3,5 dm^3.

Badanie przeprowadzono z wykorzystaniem sygnału typu *sweep* w paśmie 40–10 000 Hz przy amplitudzie wyjściowej około 90 dB. W celu zwiększenia stosunku sygnału do szumu (SNR), szczególnie istotnego w warunkach pomiaru

zewnętrznego, procedurę powtórzono 29-krotnie, a wyniki poddano uśrednianiu wektorowemu (*Vector average*).

W procesie obróbki danych zastosowano wygładzanie o rozdzielczości 1/12 oktawy. Pozwoliło to na usunięcie drobnych oscylacji wynikających z szumu tła, przy jednoczesnym zachowaniu widoczności charakterystycznych rezonansów własnych membrany głośnika. Efekt końcowy jest widoczny na rys. 3.



Rys. 3. Porównanie pomiaru odpowiedzi częstotliwościowej głośnika do modelu matematycznego

Fig. 3. Comparison of the measured frequency response (grey line) with the mathematical model (black line)

5.2. Wyniki pomiaru wraz z odniesieniem do modelu

Analiza porównawcza zmierzonej charakterystyki SPL z modelem wygenerowanym w programie MATLAB wykazuje wysoką zgodność w zakresie niskich i średnich tonów.

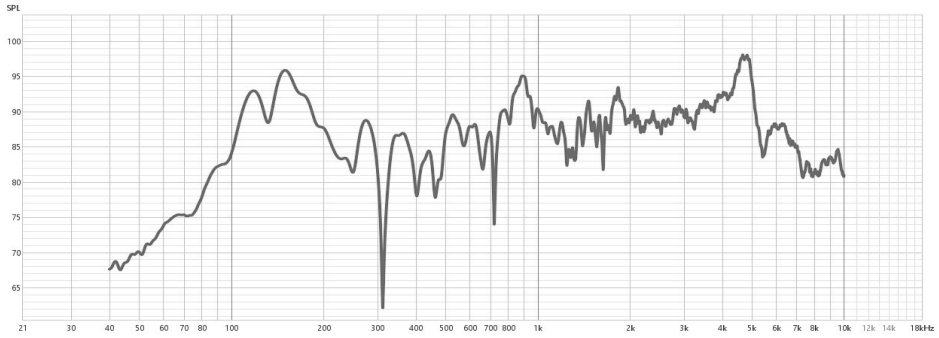
Główne obserwacje wynikające z porównania to:

1. Zgodność pasma basowego: Model poprawnie przewidział spadek efektywności poniżej częstotliwości rezonansowej F_c typowy dla obudowy zamkniętej (12 dB/oct). Widoczne na pomiarze podbicie w zakresie 100–150 Hz koresponduje z obliczoną dobrocią układu $Q_{tc} \approx 0{,}82$, co potwierdza poprawność przyjętej objętości netto 3,5 dm^3.
2. Obserwowany na pomiarze rzeczywistym wzrost efektywności w zakresie średniotonowym (200 Hz – 1 kHz), jest wynikiem zjawiska dyfrakcji (odbić) fal dźwiękowych na krawędziach obudowy oraz efektu *baffle step* (odbić od przedniej ściany obudowy), których uproszczony model matematyczny nie uwzględnia.
3. Powyżej częstotliwości 1 kHz widoczne są znaczne nieregularności pomiaru rzeczywistego, wynikające z rezonansów strukturalnych membrany (tzw. *cone breakup*. Membrana zamiast sztywno tłoczyć powietrze ulega odkształceniom). Model teoretyczny, zakładający pracę przetwornika jako sztywnego obiektu, wykazuje w tym zakresie płynny spadek wynikający

z indukcyjności cewki L_e, co stanowi naturalne ograniczenie symulacji opartej na parametrach skupionych.

5.3. Analiza błędów i ograniczeń metody pomiarowej

Podczas realizacji badań eksperymentalnych zidentyfikowano czynniki wpływające na rzetelność wyników. Najistotniejszym problemem napotkanym w pierwszej fazie pomiarów był, widoczny na rysunku 4, głęboki spadek efektywności (tzw. *dip*) w okolicy $300\ Hz$.



Rys. 4. Pomiar wstępny, wykonany w zbyt bliskiej odległości od ściany elewacji

Fig. 4. Preliminary measurement performed too close to the building face

Analiza wykazała, że błąd wynikał ze zbyt bliskiego umiejscowienia zestawu względem płaszczyzny elewacji budynku i parapetu. Powodowało to nakładanie się fali bezpośredniej z falą odbitą od twardych powierzchni (tzw. filtr grzebieniowy), co prowadziło do destruktywnej interferencji. Problem wyeliminowano poprzez zwiększenie wysięgu statywu poza obrys budynku, co pozwoliło na uzyskanie rzetelnej charakterystyki.

6. Podsumowanie i wnioski

Przeprowadzone badania oraz analiza porównawcza pozwoliły na pełną walidację modelu teoretycznego głośnika niskotonowego w obudowie zamkniętej. Na podstawie zgromadzonych danych sformułowano następujące wnioski:

Symulacja zaimplementowana w środowisku MATLAB wykazała wysoką zgodność z pomiarem rzeczywistym w zakresie niskich częstotliwości, szczególnie w przedziale do około 200 Hz. Potwierdza to wysoką użyteczność parametrów Thiele-Smalla jako fundamentu projektowania obudów typu zamkniętego.

Zaobserwowane na pomiarze podbicie w zakresie 100 – 200 Hz precyzyjnie koresponduje z obliczoną dobrocią układu $Q_{tc} \approx 0{,}82$. Potwierdza to rzetelność przyjętej objętości netto 3,5 l oraz poprawność uwzględnienia sztywności „poduszki powietrznej” w modelu matematycznym.

Rozbieżności powyżej 1 kHz wynikają z uproszczeń modelu, który traktuje membranę jako sztywny obiekt drgający. Rzeczywisty pomiar ujawnił zjawiska, których parametry skupione nie uwzględniają, takie jak: *baffle step* oraz *cone breakup* przy częstotliwości ok. 5 kHz.

Wykazano, że pomiar quasi-bezechowy zrealizowany na wysięgniku poza obrysem budynku jest skuteczną alternatywą dla profesjonalnych komór bezechowych. Analiza błędu w postaci „dipu" przy częstotliwości 300 Hz pozwoliła na optymalizację stanowiska i udowodniła, że kluczowym czynnikiem jest odpowiednie rozplanowanie stanowiska pomiarowego.

Opracowany skrypt MATLAB, dzięki wektorowemu przetwarzaniu danych i poprawnej obsłudze liczb zespolonych, stanowi skuteczne narzędzie inżynierskie pozwalające na optymalizację parametrów zestawu przed etapem budowy prototypu.

Literatura

[1] Beranek L. L., Acoustics, McGraw-Hill, New York 1954, s. 116-125; 118–121.

[2] Dickason V., *Loudspeaker Design Cookbook*, 7th Edition, Audio Amateur Publications, Peterborough 2006.

[3] Krajewski J., *Głośniki i zestawy głośnikowe*, Wydawnictwa Komunikacji i Łączności, Warszawa 2008, s. 34-40.

[4] miniDSP Ltd, *UMIK-1 User Manual*, Hong Kong 2021.

[5] Small R. H., *Direct-Radiator Loudspeaker System Analysis*, „Journal of the Audio Engineering Society", Vol. 20, No. 5, 1972, s. 383–395.

[6] The MathWorks Inc., MATLAB Statistics and Machine Learning Toolbox, Natick 2023.

[7] Thiele A. N., *Loudspeakers in Vented Boxes: Part I*, „Journal of the Audio Engineering Society", Vol. 19, No. 5, 1971, s. 382–392.

[8] Tonsil Poland, *Karta katalogowa głośnika GDN 11/30*, [dostęp: 31.03.2026].

[9] Żarski K., Kod w środowisku MATLAB (Głośnik w obudowie zamkniętej), repozytorium GitHub: https://github.com/Masterkaj2137/Kod-matlab-glosnik-w-obudowie-zamknietej [dostęp: 31.03.2026].

LOUDSPEAKER MODEL IN A CLOSED ENCLOSURE

Summary

The aim of the article is to verify the correctness of the mathematical model of a woofer implemented in the MATLAB environment. The research methodology involved determining the theoretical sound pressure level (SPL) characteristics based on Thiele-Small parameters and its validation through real-world measurements under quasi-anechoic conditions. Comparative analysis showed high agreement between the model and the measurement in the low-frequency range, which confirms the reliability of the applied numerical algorithms based on lumped parameters. Model limitations resulting from the *baffle step* effect and the *cone breakup*

phenomenon above 1 kHz, which the simplified mathematical model does not account for, were also identified. The obtained results confirm the utility of the developed script as a reliable engineering tool supporting the design process of closed-box loudspeaker systems.

Keywords: thiele-small, mathematical model, SPL characteristics, acoustic measurements.

Marcin Bator
Studenckie Koło Naukowe Informatyków „Kod”

DOSTARCZANIE INFRASTRUKTURY APLIKACJI Z UŻYCIEM TERRAFORM NA PRZYKŁADZIE CHMURY AZURE I KONTENERÓW DOCKER

Streszczenie

Rozdział przedstawia Terraform jako narzędzie stosowane w podejściu IaC (Infrastructure as Code, infrastruktura jako kod) do dostarczania infrastruktury potrzebnej do testowania lub wdrażania systemów. Autor omawia zalety narzędzia zaczynając od utworzenia prostego serwera HTTP, jego konteneryzacji i uruchomienia na platformie Azure, stopniowo korzystając z kolejnych możliwości oferowanych przez Terraform, takich jak zdefiniowanie kilku dostawców (provider) dla środowisk produkcyjnych i testowych, a następnie jednoczesne wdrożenie aplikacji w każdym z nich z użyciem modułów. Praca ukazuje na przykładach korzyści jakie przynoszą cechy charakterystyczne narzędzia Terraform, takie jak idempotentność, a więc zdolność do samodzielnego wykrycia różnic między istniejącą a pożądaną infrastrukturą, co zapobiega niepotrzebnemu tworzeniu tych samych zasobów kilkukrotnie.

Słowa kluczowe: terraform, konteneryzacja, infrastruktura, wdrożenie, chmura.

1. Wprowadzenie

W czasie, gdy zapotrzebowanie na szybkie i bezproblemowe testowanie oraz wdrażanie nowych aplikacji w zróżnicowanych środowiskach jest najwyższe w historii, a zużycie zasobów chmurowych rośnie z dnia na dzień, niezbędną wydaje się możliwość tworzenia, zarządzania i niszczenia infrastruktury w sposób deklaratywny i łatwy do zautomatyzowania. Podejście *Infrastructure as Code* (infrastruktura jako kod) zrewolucjonizowało sposób myślenia o zasobach chmurowych i stanowiskach deweloperskich, pozwalając na zdefiniowanie ich w kodzie, który może być przechowywany w tym samym repozytorium, co wdrażana aplikacja. Jednym z narzędzi IaC jest Terraform - narzędzie umożliwiające bezpieczne i efektywne dostarczanie infrastruktury [6].

2. Terraform jako narzędzie IaC

Terraform jest narzędziem stworzonym przez firmę HashiCorp. Pozwala na zdefiniowanie potrzebnych do wdrożenia danego systemu elementów (maszyn wirtualnych, sieci, kontenerów itp.) w plikach *.tf* za pomocą języka HCL (*HashiCorp Configuration Language*), zbliżonego w składni do notacji JSON (*JavaScript Object Notation*). Dzięki temu kod infrastruktury może być wersjonowany w repozytorium, np. za pomocą narzędzia Git.

Kolejną zaletą narzędzia jest jego deklaratywność. Oznacza to, że programista musi jedynie zdefiniować pożądany, końcowy stan infrastruktury, a Terraform sam ustali, jakie zmiany na platformie docelowej są konieczne do jej uzyskania. Jest to możliwe dzięki możliwości przechowywania aktualnego stanu infrastruktury w pliku *.tfstate*, który może znajdować się zarówno na lokalnym urządzeniu użytkownika, jak i na zdalnym serwerze plików. Ta cecha zapewnia idempotentność, co oznacza, że wielokrotne zastosowanie tej samej konfiguracji nie spowoduje niepożądanych zmian.

W kolejnych rozdziałach niniejszej pracy opisano przykład użycia narzędzia Terraform do dostarczenia (*provisioning*) infrastruktury dla prostego serwisu REST API w języku Python konteneryzowanego za pomocą technologii Docker [3]. Wdrożenie będzie odbywało się równolegle na platformie Azure oraz w lokalnym środowisku Docker na urządzeniu użytkownika.

3. Przygotowanie wdrażanej aplikacji

W ramach niniejszej pracy wykorzystana zostanie prosta aplikacja uruchamiająca serwer HTTP z jednym punktem końcowym (*endpoint*) REST API zwracającym tekst w formacie JSON zawierający wartość ustawioną przez zmienną środowiskową.

3.1. Kod aplikacji w Python

Aplikacja do uruchomienia serwera używa pakietu `http.server` dostępnego w środowisku Python. Serwer (przedstawiony w listingu 1.) zostaje uruchomiony na porcie 8080, a punkt końcowy zdefiniowany w metodzie `do_GET` nasłuchujący na adresie domyślnym zwraca odpowiedź w formacie JSON zawierającą pola `title` (ze stałą wartością `Hello World`) oraz `text` (z wartością wczytywaną ze zmiennej środowiskowej `TEXT_CONTENT`). Dzięki jej użyciu można sprawdzić, czy zmienne środowiskowe są odczytywane prawidłowo. W związku z tym zmienna środowiskowa `TEXT_CONTENT` musi być dostępna w środowisku uruchomieniowym aplikacji. Oznacza to, że może być ustawiona np. ręcznie w terminalu, w pliku *.env* w katalogu aplikacji, lub globalnie w systemie bądź kontenerze.

Pakiet `http.server` nie jest przeznaczony do zastosowań produkcyjnych, ponieważ implementuje jedynie podstawowe funkcje bezpieczeństwa [5]. Został wykorzystany w przykładzie ze względu na łatwość implementacji.

```
from http.server import BaseHTTPRequestHandler, HTTPServer
import json
import os
from urllib.parse import urlparse

class SimpleHandler(BaseHTTPRequestHandler):
    def do_GET(self):
        self.send_response(200)
        self.send_header('Content-Type', 'application/json')
        self.end_headers()

        text_content = os.getenv('TEXT_CONTENT', 'Default message from
             server')

        response_data = {
            "title": "hello world",
            "text": text_content,
        }

        json_response = json.dumps(response_data, ensure_ascii=False)
        self.wfile.write(json_response.encode('utf-8'))

def run(server_class=HTTPServer, handler_class=SimpleHandler):
    server_address = ('', 8080)
    httpd = server_class(server_address, handler_class)

    httpd.serve_forever()

if __name__ == '__main__':
    run()
```

Listing 1. Kod serwera REST w języku Python

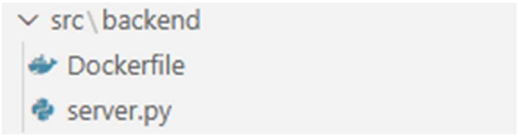
Listing 1. REST server code in Python

3.2. Konteneryzacja i publikacja obrazu

Do uruchomienia serwera wykorzystana zostanie technologia Docker pozwalająca na spakowanie środowiska aplikacji wraz z jej zależnościami do obrazu, a następnie uruchomienie go jako kontenera z wybraną konfiguracją. Pozwoli to na łatwe wdrożenie aplikacji w dowolnym środowisku, bez konieczności instalacji Pythona i jego zależności na maszynie docelowej. Jak widać na rysunku 1., plik *Dockerfile* będący „receptą" na utworzenie obrazu aplikacji znajduje się w tym samym katalogu, co plik z kodem serwera.

W listingu 2. zawierającym kod pliku *Dockerfile* widać, że jako obraz bazuje na dystrybucji Debiana z zainstalowanym Python 3.9. Po uruchomieniu komendy

`docker build` kopiowana jest zawartość folderu, w którym znajduje się plik, potem za pomocą komendy Pythona uruchamiany jest serwer, zaś na końcu na zewnątrz obrazu wystawiony zostaje port 8080 - ten, na którym serwer nasłuchuje.



Rys. 1. Struktura plików serwera zasobów
Fig. 1. Resource server file structure

Zbudowanie obrazu z pliku Dockerfile jest możliwe dzięki komendzie `docker build`. Należy podać nazwę i tag budowanego obrazu oraz ścieżkę (kontekst), w jakiej znajduje się plik. Kolejnym krokiem jest opublikowane obrazu do rejestru kontenerów, tak, aby możliwe było jego ściągnięcie na maszynie docelowej, na której zostanie uruchomiony. Po zalogowaniu się do rejestru za pomocą komendy `docker login` można opublikować w nim obraz poleceniem `docker push`.

```
FROM python:3.9-slim

WORKDIR /app
COPY ..

CMD [“python”, “server.py”]
EXPOSE 8080
```

Listing 2. Plik *Dockerfile*
Listing 2. *Dockerfile*

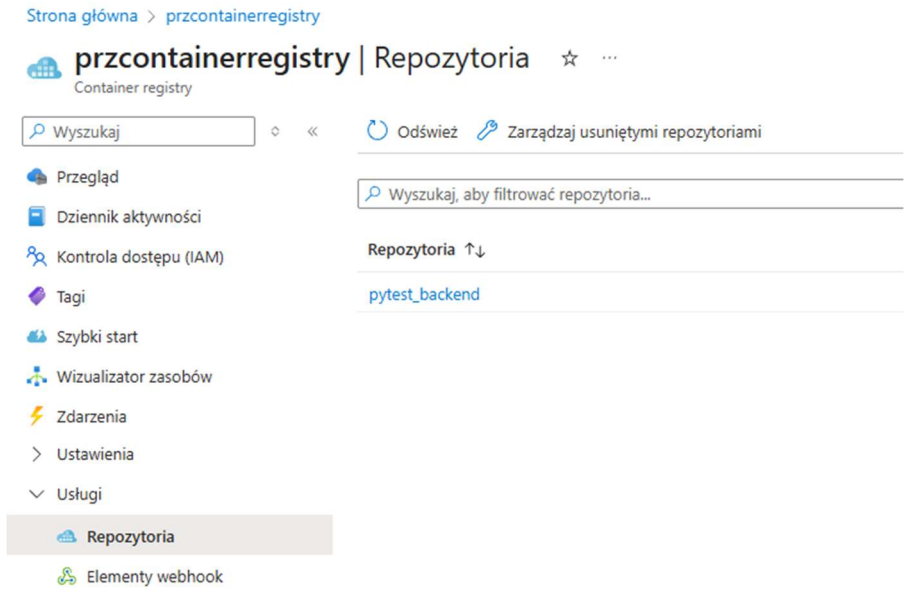
4. Przygotowanie chmury Azure

Aby możliwe było wdrażanie aplikacji do docelowej maszyny z poziomu dowolnego urządzenia, należy przechowywać obrazy, z których mają być zbudowane kontenery oraz plik stanu na zdalnym serwerze lub chmurze, do której narzędzie będzie miało dostęp [4]. Dobrym miejscem na umieszczenie powyższych elementów będzie chmura Azure [1].

4.1. Rejestr kontenerów

Rejestr kontenerów Azure (Azure Container Registry, ACR) jest elementem pozwalającym na przechowywanie i wersjonowanie obrazów Dockera. Mając do niego dostęp (unikalny adres, nazwa użytkownika i hasło) można na nim publikować obraz pobierać obrazy.

Widok interfejsu z utworzonym na platformie Azure rejestrem wraz z opublikowanym obrazem Dockera o nazwie *pytest_backend* został przedstawiony na rysunku 2.



Rys. 2. Rejestr kontenerów Azure

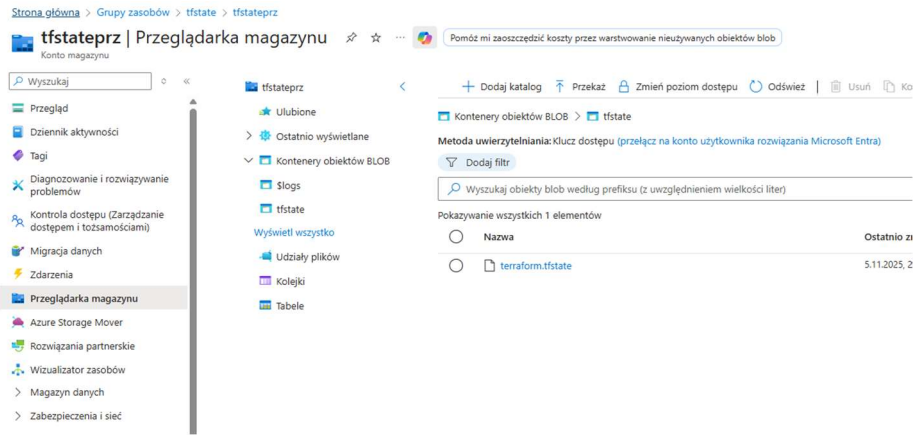
Fig. 2. Azure Container Registry

4.2. Kontener *blob* na plik stanu Terraform

Kolejnym krokiem jest utworzenie konta magazynu Azure (Storage Account). Ten zasób pozwala między innymi na przechowywanie obiektów *blob* zabezpieczonych przed jednoczesną modyfikacją, a więc doskonale nadaje się na przechowywanie pliku stanu Terraform. Utworzone konto o nazwie *tfstateprz* zawierające plik stanu (*terraform.tfstate*) zostało przedstawione na rys. 3.

5. Wdrożenie aplikacji do Azure

Po przygotowaniu elementów koniecznych do uruchomienia narzędzia Terraform można przystąpić do definiowania wymaganej infrastruktury w kodzie języka HCL.



Rys. 3. Konto magazynu Azure z plikiem stanu Terraform

Fig. 3. Azure Storage Account with Terraform state file

5.1. Konfiguracja Terraform i providera Azure

Kod Terraform składa się z bloków rozgraniczonych za pomocą klamer. Każda konfiguracja musi zawierać blok *terraform*, w którym znajdzie się definicja *backendu* (określająca dostęp do pliku *.tfstate*). Wewnątrz bloku powinna znaleźć się również konfiguracja *required_providers*, która posłuży do zainstalowania providerów - łączników pomiędzy Terraformem a docelową platformą, na którą zostanie wdrożona aplikacja. W listingu 3. przedstawiono konfigurację bloku `terraform` oraz blok *provider* `azurerm` konfigurujący dostęp do *backendu* z plikiem stanu. W głównym pliku można też umieścić definicję grupy zasobów Azure, w której zostanie umieszczony kontener. Kod Terraform może być dowolnie rozdzielony pomiędzy pliki *.tf* znajdujące się w katalogu, z którego jest uruchamiany, więc dobrą praktyką jest zdefiniowanie bloku konfigurującego w pliku *main.tf*.

5.2. Zmienne Terraform

Grupa zasobów `azurerm_resource_group` z listingu 3. korzysta ze zmiennych (wyrażenia z przedrostkiem `var`). Zmienne Terraform pozwalają na zdefiniowanie parametrów możliwych do podania podczas uruchamiania kodu. Można je przekazać jako parametry podczas uruchamiania konfiguracji z konsoli lub ustawić jako zmienne środowiskowe z odpowiednim przedrostkiem. Terraform pozwala także na ustawienie wartości domyślnej zmiennej (wtedy nie jest ona wymagana) oraz na oznaczenie jej jako wrażliwą (`sensitive`) - nie jest ona wtedy jawnie drukowana w konsoli w podsumowaniu.

```
terraform {
  required_providers {
    azurerm = {
      source  = "hashicorp/azurerm"
      version = "~> 3.0"
    }
  }

  backend "azurerm" {
    resource_group_name  = "tfstate"
    storage_account_name = "tfstateprz"
    container_name       = "tfstate"
    key                  = "terraform.tfstate"
  }
}

provider "azurerm" {
  features {}
  storage_use_azuread = true
}

resource "azurerm_resource_group" "rg" {
  name     = var.resource_group_name
  location = var.location
}
```

Listing 3. Plik *main.tf*
Listing 3. *main.tf* file

Plik *variables.tf* zawierający definicje zmiennych został przedstawiony w listingu 4. Zawiera on zmienne określające nazwę grupy zasobów i innych elementów opisanych w dalszej części pracy oraz wartości pozwalające na dostęp do rejestru kontenerów i zmienne środowiskowe, które będą przekazane aplikacji.

Wartości zmiennych używanych przez Terraform można także zapisać w pliku o nazwie *terraform_tfvars*. Powinien być to plik ignorowany przez system kontroli wersji, ponieważ może zawierać wrażliwe wartości zmiennych, takie jak hasła i klucze dostępu. Plik powinien zawierać kod w formacie:

nazwa_zmiennej = „wartosc”.

```
variable "resource_group_name" {
  type = string
}
variable "container_env_name" {
  type = string
}
variable "location" {
  type = string
}
```

```
variable "acr_server" {
  type = string
}
variable "acr_username" {
  type = string
}
variable "acr_password" {
  type      = string
  sensitive = true
}

variable "backend_image_tag" {
  type = string
}
variable "app_name" {
  type = string
}
variable "backend_port" {
  type = string
}
variable "text_content" {
  type = string
}
```

Listing 4. Plik *variables.tf*
Listing 4. *variables.tf* file

5.3. Środowisko kontenerów

Kontenery na platformie Azure do działania w obrębie danej grupy zasobów potrzebują środowiska kontenerów. Można wewnątrz niego umieścić większą liczbę pojedynczych kontenerów, jednak w darmowych wersjach subskrypcji Azure jest możliwe stworzenie tylko jednego środowiska kontenerów. Jego tworzenie i usuwanie zazwyczaj zajmuje sporo czasu ze względu na konieczność zarezerwowania zasobów chmurowych. Kod definiujący zasób środowiska kontenerów został przedstawiony w listingu 5.

```
resource "azurerm_container_app_environment" "container_env" {
  name                = var.container_env_name
  location            = azurerm_resource_group.rg.location
  resource_group_name = azurerm_resource_group.rg.name
}
```

Listing 5. Zasób środowiska kontenerów
Listing 5. Container environment resource

Po zdefiniowaniu wszystkich powyższych zasobów można wreszcie stworzyć główny zasób aplikacji, a więc sam kontener (element `azurerm_`

container_app platformy Azure). Jego kod został przedstawiony w listingu 6. Oprócz podania nazwy kontenera i nadrzędnej grupy zasobów oraz środowiska, wewnątrz bloku registry podane są dane dostępowe do rejestru ACR z którego ma zostać pobrany obraz kontenera. Najważniejszym elementem jest blok template, który określa nazwę obrazu bazowego dla kontenera, a także zasoby procesora i pamięci, jakie mają zostać mu przyznane. Ustawiana w nim jest także wartość zmiennej środowiskowej TEXT_CONTENT, która zostanie użyta przez aplikację serwera HTTP. Ostatnim blokiem jest ingress, który pozwala na otwarcie ruchu przychodzącego na określony port kontenera i wystawienie go na publiczny adres URL.

5.4. Zatwierdzenie infrastruktury

Po zdefiniowaniu całej niezbędnej infrastruktury można przejść do wykonywania kodu Terraform. Pierwszą komendą, jaka jest potrzebna do inicjalizacji *backendu* i pobrania *providerów* jest terraform init. Przed jej wykonaniem należy upewnić się, że zmienne środowiskowe zawierające dane dostępowe konieczne do połączenia z plikiem *backendu* (w tym przypadku z platformą Azure) są dostępne w kontekście konsoli. Można je załadować poprzez utworzenie pliku *.env* ze zmiennymi *ARM_CLIENT_ID*, *ARM_CLIENT_SECRET*, a także ARM_SUBSCRIPTION_ID oraz ARM_TENANT_ID, a następnie uruchomienie komendy SetEnvironmentVariable (na Windows), lub export na Linux. Po poprawnym zainicjalizowaniu *backendu* Terraform na konsoli powinien pojawić się komunikat *Terraform has been successfully initialized.*

Kolejnym krokiem jest zaplanowanie wdrożenia infrastruktury. Terraform porówna wtedy istniejącą na platformie docelowej konfigurację z tą, jaka została zdefiniowana w plikach *.tf* i pokaże zmiany konieczne do wykonania, by były one identyczne. Można to zrobić za pomocą komendy terraform plan lub od razu uruchomić terraform apply (narzędzie zapyta wtedy o potwierdzenie wykonania zaplanowanych akcji).

```
resource "azurerm_container_app" "backend" {
  name                         = "${var.app_name}backend"
  container_app_environment_id = az-urerm_container_app_environ-
            ment.container_env.id
  resource_group_name          = azurerm_resource_group.rg.name

  registry {
    server               = var.acr_server
    username             = var.acr_username
    password_secret_name = "registry-password"
  }
  secret {
    name  = "registry-password"
    value = var.acr_password
```

```
  }
  template {
    container {
      name   = "pytest-backend"
      image  = var.backend_image_tag
      cpu    = 0.25
      memory = "0.5Gi"

      env {
        name  = "TEXT_CONTENT"
        value = var.text_content
      }
    }

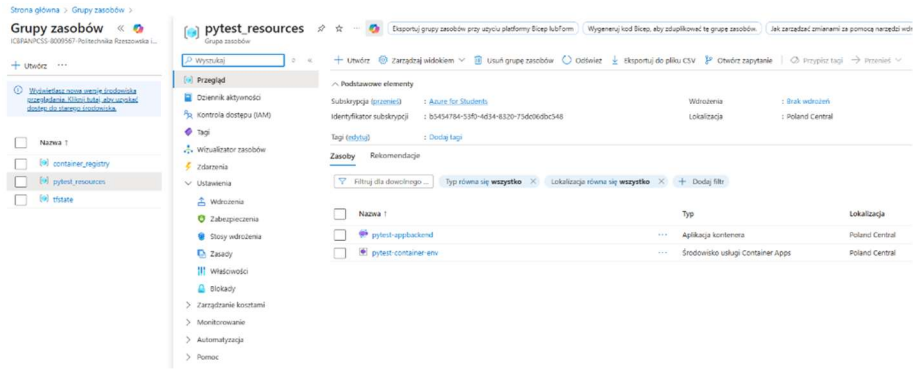
    min_replicas = 1
    max_replicas = 3
  }
  ingress {
    external_enabled = true
    target_port      = var.backend_port
    transport        = "http"

    traffic_weight {
      latest_revision = true
      percentage      = 100
    }
  }
}
```

Listing 6. Definicja kontenera

Listing 6. Container definition

Po poprawnym wykonaniu wszystkie zasoby zostały pomyślnie utworzone, co zostało przedstawione na rys. 4.



Rys. 4. Infrastruktura pomyślnie wdrożona na chmurę Azure

Fig. 4. Infrastructure correctly deployed to Azure platform

Po kliknięciu w zasób kontenera *pytest-appbackend* po prawej stronie dostępny jest adres URL, który przekierowuje na otwarty przez kontener port. Otworzenie go w przeglądarce spowoduje wyświetlenie odpowiedzi serwera HTTP w formacie JSON zawierającej ustawione odpowiednio pola, co przedstawiono na rys. 5.



Rys. 5. Rezultat wywołania endpointu wdrożonego serwera HTTP

Fig. 5. The result of calling deployed HTTP server endpoint

6. Wdrożenie aplikacji w lokalnym Dockerze

Terraform może być używany nie tylko do wdrażania produkcyjnych i testowych środowisk na chmurze lub serwerze, ale także do przyspieszenia uruchamiania stanowiska deweloperskiego na lokalnym urządzeniu. Niniejszy rozdział zawiera opis kroków koniecznych do wyodrębnienia istniejącej konfiguracji wdrożenia na Azure do oddzielnego modułu Terraform (aby zachować czytelność kodu *.tf*) oraz utworzenia nowego modułu uruchamiającego kontener z serwerem HTTP wewnątrz lokalnej usługi Dockera.

6.1. Wyodrębnienie istniejącego wdrożenia do modułu Terraform

Moduły Terraform są szczególnie użyteczne w przypadku konieczności zdefiniowania wielu wdrożeń tego samego systemu używających różnych dostawców [2]. Aby wyodrębnić opisaną w poprzednim rozdziale konfigurację wdrożenia na Azure do oddzielnego modułu, można przenieść pliki zawierające opis grupy zasobów, środowiska kontenerów i samego kontenera do nowego katalogu *./modules/azure*. Oprócz tego konieczne jest utworzenie w tym samym katalogu pliku *variables.tf* zawierającego definicje zmiennych koniecznych do działania modułu Azure.

Ostatnim krokiem koniecznym do rozpoznania i użycia przez Terraform utworzonego modułu Azure jest dodanie w pliku *main.tf* – znajdującego się w głównym katalogu Terraform – definicji modułu zawierającej ścieżkę do niego oraz przekazanie koniecznych zmiennych (wymaganych przez plik *variables.tf* w katalogu modułu). Kod defincji modułu został przedstawiony w listingu 7.

```
module "azure" {
  source = "./modules/azure"

  resource_group_name = var.resource_group_name
  location            = var.location
```

```
  acr_server          = var.acr_server
  acr_username        = var.acr_username
  acr_password        = var.acr_password
  container_env_name  = var.container_env_name
  backend_image_tag   = var.backend_image_tag
  app_name            = var.app_name
  backend_port        = var.backend_port
  text_content        = var.text_content
}
```

Listing 7. Definicja modułu Azure

Listing 7. Azure module definition

Wyodrębniony w ten sposób moduł przy uruchomieniu komend Terraform zachowa się w identyczny sposób jak ukazano w poprzednim rozdziale (komenda `terraform apply` nie wykryje żadnych zmian do wykonania, a infrastruktura nie ulegnie zmianie). Należy jedynie przed jej wykonaniem uruchomić `terraform init --reconfigure`, ponieważ pojawił się nowy moduł, który należy wpisać do pliku stanu.

6.2. Wdrożenie z użyciem providera Docker

Po wyodrębnieniu modułu Azure można dodać kolejny moduł, który będzie definiował uruchomienie aplikacji w Dockerze. W tym celu należy utworzyć strukturę katalogów *./modules/docker*, a wewnątrz niego plik *container.tf* z definicją kontenera, której kod przedstawiono w listingu 8.

```
resource "docker_container" "backend" {
  name  = var.app_name
  image = var.backend_image_tag

  ports {
    internal = var.backend_port
    external = var.backend_port
  }

  env = [
    "TEXT_CONTENT=${var.text_content}",
  ]
}
```

Listing 8. Definicja kontenera Docker

Listing 8. Docker container definition

Należy także, tak jak poprzednio, utworzyć plik *variables.tf* zawierający zmienne konieczne do połączenia z ACR (w celu pobrania obrazu) oraz te dotyczące konfiguracji samego kontenera.

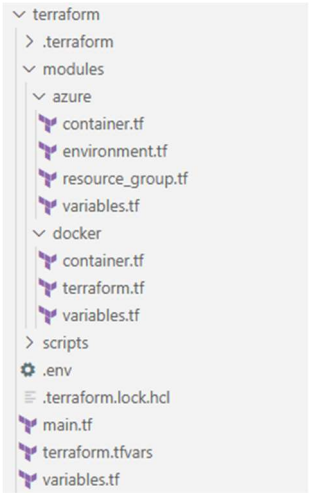
Kolejnym krokiem będzie dodanie nowego dostawcy (provider) w pliku *main.tf* w głównym katalogu Terraform. Do listy `required_providers` należy dodać element o źródle `kreuzwerker/docker`, a następnie utworzyć nowy blok `provider` zawierający konfigurację hosta (wskazującą na lokalny silnik Dockera na Windows) oraz dane uwierzytelniające potrzebne do połączenia z rejestrem ACR.

W odróżnieniu od modułu Azure (którego provider pochodził z rejestru *hashicorp* będącego rejestrem domyślnym), w przypadku modułu Docker należy w katalogu *./modules/docker* również dodać blok *terraform* z definicją `required_providers`, dzięki której Terraform będzie wiedział, w jakim rejestrze szukać dostawcy o nazwie `docker`. Brak tej definicji spowodowałby próbę odnalezienia nieistniejącego dostawcy *hashicorp/docker* (z domyślnego rejestru), co skończyłoby się błędem. Można to zrobić na przykład w pliku o nazwie *terraform.tf* w katalogu modułu. Tym razem nie jest konieczne podawanie wersji. Kod pliku został przedstawiony w listingu 9.

```
terraform {
  required_providers {
    docker = {
      source = "kreuzwerker/docker"
    }
  }
}
```

Listing 9. Provider w katalogu modułu Docker

Listing 9. Provider in Docker module folder



Rys. 6. Końcowa struktura plików Terraform

Fig. 6. Final Terraform file structure

Ostatnim krokiem jest zdefiniowanie ścieżki do modułu Docker w pliku *main.tf* i przekazanie mu wymaganych zmiennych. Końcowa struktura katalogów i plików kodu Terraform podzielonego na moduły została przedstawiona na rys. 6.

Pełny kod pliku *main.tf* po koniecznych modyfikacjach został przedstawiony w listingu 10.

```
terraform {
  required_providers {
    azurerm = {
      source  = "hashicorp/azurerm"
      version = "~> 3.0"
    }
    docker = {
      source  = "kreuzwerker/docker"
      version = "~> 3.0"
    }
  }

  backend "azurerm" {
    resource_group_name  = "tfstate"
    storage_account_name = "tfstateprz"
    container_name       = "tfstate"
    key                  = "terraform.tfstate"
  }
}

provider "azurerm" {
  features {}
  storage_use_azuread = true
}

provider "docker" {
  host = "npipe:////./pipe/docker_engine"

  registry_auth {
    address  = var.acr_server
    username = var.acr_username
    password = var.acr_password
  }
}

module "azure" {
  source = "./modules/azure"

  resource_group_name = var.resource_group_name
  location            = var.location
  acr_server          = var.acr_server
  acr_username        = var.acr_username
  acr_password        = var.acr_password
  container_env_name  = var.container_env_name
  backend_image_tag   = var.backend_image_tag
```

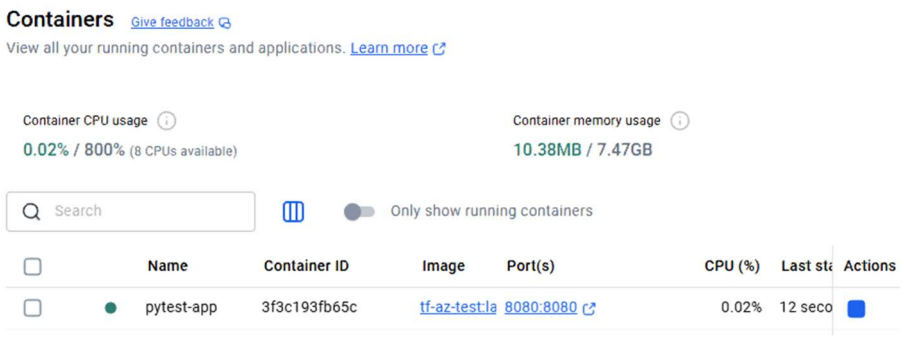
```
  app_name            = var.app_name
  backend_port        = var.backend_port
  text_content        = var.text_content
}

module "docker" {
  source = "./modules/docker"

  backend_image_tag = var.backend_image_tag
  backend_port      = var.backend_port
  text_content      = var.text_content
  app_name          = var.app_name
  acr_server        = var.acr_server
  acr_username      = var.acr_username
  acr_password      = var.acr_password
}
```

Listing 10. Plik *main.tf* po modyfikacjach
Listing 10. Modified *main.tf* file

Przed wykonaniem komendy `terraform init --reconfigure` należy upewnić się, że na komputerze uruchomiony jest demon Docker. Komenda `terraform apply` utworzy (lub w przypadku wdrożenia na Azure – zachowa istniejącą) infrastrukturę i uruchomi kontener, który można zobaczyć w aplikacji Docker Desktop na rys. 7. Jego działanie można przetestować otwierając w przeglądarce adres *http://localhost:8080/* - zwrócony wynik w postaci JSON będzie identyczny, jak w przypadku wdrożenia na chmurze Azure.



Rys. 7. Uruchomiony kontener Dockera
Fig. 7. Deployed Docker container

7. Podsumowanie

W ramach niniejszej pracy zaprezentowano możliwości narzędzia IaC Terraform w zakresie dostarczania infrastruktury i automatyzacji wdrożeń. Ukazano

zalety narzędzia w kontekście bezpieczeństwa, idempotentności i powtarzalności na przykładzie konteneryzacji i wdrożenia prostego serwera HTTP na chmurę Azure i w lokalnym Dockerze. Wraz z dodaniem kolejnego *providera* wykonano także podział kodu Terraform na moduły, umożliwiając oddzielenie wdrożeń na różnych platformach, udowadniając tym samym uniwersalność i skalowalność narzędzia.

Przedstawione rozwiązanie dowodzi, że technologia Terraform rozwiązuje problemy związane z deterministycznością dostarczania infrastruktury dla wdrożeń zarówno prostych, jak i bardziej złożonych systemów i aplikacji. Dzięki zarządzaniu stanem i możliwości zapisu konfiguracji w kodzie zapobiega ona powstawaniu bałaganu w środowiskach chmurowych, a także pozwala znacznie łatwiej zarządzać kosztami i optymalizacją w dynamicznie dostosowującej się do wymagań infrastrukturze.

Literatura

[1] *Azure – dokumentacja*. https://azure.microsoft.com/pl-pl (dostęp: 20.02.2026):

[2] Brikman, Yevgeniy. *Terraform: up and running: writing infrastructure as code*. O'Reilly Media, Inc, 2022.

[3] *Docker - dokumentacja.* https://www.docker.com (dostęp: 04.02.2026).

[4] Howard, Michael. *Terraform – Automating Infrastructure as a Service.* arXiv preprint arXiv:2205.10676 (2022).

[5] *Python – dokumentacja. https://docs.python.org/3/library/http.server.html* (dostęp: 20.06.2026).

[6] *Terraform – dokumentacja*. https://developer.hashicorp.com/terraform (dostęp: 04.02.2026).

PROVISIONING THE INFRASTRUCTURE OF AN APPLICATION USING TERRAFORM ON THE EXAMPLE OF THE AZURE CLOUD AND DOCKER CONTAINERS

Summary

Article describes Terraform as a tool used in IaC (Infrastructure as Code) paradigm for provisioning of the infrastructure required to test or deployment of systems. The author discusses the advantages of the tool starting with implementation of simple HTTP server, its containerization and deployment on Azure platform, gradually utilizing the opportunities given by Terraform, such as multi-provider configuration for both production and testing environments, ending with simultaneous deployment in each of them using TF modules. The paper shows advantages of using Terraform features such as idempotence – the ability to compare desired and current states and perform only changes that are needed to reach the goal, without creating resources unnecessarily.

Keywords: terraform, containerization, infrastructure, deployment, cloud.

Maja Lis
Koło Naukowo-Badawcze Turystyki „Tuptuś”

WPŁYW TURYSTYKI KOSMICZNEJ NA ŚRODOWISKO

Streszczenie

Dynamiczny rozwój turystyki kosmicznej stanowi jeden z najbardziej nieprzewidywalnych i kontrowersyjnych tematów współczesnego społeczeństwa. W konsekwencji prywatyzacji sektora kosmicznego powstają liczne pozarządowe przedsiębiorstwa, takie jak SpaceX, Blue Origin czy Virgin Galactic. Ich działalność ma na celu upowszechnienie możliwości eksploracji kosmosu, oferując loty suborbitalne, jak i orbitalne dla klientów indywidualnych, nie tylko dla naukowców. Celem rozdziału jest analiza następstw tego zjawiska na tle środowiskowym. Obiektem badań jest realizacja komercyjnych lotów kosmicznych w aspekcie technicznym, a zastosowaną metodą – przegląd literatury naukowej oraz raportów branżowych. Wykazano, że emisje spalin rakietowych oddziałują na wyższe warstwy atmosfery, a rosnąca liczba ich startów nasila wymiar problemu odpadów kosmicznych. Wyniki implikują potrzebę połączenia innowacji technologicznych z odpowiedzialnymi regulacjami prawnymi i refleksją etyczną nad kierunkiem rozwoju turystyki kosmicznej.

Słowa kluczowe: turystyka kosmiczna, emisje rakietowe, warstwa ozonowa, śmieci kosmiczne, ochrona środowiska.

1. Wprowadzenie

Człowiek jeszcze do niedawna mógł jedynie śnić o wzbiciu się ponad atmosferę Ziemi, jednak marzenia szybko stały się rzeczywistością. Podbój kosmosu jest stosunkowo świeżym tematem, w końcu dopiero ponad 100 lat temu bracia Wright skonstruowali swój pierwszy samolot wznoszący człowieka w powietrze, otwierając furtkę do rozwoju sektorów lotnictwa, a następnie i kosmonautyki.

Załogowe misje kosmiczne należą do jednych z najbardziej znaczących osiągnieć cywilizacji ludzkiej, jednak dostępne dla bardzo niewielkiej garstki uprzywilejowanych. Początkowo osiągalne jedynie wysoko wykształconym i wyszkolonym astronautom, w dobie postępowego XXI wieku stają się coraz szerzej wykorzystywane również wśród jednostek nie związanych w przeszłości z badawczo-naukowymi podróżami w kosmos. Turystykę kosmiczną tłumaczy się jako:

„jakakolwiek działalność komercyjna oferująca klientom bezpośrednie lub pośrednie doświadczenie związane z podróżami kosmicznymi" [1]. Jest to forma podróży opłacana przez tzw. „turystę kosmicznego" w celach czysto rozrywkowych, poznawczych, bądź dla zaspokojenia satysfakcji. Rozróżnia się kilka rodzajów eksploracji, m.in. suborbitalną, orbitalną, a także międzyplanetarną (wciąż pozostającą w fazie rozwoju). Pierwszą z nich rozumie się jako przekroczenie przez statek kosmiczny na krótki okres czasu granicy kosmosu – umownej linii Kármána [2], rozpościerającej się na ponad 100 km n.p.m. Natomiast nieco bardziej skomplikowane są loty orbitalne polegające na okrążaniu orbity Ziemi i trwające od kilku do nawet kilkunastu dni.

Za pierwszego turystę kosmicznego uznaje się amerykańskiego multimilionera Dennis Tito, który w 2001 r. płacąc Rosjanom koszt w wysokości 20 milionów dolarów [3], odbył podróż na ISS (Międzynarodową Stację Kosmiczną) po czym spędził tam kilka dni. W późniejszych latach, aż do 2009 roku, siedmiu nieastronautów odbyło loty, w tym pierwsza kosmoturystka, Anousheh Ansari. Postęp tej dziedziny umożliwił w dużej mierze rozkwit prywatnych przedsiębiorstw takich jak Space X, Blue Origin oraz Virgin Galactic, które w 2021 roku zapewniły swoje pierwsze komercyjne loty w przestrzeń kosmiczną. W doświadczeniach brali udział m.in. Jeff Bezos – założyciel Amazon i Blue Origin – oraz rekordzista William Shatner, zostając w wieku 90 lat najstarszym człowiekiem, który poleciał w kosmos [4]. Jak podaje oficjalna strona Virgin Galactic, obecnie cena lotu suborbitalnego wynosi 600 tysięcy dolarów, jednak przewiduje się znaczny spadek kosztów wraz z postępem inżynierii kosmicznej i wzrastającej konkurencji na rynku. Popularyzacja turystyki kosmicznej wynikająca z rosnącego zainteresowania szczególnie wśród zamożnej części społeczeństwa - wiąże się z chęcią adaptacji usługowych działalności sektora kosmicznego. Dzięki temu w 2025 roku z oferty skorzystała np. amerykańska piosenkarka Katy Perry oraz Lauren Sánchez [5]. Jednocześnie uwagę skupia długoterminowa wizja Elona Muska o kolonizacji Marsa, zakładająca przewóz około miliona ludzi na czerwoną planetę.

Zaangażowanie prywatnych jednostek w rozwój obszaru kosmicznego budzi szereg pytań i kontrowersji na tle etycznym i środowiskowym. Mimo swych szczytnych celów i rzekomego zapewnienia bezpieczeństwa operacji, nie można zignorować faktu, iż intensyfikacja rozrywkowych lotów doświadczalnych idzie w sprzeczności z globalnymi ideami ochrony środowiska. Coraz większa część społeczeństwa zaczyna poruszać kwestie moralności i zasadniczego celu podróży w przestrzeń pozaziemską – z jakimi konsekwencjami ekologicznymi wiążą się te działania? Jakie są granice eksploracji kosmosu?

Celem niniejszego rozdziału jest dokonanie analizy skutków turystyki kosmicznej na środowisko naturalne. Zastosowana metoda badawcza została przeprowadzona w oparciu o przegląd dostępnej literatury naukowej.

2. Emisje spalin rakietowych

Materiał Rozwój turystyki kosmicznej bezpośrednio pociąga za sobą wzrost liczby startów rakiet, a to równa się zwiększonej ilości emisji niekorzystnych dla atmosfery substancji chemicznych. Czy ludzkość powinna godzić się na rekreacyjne operacje kosmiczne kosztem tak ogromnego ryzyka? Sektor kosmiczny wciąż znajduje się we wczesnej fazie rozkwitu, lecz staje się porównywalnie niebezpieczny co sfera lotnictwa, a zanieczyszczenia uwalniane ze spalania paliw rakietowych znacząco wpływają na zmiany klimatyczne planety, a także na warstwę ozonową. Wynika to głównie z faktu, że zanieczyszczenia te dostają się do wyższych sfer atmosfery – podczas startów m.in. do stratosfery, czyli siedziby ozonosfery na wysokości od około 15 do 50 kilometrów nad powierzchnią Ziemi. Substancje te to przede wszystkim dwutlenek węgla, sadza, tlenki azotu, para wodna oraz chlor czy alumina.

Intensywność wydzielania zależy jednak w dużym stopniu od rodzaju stosowanego paliwa. Paliwa stałe oparte na nadchloranie amonu (NH_4CIO_4) i polimer HTPB produkują istotne ilości związków chloru oraz cząstek tlenku glinu (Al_2O_3), które niekorzystnie wpływają na atmosferę. Z kolei paliwa kerozynowe, między innymi RP-1 wykorzystywane w statkach kosmicznych Falcon 9 od SpaceX oraz rosyjskim Sojuz, generują CO_2 oraz sadzę silnie absorbującą promieniowanie słoneczne. Natomiast używane przez firmę Blue Origin w rakiecie New Shepard paliwa na bazie ciekłego wodoru LH_2 połączonego z tlenem w formie cieczy nie tworzą dwutlenku węgla, jednak prowadzą do emisji pary wodnej bezpośrednio do atmosfery [6]. Choć dzięki ciągłym dążeniom do unowocześnienia rynku inżynierii kosmicznej, topowe firmy turystyki kosmicznej promują dostosowywanie paliw oraz ogólnych potrzeb misji do założeń ekologicznych, nie zmienia to faktu, iż skuteczność tych rozwiązań wciąż pozostaje ciężka do oceny. Mimo to, dzięki dostępnym dla informacji publicznej źródłom, można w stopniu uproszczonym określić ogólny obraz zagrożeń głównych kompozycji potrzebnych do wyniesienia statku ponad granicę atmosfery ziemskiej.

Znaczne zagrożenia w kontekście globalnym wynikają z emisji chloru i tlenków glinu pochodzących z paliw stałych, a także z emisji cząstek sadzy głównie z silników zasilanych naftą (RP-1) [7]. Dodatkowo istotną rolę odgrywają tlenki azotu, które kreują się w wyniku spalania wszystkich paliw w atmosferze, gdyż ta charakteryzuje się wysokim stężeniem azotu atmosferycznego [8]. W połączeniu z parą wodną, tlenki azotu stają się destruktorem ozonu. Reakcje tych substancji chemicznych prowadzą do niszczenia warstwy ozonowej, szczególnie jeśli zsumuje się ich nakład na przestrzeni lat.

Warstwa ozonowa jest wrażliwą osłoną naszej planety, ponieważ odpowiada za pochłanianie szkodliwego promieniowania ultrafioletowego (UV) emitowanego przez słońce. Powstawanie tzw. dziury ozonowej skutkuje natężeniem nowotworów skóry wśród populacji ludzkiej, czy zachorowań na zaćmę, a ponadto skutkuje zmianami w ekosystemach lądowych i problemami w przemyśle rolniczym.

Jej ochronę naznaczono już w 1987 roku, w ramach Protokołu Montrealskiego, co tylko podkreśla jej fundamentalność. W wyniku niepowstrzymanych działań, prognozuje się przy znaczącym wzroście liczby podróży kosmicznych mogłoby dojść do zauważalnych zmian regionalnych w pokrywie ozonowej, zwłaszcza w strefach tropikalnych (utrata o około 1,7%) i polarnych (wzrost o 5-6%)[9]. Dlatego zwracając uwagę na przewidywania co do popularyzacji komercyjnych lotów w kosmos, kumulacja tych elementów poprowadzi do długoterminowych konsekwencji.

Zubożenie warstwy ozonowej idzie w parze ze zmianami wymuszenia radiacyjnego, czyli różnic energii, którą Ziemia otrzymuje od Słońca oraz tą zwracaną w kosmos. Im więcej energii zostaje, tym planeta się ociepla, a im mniej pochłania, tym planeta się ochładza. Dlatego gdy zanika „filtr ochronny" w postaci ozonosfery, powstaje zjawisko zaburzeń klimatycznych. Z uwagi na pochłanianie promieniowania podczerwonego oraz jego emitowanie, również gazy cieplarniane będąc istotnym inicjatorem efektu cieplarnianego mogą powodować coraz to gorsze skutki planetarne. Przewiduje się, że emisja dwutlenków węgla (które kwalifikują się jako jeden z ważniejszych gazów cieplarnianych) przypadająca na tak niewielu pasażerów prywatnego statku kosmicznego będzie od 50 do 100 razy większa niż emisja prawie 3 ton na osobę w przypadku lotów długodystansowych samolotem [10].

Zjawisko to potęguje również kolejny GSG (gaz cieplarniany), jakim jest para wodna. Para wodna pochodząca głównie z paliw opartych na wodorze długo utrzymuje się w stratosferze, która w warunkach naturalnych jest warstwą bardzo suchą. Jej wzbogacenie o H_2O znacznie odbija się na stanie klimatycznym, również sprzyjając powstawaniu polarnych chmur mezosferycznych.

3. Kosmiczne śmieci

Wskutek ponad 60-letniej działalności człowieka w przestrzeni kosmicznej pojawiło się kolejne poważne wyzwanie dla środowiska jakim są kosmiczne śmieci (ang. *Space Debris*). Kosmosu - szczególnie orbity okołoziemskiej, nie można już zdefiniować jako pustej przestrzeni, gdyż systematycznie ulega ona gwałtownym zanieczyszczeniom. Termin kosmiczne śmieci odnosi się do wszystkich nieaktywnych obiektów, ich części lub fragmentów pozostających w orbicie planety lub jej atmosferze. Stanowią one m.in. zużyte elementy rakiet, niesprawne już satelity, elementy pochodzące z kolizji np. w wyniku nieudanej misji, a często są to drobne, pozornie nieistotne kawałki pozostawiane przez kosmonautów [11]. Z roku na rok, ich liczba niebezpiecznie wzrasta – w 2024, Europejska Agencja Kosmiczna (ESA) śledziła ponad 45 300 szczątków o średnicy przekraczającej 10 cm, oraz około 130 milionów cząstek mniejszych niż 1 cm, które obecnie krążą wokół Ziemi [12]. Nie wszystkie elementy można skutecznie monitorować, jednak skala problemu nie ulega wątpliwości. Nawet najmniejsze fragmenty z uwagi na ogromne prędkości potrafią wyrządzić ogromną szkodę.

Kumulacja odpadów orbitalnych niesie ze sobą ryzyko tzw. efektu kaskadowego, znanego jako syndrom Kesslera. Zjawisko to polega na tym, że pojedyncza kolizja dwóch obiektów prowadzi do powstania setek nowych śmieci kosmicznych, które z kolei zwiększają prawdopodobieństwo kolejnych zderzeń. W konsekwencji może dojść do samonapędzającego się procesu, znacząco ograniczającego możliwość bezpiecznego korzystania z przestrzeni kosmicznej. W zgodzie z przyszłymi planami na rozbudowę prywatnych firm kosmicznych SpaceX, czy Virgin Galactic, które zapowiedziało ofertę 400 lotów rocznie, istnieje ryzyko spotęgowania tego zjawiska.

Kosmiczne śmieci stanowią bezpośrednie zagrożenie dla aktywnych satelitów odpowiedzialnych za łączność, nawigację, obserwację Ziemi czy monitorowanie zmian klimatycznych. Uszkodzenie tego typu infrastruktury mogłoby prowadzić do poważnych zakłóceń w funkcjonowaniu współczesnego społeczeństwa, mocno uzależnionego od technologii satelitarnych. Zagrożenie to dotyczy również przyszłych misji załogowych, a szczególnie tych komercyjnych, skoncentrowanych jedynie na zarobku. Statki przewożące pasażerów, zarówno w lotach suborbitalnych, jak i orbitalnych, muszą przecinać warstwy przestrzeni, w których ogniskują śmieci z poprzednich ekspedycji. Jak podaje ESA, niewielkie fragmenty osiągają prędkość nawet 15 km/s [14], co przy takich liczbach może doprowadzić do katastrofalnych uszkodzeń. Przykładowo, 1 centymetrowy kawałek z taką szybkością ma moc porównywalną do eksplodującego granatu [15].

4. Perspektywy zrównoważonej turystyki kosmicznej

W Prywatyzacja lotów poza atmosferę ziemską rodzi szereg wątpliwości czy branża jest w stanie pogodzić ambicje rozwojowe z zasadami zrównoważonego rozwoju. Wraz ze wzrostem zainteresowania dziedziną, coraz więcej przedsiębiorstw deklaruje wprowadzanie „zielonych" paliw, takie jak eksperymentalne biopaliwa syntetyczne, czy ciekły wodór (LH_2), o którym wspomniano już wcześniej – podczas spalania nie uwalnia dwutlenku węgla. Warto jednak pamiętać, że nawet paliwa wodorowe nie są rozwiązaniem na dłuższą metę, gdyż emitują znaczące ilości pary wodnej. Dostosowywanie się do potrzeb środowiskowych nie oznacza zatem kompletnej neutralności klimatycznej, lecz tylko ograniczenie najbardziej szkodliwych substancji chemicznych.

Istotnym krokiem w kierunku ograniczenia zanieczyszczeń środowiskowych jest używanie rakiet wielokrotnego użytku. Takie projekty umożliwiają odzyskiwanie i ponowną eksploatację rakiet, bez konieczności obciążania planety nową produkcją oraz pozostawiania ich członów lewitujących poza atmosferą. Równolegle w ramach ekologistyki kosmicznej powstają koncepcje „kosmicznych śmieciarek". Inicjatywy takie jak pojazdy ClearSpace-1 [16], których zadaniem jest sprzątnięcie pozostałości po działalności człowieka, namierzają, a następnie razem z przechwyconym ładunkiem – deorbitują.

Warto zauważyć, że przemysł kosmiczny wyróżnia się licznymi lukami prawnymi, a szczególnie w zestawieniu z restrykcyjnymi normami lotnictwa cywilnego. Transport lotniczy jest objęty wieloma regulacjami w celu zredukowania emisji gazów cieplarnianych oraz ich neutralizacji, takimi jak CORSIA (Carbon Offsetting and Reduction Scheme for International Aviation). W przypadku statków kosmicznych nie istnieje obecnie limit na skalę globalną, który mógłby zahamować ich oddziaływanie na ekosystem.

W długofalowej perspektywie istotne znaczenie może mieć koncepcja ISRU (In-Situ Resource Utilization), zakładająca wykorzystanie surowców dostępnych na ciałach niebieskich zamiast transportowania ich z Ziemi [17]. Ograniczenie liczby startów przełożyłoby się na mniejszą emisję oraz redukcję obciążenia atmosfery i infrastruktury orbitalnej. Uzupełnieniem tego podejścia jest ISFR (In-Situ Fabrication and Repair), czyli wytwarzanie i naprawa elementów infrastruktury kosmicznej na miejscu, np. z użyciem druku 3D. Rozwiązania te mogą w przyszłości wspierać bardziej zrównoważony model lotów orbitalnych, a więc tych zakładających niekrótki czas trwania.

Fundamentalną rolę odgrywa również wymiar etyczny. Turystyka kosmiczna ma dziś charakter przede wszystkim komercyjny i elitarny, podczas gdy jej środowiskowe konsekwencje mają wymiar globalny. W dobie rosnącej świadomości społecznej pojawia się pytanie o etyczne uzasadnienie rekreacyjnych lotów poza Ziemię, jeśli wiążą się one z dodatkową degradacją przyrodniczą i pogłębianiem obciążenia środowiska orbitalnego. Odpowiedzialny rozwój branży wymaga nie tylko innowacji technologicznych, lecz także przejrzystości danych emisyjnych, międzynarodowych regulacji oraz stosowania zasady ostrożności.

Można zatem stwierdzić, że turystyka kosmiczna podejmuje pierwsze działania w kierunku większej odpowiedzialności środowiskowej, jednak ich skuteczność zależy od równowagi między tempem komercjalizacji a wdrażaniem realnych, gruntownych standardów ekologicznych i etycznych.

5. Podsumowanie

Turystyka kosmiczna, rozwijająca się w szybkim tempie dzięki prywatnym firmom, staje się nowym wyzwaniem dla ochrony środowiska. Rosnąca liczba startów rakiet wiąże ze sobą wprowadzanie do stratosfery substancji, które osłabiają warstwę ozonową i przyczyniają się do lokalnych zmian klimatycznych. Nawet krótkie loty suborbitalne generują emisje mające długotrwały wpływ na atmosferę, a dalsza intensyfikacja działalności może z czasem doprowadzić do poważnych i trudnych do odwrócenia konsekwencji.

Co więcej, dookoła kuli Ziemskiej wciąż nagromadzają się z rosnącą intensywnością śmieci kosmiczne. Nawet jako niepozornie drobne fragmenty poruszające się z ogromną prędkością orbitalną mogą być przyczyną kolizji i inicjować raptowny przyrost liczby obiektów niepożądanych, co zagraża bezpiecznemu i płynnemu funkcjonowaniu satelitów. Jeśli zbagatelizowane, będzie stanowić

również ogromne wyzwanie dla następnych pokoleń, w znacznym stopniu ograniczając swobodne poruszanie się w kosmosie.

Podsumowując, jeśli ludzkość dalej dąży do koegzystencji w zgodzie z naturą i realizacji postulatów zrównoważonego rozwoju, powinniśmy jako społeczeństwo wszcząć działania na rzecz ograniczenia samowoli i globalnie stłumić dynamiczną realizację komercyjnych lotów (sub)orbitalnych. Obecny niedomiar regulacji prawnych, ułatwia tylko te zachowania. Może obecnie turystyka kosmiczna nie jest na tyle krytycznym wyzwaniem, jednak jeśli wziąć pod uwagę narastające zaangażowanie i inwestycje finansowe w przemysł kosmiczny, gdzie dominują zyski ponad ludzką dociekliwość i badania naukowe, rozpatrzenie tej kwestii staje się krytyczne. Dla zachowania obiektywnego bezpieczeństwa klimatycznego i ekologicznego, priorytetowo należy rozpocząć od wdrożenia ścisłych restrykcji prawnych, stosowania rakiet wielokrotnego użytku, paliw o mniejszym wpływie na atmosferę oraz skutecznych systemów aktywnego usuwania odpadów orbitalnych.

Literatura i źródła internetowe

[1] Freeland S., *Fly Me to the Moon: How Will International Law Cope with Commercial Space Tourism*, "Melbourne Journal of International Law" 2010, vol. 11, No. 1, s. 98.

[2] https://innspace.pl/loty-suborbitalne/

[3] Polkowska M., *Space tourism challenges, Review of european and comparative law* 2021, vol. 45, s. 156.

[4] Howell E., *Space tourism took a giant leap in 2021: Here's 10 milestones from the year*, published December 27, 2021.

[5] https://www.blueorigin.com/news/new-shepard-ns-31-mission

[6] U.S. Task Force, Space Generation Advocacy & Policy Platform, *Environmental Impacts of Rocket Fuel*, updated January 2023.

[7] Ross M., Vedda J.A, *The Policy and Science of Rocket Emissions*, April 2018.

[8] W. Hoffmann, *Wpływ emisji rakiet nośnych na atmosferę ziemską*, Warszawa 2022.

[9] Tkaczyński M., *Jak nowe technologie wpływają na zmiany w turystyce kosmicznej. Aspekty technologiczne, etyczne i środowiskowe.*

[10] https://theconversation.com/space-tourism-rockets-emit-100-times-more-co-per-passenger-than-flights-imagine-a-whole-industry-164601

[11] Sala K., *Kosmiczne śmieci jako globalne wyzwanie ekologiczne*, 2023.

[12] Filho W.L., Abubakar I.R., Hunt J., Dinis M.A.P., *Managing space debris: Risks, mitigation measures, and sustainability challenges*, 2025.

[13] European Space Agency, *Hypervelocity impacts and protecting spacecraft*, https://www.esa.int/Space_Safety/Space_Debris/Hypervelocity_impacts_and_protecting_spacecraft

[14] Polkowska M., Toumi M., *Odpowiedzialność za Zanieczyszczenie Przestrzeni Kosmicznej*, 2023.

[15] Vasconcelos J., Gaggi S., Amaral T., Bakouche C., Cotuna A., Friaças A., *Close-Proximity Operations Design, Analysis, and Validation for Non-Cooperative Targets with an Application to the ClearSpace-1 Mission*, 2025.

[16] Baranicka A., *Znaczenie międzyplanetarnych łańcuchów dostaw w zrównoważonej eksploracji kosmosu*, 2021.

ENVIRONMENTAL IMPACT OF SPACE TOURISM

Summary

The dynamic development of space tourism constitutes one of the most unpredictable and controversial topics in contemporary society. As a result of the privatization of the space sector, numerous non-governmental enterprises have emerged, such as SpaceX, Blue Origin, and Virgin Galactic. Their activities aim to popularize access to space exploration by offering both suborbital and orbital flights to individual customers, not only to scientists. The purpose of this article is to analyze the consequences of this phenomenon from an environmental perspective. The subject of the study is the implementation of commercial space flights in technical terms, while the applied method is a review of scientific literature and industry reports. It has been demonstrated that rocket emissions affect the upper layers of the atmosphere, and the growing number of launches intensifies the problem of space debris. The findings imply the need to combine technological innovation with responsible legal regulations and ethical reflection on the direction of space tourism development.

Keywords: space tourism, rocket emissions, ozone layer, space debris, environmental protection.

Julia Widz
Koło Naukowo-Badawcze Turystyki „Tuptuś”

METODY ANALIZY I OCENY EFEKTYWNOŚCI EKONOMICZNEJ DZIAŁALNOŚCI PRZEDSIĘBIORSTW TURYSTYCZNYCH

Streszczenie

Efektywność ekonomiczna przedsiębiorstw turystycznych odgrywa istotną rolę w kształtowaniu ich konkurencyjności i rentowności. W rozdziale omówiono metody analizy i oceny efektywności, uwzględniając aspekty finansowe, techniczne i organizacyjne oraz specyfikę branży turystycznej, w tym sezonowość i zmienność popytu. Dokonano klasyfikacji przedsiębiorstw turystycznych oraz przedstawiono trzy główne modele oceny: analizę wskaźnikową, analizę kosztową i metody porównawcze. Wskazano ich zalety, ograniczenia oraz możliwości zastosowania w praktyce. Wnioski wskazują, że kompleksowa ocena wymaga podejścia wielowymiarowego i stosowania różnych metod analitycznych, co wspiera podejmowanie decyzji strategicznych, poprawę wyników finansowych i zwiększenie konkurencyjności w sektorze turystycznym.

Słowa kluczowe: efektywność ekonomiczna, sektor turystyczny, konkurencyjność, rentowność.

1. Wprowadzenie

Turystyka jest jednym z najbardziej dynamicznie rozwijających się sektorów współczesnej gospodarki i pełni ważną funkcję w generowaniu wzrostu gospodarczego. Funkcjonowanie przedsiębiorstw turystycznych odbywa się jednak w warunkach znacznej niestabilności otoczenia rynkowego, wyraźnej sezonowości popytu oraz silnej konkurencji. Dodatkowo ich działalność silnie uzależniona jest od czynników makroekonomicznych, społecznych i politycznych, co zwiększa poziom niepewności prowadzenia działalności gospodarczej. W takich okolicznościach kluczowego znaczenia nabiera efektywność ekonomiczna, będąca podstawowym miernikiem oceny funkcjonowania przedsiębiorstwa. Jej właściwy pomiar pozwala na racjonalne wykorzystanie zasobów, monitorowanie kosztów oraz ocenę skuteczności przyjętej strategii działania. W sektorze turystycznym efektywność ma charakter złożony, gdyż obejmuje nie tylko wymiar finansowy i kosztowy, lecz także organizacyjny. Celem niniejszego opracowania jest porównanie

i omówienie wybranych metod stosowanych w ocenie efektywności ekonomicznej przedsiębiorstw świadczących usługi w sektorze turystycznym. Dokonano przeglądu literatury przedmiotu w zakresie stosowanych klasyfikacji – zarówno samych podmiotów gospodarczych, jak i wykorzystywanych metod analitycznych. Ponadto przedstawiona została konstrukcja oraz zmienne składowe trzech wybranych modeli stosowanych w praktyce gospodarczej, co pozwala na wskazanie najważniejszych różnic między nimi oraz wynikających z tego korzyści i ograniczeń związanych z wyborem konkretnego narzędzia.

2. Specyfika ekonomiczna przedsiębiorstw turystycznych

2.1. Charakterystyka sektora turystycznego jako sektora usługowego

Sektor turystyczny stanowi jeden z kluczowych obszarów współczesnej gospodarki usługowej, którego znaczenie systematycznie wzrasta wraz z postępującą globalizacją, rozwojem transportu oraz wzrostem mobilności społeczeństwa. Zgodnie z definicją przyjętą przez Światową Organizację Turystyki, turystyka obejmuje aktywność osób podróżujących i przebywających poza swoim stałym miejscem zamieszkania przez okres nieprzekraczający 12 miesięcy, w celach innych niż podjęcie stałej pracy zarobkowej [17, s. 5-7]. Takie ujęcie podkreśla usługowy charakter turystyki, której istotą jest dostarczanie niematerialnych korzyści odpowiadających na potrzeby wypoczynkowe, rekreacyjne, poznawcze, zdrowotne czy biznesowe uczestników ruchu turystycznego [12, s. 17].

Turystyka, jako część sektora usług, wykazuje cechy typowe dla tej kategorii działalności. Jej istota nie sprowadza się do wytworzenia dobra materialnego, lecz do dostarczenia określonych przeżyć i doświadczeń, które kształtują satysfakcję odbiorcy. Proces świadczenia usług jest jednocześnie procesem jej konsumpcji, co oznacza nierozdzielność tych dwóch etapów. Charakterystyczna jest również nietrwałość - niewykorzystane miejsce noclegowe czy wolny fotel w samolocie oznaczają bezpowrotnie utracony potencjalny przychód [5, s. 25-28]. Dodatkowo jakość usług pozostaje zmienna, iż zależy od czynnika ludzkiego, warunków infrastrukturalnych oraz indywidualnych oczekiwań klientów.

Specyfika turystyki jako sektora usługowego przejawia się ponadto w jej złożonym systemowym charakterze. Produkt turystyczny nie ma postaci pojedynczej usługi, lecz stanowi zestaw wzajemnie uzupełniających się świadczeń, obejmujących m.in. transport, noclegi, wyżywienie, program rekreacyjny oraz możliwość korzystania z walorów kulturowych i przyrodniczych [1, s. 61-63]. Oznacza to, że w jego tworzeniu uczestniczy szerokie grono podmiotów działających w powiązanej strukturze rynkowej.

2.2. Klasyfikacja przedsiębiorstw turystycznych

Przedsiębiorstwa działające w sektorze turystycznym można podzielić według rodzaju świadczonych usług, co pozwala na precyzyjne określenie ich

efektywności ekonomicznej. W publikowanych raportach z badań najczęściej spotyka się klasyfikację strukturalno-branżową przedsiębiorstw turystycznych (bez osób samozatrudnionych) na podstawie właściwych kodów PKD, tj. I.55 Działalność związana z zakwaterowaniem, I.56 Działalność związana z wyżywieniem oraz N79 Działalność związana z organizacją turystyki i pośrednictwem turystycznym [6, 9]. Do podmiotów tworzących tzw. bezpośrednią gospodarkę turystyczną zaliczają się te, których działalność koncentruje się na organizowaniu i obsłudze ruchu turystycznego. A. Górka-Chowaniec klasyfikuje je w strukturalnym układzie podmiotowym obejmującym następujące branże:

I. hotelarstwo,
II. gastronomia,
III. transport osobowy,
IV. biura podróży,
V. kultura,
VI. sport i kultura fizyczna [4, s. 17].

W sektorze hotelarskim funkcjonują różnorodne jednostki, w tym obiekty *stricte* hotelowe, pensjonaty, hostele, ośrodki wypoczynkowe oraz apartamenty turystyczne. Ich podstawowym zadaniem jest zapewnienie zakwaterowania oraz świadczenie usług uzupełniających, takich jak centra konferencyjne czy obiekty rekreacyjne. Efektywność ekonomiczna w tym segmencie zależy w dużej mierze od jakości obsługi oraz racjonalnego zarządzania zasobami ludzkimi i infrastrukturą. Gastronomia turystyczna obejmuje restauracje, bary i punkty żywieniowe, a jej wartość wynika z kreowania doświadczenia turystycznego. Rentowność sektora jest kształtowana przez kontrolę kosztów operacyjnych i satysfakcji klientów [3, s. 310, 343]. Transport turystyczny i rekreacyjny obejmuje zarówno przewoźników, jak i podmioty oferujące aktywne formy wypoczynku, np. wypożyczalnie sprzętu sportowego czy organizatorów wycieczek fakultatywnych. Biura podróży i touroperatorzy odpowiadają za organizację i sprzedaż pakietów turystycznych, działając zarówno jako pośrednicy, jak i kompleksowi organizatorzy wyjazdów. Analiza efektywności ich działalności uwzględnia strukturę kosztów operacyjnych, poziom sprzedaży oraz rentowność poszczególnych ofert [3, s. 286-290]. W efekcie klasyfikacja przedsiębiorstw turystycznych według rodzaju świadczonych usług pozwala precyzyjniej dobierać metody oceny efektywności ekonomicznej i wspiera podejmowanie decyzji strategicznych w zarządzaniu sektorem.

2. Teoretyczne podstawy pomiaru efektywności ekonomicznej

2.1. Pojęcie efektywności ekonomicznej

Efektywność ekonomiczna należy do podstawowych kategorii badawczych w ekonomii i odnosi się do oceny wyników działalności gospodarczej w relacji do poniesionych nakładów. W ujęciu encyklopedycznym efektywność (ekonomiczna) definiowana jest jako rezultat działalności gospodarczej określony przez stosunek

osiągniętego wyniku do nakładów [2, s. 35]. Zbliżone ujęcie prezentują L. Pasieczny i J. Więckowski, wskazując, że efektywność ocenia się na podstawie stosunku osiągniętych wyników do nieodzownych nakładów poniesionych w celu ich uzyskania [7, s. 14]. Można również odnaleźć szersze podejście, które podkreśla, że efektywność jest sposobem pomiaru skuteczności i celowości działalności gospodarczej, wyrażającym się porównaniem wartości uzyskanych efektów do nakładów czynników wykorzystanych w procesie ich osiągania [11, s. 21]. Takie ujęcie rozszerza rozumienie efektywności o aspekt racjonalności działania oraz zgodności podejmowanych działań z założonymi celami.

Analiza przytoczonych definicji pozwala wyodrębnić ich wspólne elementy. Wszystkie odwołują się do kategorii rezultatu działalności oraz konieczności odniesienia go do zużytych zasobów. Efektywność ma zatem charakter relacyjny i porównawczy, co umożliwia jej ilościowy pomiar. Wspólnym mianownikiem definicji jest traktowanie efektywności jako miary racjonalności gospodarowania, wyrażającej się dążeniem do maksymalizacji efektów przy danych nakładach lub minimalizacji nakładów przy założonym poziomie efektów.

W ujęciu *stricte* ekonomicznym pojęcie efektywności zostaje doprecyzowane poprzez odniesienie do mechanizmu alokacji ograniczonych zasobów. Jak wskazywali P. A. Samuelson oraz W. D. Nordhaus, efektywności występuje wówczas, gdy nie jest możliwe zwiększenie produkcji jednego dobra bez ograniczenia produkcji innego, co odpowiada koncepcji efektywności Pareto [13, s. 42]. Perspektywa ta rozszerza rozumienie efektywności o aspekt optymalnego wykorzystania ograniczonych zasobów w skali całej gospodarki, podkreślając problem rzadkości zasobów i konieczność dokonywania wyborów.

2.2. Rodzaje efektywności ekonomicznej

Zarówno w teorii, jak i w praktyce gospodarczej, efektywność ekonomiczna przedsiębiorstwa uznawana jest za jeden z kluczowych czynników determinujących jego konkurencyjność oraz trwałość funkcjonowania na rynku. W literaturze wyróżnia się kilka podstawowych wymiarów efektywności, które pozwalają na kompleksową ocenę działalności podmiotu gospodarczego. Do najczęściej wskazywanych rodzajów należą: efektywność finansowa, techniczna oraz organizacyjna [18, s. 45].

Efektywność finansowa odnosi się do zdolności przedsiębiorstwa do generowania zysku oraz utrzymania stabilnej sytuacji finansowej. Jest ona oceniana przede wszystkim przy wykorzystaniu metod wskaźnikowych, obejmujących analizę rentowności, płynności finansowej, zadłużenia czy sprawności działania [14, s. 31]. Wymiar ten odzwierciedla rezultaty działalności przedsiębiorstwa wyrażone w kategoriach pieniężnych.

Efektywność techniczna odnosi się do relacji między nakładami a osiąganymi efektami rzeczowymi i oznacza zdolność przedsiębiorstwa do uzyskiwania maksymalnych rezultatów przy danym poziomie zasobów lub minimalizacji nakładów

przy założonym poziomie produkcji usług. W przeciwieństwie do analizy finansowej, która koncentruje się na wynikach pieniężnych, pomiar efektywności technicznej opiera się na mikroekonomicznym ujęciu procesu produkcji, uwzględniającym zależność między czynnikami wytwórczymi a wielkością uzyskiwanych efektów [18, s. 49-51].

W teorii organizacji i zarządzania jednym z kluczowych zagadnień jest efektywność organizacyjna. Definiuje się ją jako zdolność przedsiębiorstwa do bieżącego i strategicznego przystosowania się do zmian w otoczeniu oraz produktywnego i oszczędnego wykorzystania posiadanych zasobów dla realizacji przyjętej struktury celów [18, s. 61]. Pomiar efektywności organizacyjnej polega na określeniu efektywności działania organizacji na tle innych podmiotów tego sektora przez porównanie ich wyników między sobą.

Tak ujęta klasyfikacja pozwala na wielowymiarowe spojrzenie na efektywność ekonomiczną przedsiębiorstw turystycznych i stanowi punkt wyjścia do omówienia metod jej pomiaru przedstawionych w rozdziale trzecim.

2.3. Czynniki wpływające na efektywność ekonomiczną w turystyce

Efektywność ekonomiczna przedsiębiorstw turystycznych jest wypadkową działania wielu czynników wewnętrznych i zewnętrznych, które kształtują zarówno wyniki finansowe, jak i sprawność operacyjną. Wśród czynników wewnętrznych kluczowe znaczenie mają kwalifikacje i doświadczenie personelu, który bezpośrednio wpływa na jakość świadczonych usług i poziom satysfakcji turystów. Istotne jest również zarządzanie zasobami i procesami operacyjnymi, w tym koordynacja działań w obszarach hotelarstwa, gastronomii i rekreacji, co pozwala na optymalizację kosztów i elastyczne reagowanie na zmiany w popycie [4, s. 35-37]. Stabilna sytuacja finansowa oraz dostęp do nowoczesnych technologii wspierają podejmowanie decyzji strategicznych i zwiększają efektywność procesów wewnętrznych. Szczególne znaczenie w turystyce ma także zdolność szybkiego dostosowania się do sezonowości oraz zmieniających się preferencji konsumentów.

W zakresie czynników zewnętrznych największy wpływ na efektywność mają uwarunkowania rynkowe i gospodarcze, w tym poziom konkurencji, kursy walut oraz zmiany w preferencjach turystów. Stabilność polityczna i prawna kraju, rozwój infrastruktury turystycznej oraz wsparcie ze strony instytucji lokalnych umożliwiają przedsiębiorstwom lepszą organizację usług i zwiększenie przychodów. Równie istotne są uwarunkowania społeczne i środowiskowe, zmiany demograficzne, wzrost zamożności społeczeństwa czy dostępność atrakcji turystycznych kształtują popyt i determinują strukturę oferty [4, s. 35-37]. Świadome uwzględnianie tych czynników w procesach decyzyjnych pozwala przedsiębiorstwom turystycznym efektywnie zarządzać zasobami, optymalizować koszty operacyjne oraz zwiększać konkurencyjność i rentowność działalności.

3. Wybrane modele analizy efektywności ekonomicznej przedsiębiorstw turystycznych

3.1. Analiza wskaźnikowa

Analiza wskaźnikowa należy do podstawowych i najczęściej wykorzystywanych metod oceny efektywności ekonomicznej przedsiębiorstw, zwłaszcza w obszarze finansowym. Istota tej metody polega na wyznaczaniu oraz interpretowaniu określonych relacji liczbowych pomiędzy wybranymi kategoriami ekonomicznymi. Pozwala to na syntetyczną ocenę sytuacji majątkowej, finansowej i wynikowej podmiotu [8, s. 190]. Odpowiednio dobrane wskaźniki umożliwiają uchwycenie zależności między elementami sprawozdań finansowych oraz ocenę stopnia racjonalności gospodarowania zasobami.

Podstawą tej analizy są informacje zawarte w sprawozdaniach finansowych, przede wszystkim w bilansie, rachunku zysków i strat oraz rachunku przepływów pieniężnych. W odniesieniu do przedsiębiorstw turystycznych zakres wykorzystywanych danych może być szerszy i obejmować również raporty branżowe, statystyki publikowane przez Główny Urząd Statystyczny, dane udostępniane przez organizacje turystyczne, a także wewnętrzne raporty operacyjne jednostki. Pozwala to na pełniejsze ujęcie specyfiki działalności usługowej oraz sezonowego charakteru popytu.

W praktyce analiza wskaźnikowa obejmuje kilka zasadniczych grup mierników. Do najistotniejszych zalicza się wskaźniki rentowności, odzwierciedlające zdolność do generowania zysku w relacji do sprzedaży, aktywów lub kapitału własnego, wskaźniki płynności określające możliwość terminowego regulowania zobowiązań krótkoterminowych, wskaźniki zadłużenia obrazujące strukturę finansowania działalności oraz wskaźniki sprawności działania, które informują o efektywności wykorzystania majątku i kapitału obrotowego [10, s. 95].

W sektorze turystycznym metoda ta znajduje szerokie zastosowanie w ocenie kondycji finansowej hoteli, biur podróży, przedsiębiorstw gastronomicznych czy podmiotów transportowych. Umożliwia analizę zmian wyników w czasie, co ma szczególne znaczenie w warunkach sezonowości, a także porównanie rezultatów z innymi podmiotami funkcjonującymi w branży. W efekcie stanowi narzędzie identyfikacji obszarów wymagających usprawnień oraz oceny skuteczności realizowanej strategii zarządczej.

Analiza wskaźnikowa cechuje się szeregiem istotnych zalet. Wyróżnia się przede wszystkim prostotą zastosowania, powszechnością wykorzystania w różnych sektorach gospodarki oraz możliwością porównywania wyników w czasie i względem innych podmiotów. Jej zaletą jest także uniwersalność wskaźników oraz możliwość szybkiej identyfikacji kluczowych obszarów wymagających poprawy, co ma istotne znaczenie w przedsiębiorstwach turystycznych funkcjonujących w warunkach sezonowości i silnej konkurencji. Do najważniejszych ograniczeń należy natomiast jej statyczny charakter, oparcie na danych historycznych

oraz brak bezpośredniego wskazania przyczyn niekorzystnych zjawisk. Wyniki mogą być również zniekształcone przez różnice w polityce rachunkowości lub ograniczoną porównywalność danych, dlatego analiza wskaźnikowa powinna być uzupełniana innymi metodami oceny efektywności [18, s. 48].

3.2. Analiza efektywności kosztowej

Efektywność kosztowa odnosi się do optymalizacji relacji między poniesionymi kosztami a osiąganymi korzyściami z działalności. Zasadniczym celem tej metody jest określenie, czy przedsiębiorstwo realizuje swoje zadania w sposób ekonomiczny, tzn. przy możliwie racjonalnym poziomie kosztów, bez pogorszenia jakości oferowanych usług. Podejście to umożliwia bardziej szczegółową ocenę działalności, pozwalając na wnikliwą analizę struktury kosztów oraz mechanizmów ich kształtowania. W praktyce analiza ta rozpoczyna się od identyfikacji oraz podziału kosztów (np. na stałe i zmienne, bezpośrednie i pośrednie), a następnie odniesienia ich do określonych efektów działalności, takich jak wielkość sprzedaży, liczba klientów czy skala świadczonych usług [16, s. 140-142].

Podstawą przeprowadzenia analizy są dane pochodzące z ewidencji księgowej, rachunku zysków i strat, kalkulacji kosztowych oraz raportów controllingowych. W przedsiębiorstwach turystycznych szczególnego znaczenia nabierają także dane operacyjne, takie jak poziom obłożenia w hotelach, liczba sprzedanych imprez turystycznych, liczba pasażerów czy średni przychód przypadający na jednego klienta.

W sektorze turystyki analiza efektywności kosztowej ma szczególne znaczenie ze względu na wysoki udział kosztów stałych oraz sezonowy charakter popytu. W hotelarstwie może dotyczyć kosztu utrzymania pokoju czy obsługi jednego gościa, w biurach podróży będzie to opłacalność konkretnych pakietów turystycznych, natomiast w transporcie koszt realizacji pojedynczego kursu lub przewozu jednego pasażera. Dzięki temu możliwe jest szybkie wykrycie obszarów generujących nadmierne wydatki i podjęcie działań optymalizacyjnych [15, s. 127-146].

Konieczne jest podkreślenie, że wyniki tej analizy są silnie uzależnione od przyjętego systemu rachunku kosztów oraz sposobu rozliczania kosztów pośrednich. W przedsiębiorstwach o zróżnicowanej ofercie trudne bywa precyzyjne przypisanie kosztów do poszczególnych usług. Ponadto koncentracja wyłącznie na redukcji kosztów może prowadzić do pomijania aspektów jakościowych, takich jak standard obsługi czy satysfakcja klientów. W branży turystycznej, gdzie doświadczenie klienta odgrywa kluczową rolę, nadmierne ograniczanie wydatków może w dłuższej perspektywie osłabić konkurencyjność przedsiębiorstwa.

3.3. Metody porównawcze

Metody porównawcze, zwane benchmarkingiem, stanowią narzędzie oceny efektywności ekonomicznej przedsiębiorstwa poprzez zestawienie jego wyników z wynikami innych podmiotów działających w tym samym sektorze lub z ustalonymi standardami rynkowymi [19, s. 201-212]. Celem tej analizy jest identyfikacja

najlepszych praktyk, określenie pozycji jednostki na tle konkurencji oraz wskazanie obszarów wymagających poprawy, co pozwala na wdrożenie działań zwiększających efektywność organizacyjną.

Tabela 1. Porównanie wybranych modeli analizy efektywności ekonomicznej

Table 1. Comparison of selected models of economic efficiency analysis

Cecha / model	Analiza wskaźnikowa	Analiza efektywności kosztowej	Metody porównawcze
Główna idea	Ocena na podstawie zestawu wskaźników finansowych	Ocena racjonalności ponoszonych kosztów w relacji do efektów działalności i jakości oferowanych usług	Porównanie wyników jednostki z innymi podmiotami w sektorze lub z ustalonymi standardami rynkowymi
Typowe rodzaje	Wskaźniki płynności, zadłużenia, rentowności, sprawności działania	Koszty całkowite, koszty jednostkowe, koszty stałe i zmienne w relacji do efektów działalności (np. sprzedaży, liczby klientów)	Benchmarking, rankingi, ocena atrakcyjności/konkurencyjności
Główne zalety	Prosta, powszechność w różnych sektorach, oparta na standardowym sprawozdaniu finansowym; szybka ocena kondycji	Szczegółowa analiza struktury kosztów, identyfikacja źródeł nadmiernych wydatków	Umożliwia ocenę pozycji rynkowej i względnej efektywności wobec konkurentów
Główne wady	Duża wrażliwość na przyjęte progi i interpretację, słabo uwzględnia jakość i czynniki niefinansowe	Uzależnienie wyników od systemu rachunku kosztów i sposobu rozliczania kosztów pośrednich	Wyniki zależą od wyboru grupy porównawczej i wskaźników, trudność uwzględnienia kontekstu rynkowego i specyfiki regionu
Trudność pozyskania danych	Niska – głównie dane ze sprawozdania finansowego (bilans, RZiS)	Średnia/wysoka – potrzebne szczegółowe dane kosztowe, często z rachunku kosztów działań	Średnia – wymaga danych własnych i danych konkurentów/rynku
Sposób pozyskiwania zmiennych	Z systemu księgowego, sprawozdań finansowych, ewentualnie z raportów branżowych	Z rachunku kosztów, ewidencji działań, systemów controllingu	Z baz statystycznych, raportów rynkowych, badań ankietowych, sprawozdań firm
Kluczowe korzyści zastosowania	Identyfikacja mocnych i słabych stron finansowych, monitorowanie zmian w czasie	Optymalizacja struktury kosztów, wzrost rentowności przy danym poziomie przychodów	Ustalenie pozycji konkurencyjnej, identyfikacja dobrych praktyk do naśladowania

Źródło: opracowanie własne na podstawie R. Polaczek, *Analiza wskaźnikowa jako podstawowa metoda oceny efektywności przedsiębiorstwa* [w:] *Współczesne problemy analizy ekonomicznej przedsiębiorstwa,* Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 9 (1209), Wrocław 2008, s. 190, B. Pomykalska, P. Pomykalski, *Analiza finansowa przedsiębiorstwa,* Wydawnictwo Naukowe PWN, Warszawa 2017, s. 95, R. R. Stolarski, *Efektywność kosztowa jako czynnik oddziaływania na rentowność przedsiębiorstwa turystycznego* [w:] *Turystyka i Rozwój Regionalny*, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, Warszawa 2025, s. 140-142, M. Sołtysik, W. Fedyk, B. Walas, *Metody i narzędzia oceny skuteczności wdrażania strategii rozwoju turystyki oraz produktów turystycznych. Zróżnicowania regionalne i lokalne* [w:] *Ekonomiczne Problemy Turystyki*, Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 3 (27), Szczecin 2014, s. 127-146, B. Ziębicki, *Metodologiczne problemy oceny efektywności organizacyjnej*, Studia i Prace Kolegium Zarządzania i Finansów, Szkoła Główna Handlowa, z. 115, 2012, s. 201-212.

Proces benchmarkingu obejmuje wybór odpowiednich wskaźników efektywności, takich jak koszty jednostkowe, obłożenie miejsc noclegowych czy rentowność poszczególnych usług. Wyniki badanej jednostki porównuje się następnie z wynikami innych przedsiębiorstw lub wartościami referencyjnymi, aby określić poziom efektywności i zidentyfikować luki w funkcjonowaniu organizacji.

W sektorze turystycznym metody porównawcze znajdują zastosowanie w hotelarstwie, biurach podróży i przedsiębiorstwach transportowych, umożliwiając porównanie kosztów utrzymania pokoi, przychodów z usług, wydajności personelu czy jakości obsługi klientów między różnymi podmiotami. Dane do analizy pochodzą zarówno z systemów księgowych i raportów operacyjnych badanej jednostki, jak i z publicznych zestawień branżowych, raportów stowarzyszeń turystycznych czy platform analitycznych. Zastosowanie metod porównawczych pozwala przedsiębiorstwom turystycznym identyfikować najlepsze praktyki w branży, wdrażać działania doskonalące, optymalizować koszty oraz podnosić jakość usług.

Po omówieniu trzech podstawowych modeli analizy efektywności ekonomicznej przedsiębiorstw turystycznych warto zestawić ich kluczowe cechy. Umożliwia to szybką ocenę, które podejście najlepiej odpowiada specyfice danej jednostki turystycznej, a także ułatwia wybór odpowiednich narzędzi analitycznych w praktyce zarządczej. W tabeli 1. przedstawiono porównanie głównych idei, typowych wskaźników, zalet oraz ograniczeń poszczególnych metod. Analiza wskaźnikowa koncentruje się na ocenie kondycji finansowej poprzez wskaźniki płynności, zadłużenia i rentowności. Analiza efektywności kosztowej pozwala ocenić racjonalność ponoszonych kosztów oraz ich optymalizację. Natomiast metody porównawcze, takie jak benchmarking czy rankingi, umożliwiają określenie pozycji rynkowej przedsiębiorstwa w stosunku do konkurencji i przyjętych standardów. Tabela uwzględnia również typowe wskaźniki, mocne i słabe strony metod, trudności w pozyskiwaniu danych oraz korzyści wynikające z ich stosowania, co pozwala na kompleksowe porównanie narzędzi analitycznych.

Literatura

[1] Altkorn J.: Marketing w turystyce, Wydawnictwo Naukowe PWN, Warszawa 1995.

[2] Encyklopedia popularna, Wydawnictwo Naukowe PWN, t. 3, Warszawa 1997.

[3] Gaworecki W.: Turystyka, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003.

[4] Górka-Chowaniec A.: Zarządzanie przedsiębiorstwem turystycznym w warunkach sieci. Część I: Kompendium wiedzy teoretycznej, Wydawnictwo Akademii Wychowania Fizycznego w Katowicach, Katowice 2017.

[5] Panasiuk A.: Marketing usług turystycznych, Wydawnictwo Naukowe PWN, Warszawa 2007.

[6] PARP: Branżowy Bilans Kapitału Ludzkiego II edycja – Branża turystyczna, 2022.

[7] Pasieczny L., Więckowski J.: Ekonomika i analiza działalności przedsiębiorstwa, Polskie Wydawnictwo Ekonomiczne, Warszawa 1987.

[8] Polaczek R.: Analiza wskaźnikowa jako podstawowa metoda oceny efektywności przedsiębiorstwa [w:] Współczesne problemy analizy ekonomicznej przedsiębiorstwa, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 9 (1209), Wrocław 2008.

[9] Polski Instytut Ekonomiczny: Branża turystyczna w Polsce, 2020.

[10] Pomykalska B., Pomykalski P.: Analiza finansowa przedsiębiorstwa, Wydawnictwo Naukowe PWN, Warszawa 2017.

[11] Popularna Encyklopedia Powszechna, t. 5, Oficyna Wydawnicza „FORGA", Kraków 1995.

[12] Rapacz A.: Nowe wyzwania i możliwości rozwoju turystyki, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2021.

[13] Samuelson P. A., Nordhaus W.D.: Ekonomia. Tom I, Wydawnictwo Naukowe PWN, Warszawa 1998.

[14] Siemińska E.: Metody pomiaru i oceny kondycji finansowej przedsiębiorstwa, Wyd. TNOiK „Dom Organizatora", Toruń 2002.

[15] Sołtysik M., Fedyk W., Walas B.: Metody i narzędzia oceny skuteczności wdrażania strategii rozwoju turystyki oraz produktów turystycznych. Zróżnicowania regionalne i lokalne [w:] Ekonomiczne Problemy Turystyki, Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 3 (27), Szczecin 2014.

[16] Stolarski R. R.: Efektywność kosztowa jako czynnik oddziaływania na rentowność przedsiębiorstwa turystycznego [w:] Turystyka i Rozwój Regionalny, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, Warszawa 2025.

[17] World Tourism Organization: Terminologia turystyczna. Zalecenia WTO, Instytut Turystyki, Warszawa 1995.

[18] Zielińska-Chmielewska A.: Ocena efektywności ekonomicznej przedsiębiorstw mięsnych w Polsce, Uniwersytet Ekonomiczny w Poznaniu, Poznań 2020.

[19] Ziębicki B.: Metodologiczne problemy oceny efektywności organizacyjnej, Studia i Prace Kolegium Zarządzania i Finansów, Szkoła Główna Handlowa, z. 115, 2012.

METHODS OF ANALYSIS AND EVALUATION OF THE ECONOMIC EFFICIENCY OF TOURISM ENTERPRISES

Summary

The economic efficiency of tourism enterprises plays an important role in shaping their competitiveness and profitability. The article discusses methods of analysis and evaluation of effectiveness, taking into account financial, technical and organizational aspects as well as the specificity of the tourism industry, including seasonality and variability of demand. Tourism enterprises are classified and three main evaluation models are presented: ratio analysis, cost analysis and comparative methods. Their advantages, limitations and possibilities of application in practice are indicated. The conclusions indicate that a comprehensive assessment requires a multidimensional approach and the use of different analytical methods, which supports strategic decision-making, improved financial performance and increased competitiveness in the tourism sector.

Keywords: economic efficiency, tourism sector, competitiveness, profitability.

Karolina Czapka
Koło Naukowo-Badawcze Turystyki „Tuptuś”

CZYNNIKI WPŁYWAJĄCE NA WYBÓR UBEZPIECZENIA TURYSTYCZNEGO PRZEZ OSOBY PODRÓŻUJĄCE

Streszczenie

Ubezpieczenia turystyczne pełnią istotną rolę w ochronie podróżnych przed nieprzewidzianymi kosztami oraz ryzykiem związanym z wyjazdami, zapewniając zarówno bezpieczeństwo finansowe, jak i komfort psychiczny podczas podróży. Badanie pozwoliło zidentyfikować czynniki wpływające na wybór ubezpieczenia turystycznego oraz rolę posiadanej polisy w poczuciu bezpieczeństwa respondentów. Najważniejszym kryterium przy podejmowaniu decyzji okazał się zakres ochrony i jego wysokość, który zdecydowanie przewyższał inne aspekty, takie jak cena, renoma firmy czy opinie znajomych. Większość uczestników badania docenia funkcję ochronną ubezpieczenia i jest skłonna dopłacić za rozszerzony zakres ochrony, przy czym cena wciąż pozostaje istotnym czynnikiem w procesie decyzyjnym. Wzrost znaczenia odnotowano również w przypadku wygody zakupu online, co odzwierciedla zmieniające się oczekiwania podróżnych wobec dostępności i prostoty usług. Respondenci najczęściej pozyskują informacje z Internetu lub od rodziny i znajomych, natomiast regularne zapoznawanie się z warunkami polis wciąż nie jest powszechne. Analiza odpowiedzi otwartych wykazała, że podróżni oczekują od polis nie tylko praktycznej ochrony finansowej, ale również wsparcia w różnych sytuacjach kryzysowych, przejrzystych i jasnych warunków umowy oraz adekwatnej ceny do oferowanego zakresu ochrony. Choć większość respondentów nie korzystała faktycznie z ubezpieczenia, jego posiadanie znacząco podnosi poczucie bezpieczeństwa, pełniąc funkcję zarówno praktyczną, jak i psychologiczną. Ankieta przeprowadzona wśród 100 respondentów, z czego 64 faktycznych nabywców polis, potwierdziła, że decyzje zakupowe mają charakter racjonalny, opierając się głównie na analizie zakresu ochrony i kosztów przy rosnącym znaczeniu dostępności i wygody usług online.

Słowa kluczowe: ubezpieczenie, turystyka, podróże, zakres ochrony, zabezpieczenie wyjazdów.

1. Wprowadzenie

Współczesna turystyka, zarówno krajowa, jak i zagraniczna, charakteryzuje się wysoką mobilnością oraz dynamicznie rosnącą liczbą podróżnych. Wraz ze wzrostem liczby wyjazdów rośnie również potrzeba ochrony finansowej i bezpieczeństwa podczas podróży. Ubezpieczenia turystyczne pełnią w tym kontekście istotną funkcję, zabezpieczając podróżnych przed negatywnymi skutkami zdarzeń losowych, takich jak wypadki, nagłe zachorowania czy utrata bagażu. Pomimo powszechnej dostępności polis turystycznych, decyzje o ich zakupie są wciąż zróżnicowane, a wiele osób podejmuje je w oparciu o subiektywne odczucie ryzyka oraz indywidualne preferencje.

Celem niniejszego rozdziału jest identyfikacja czynników wpływających na wybór ubezpieczenia turystycznego przez osoby podróżujące oraz analiza znaczenia tych czynników w procesie podejmowania decyzji nabywczych. W pracy uwzględniono zarówno aspekty teoretyczne dotyczące zachowań konsumentów i charakterystyki rynku ubezpieczeń turystycznych w Polsce, jak i badania empiryczne oparte na ankiecie przeprowadzonej wśród podróżnych. Przeprowadzone analizy pozwalają zrozumieć, które elementy polisy, takie jak zakres ochrony, cena czy opinie innych użytkowników, mają największe znaczenie przy podejmowaniu decyzji zakupowej, a także w jakim stopniu ubezpieczenie zwiększa poczucie bezpieczeństwa podróżnych.

2. Istota i znaczenie ubezpieczeń turystycznych

Ubezpieczenia turystyczne stanowią istotny element zabezpieczenia finansowego osób podróżujących, zarówno w kraju, jak i za granicą. Ich głównym celem jest ochrona ubezpieczonego przed negatywnymi skutkami zdarzeń losowych, które mogą wystąpić w trakcie podróży [3, s. 48]. Do takich zdarzeń zalicza się m.in. nagłe zachorowania, wypadki, utratę bagażu, odpowiedzialność cywilną wobec osób trzecich czy konieczność skorzystania z pomocy assistance [6, s. 162] [12, s. 92].

Jest to dobrowolna forma ochrony, zawierana na określony czas, najczęściej na czas trwania podróży [3, s. 49]. W ramach jednej polisy możliwe jest objęcie ochroną kilku różnych ryzyk, co czyni ten rodzaj ubezpieczenia elastycznym i dopasowanym do indywidualnych potrzeb podróżnych [7, s. 241]. Zakres ochrony zależy od wybranego wariantu oraz warunków określonych w Ogólnych Warunkach Ubezpieczenia (OWU).

Rosnąca liczba podróży zagranicznych oraz wysokie koszty leczenia poza krajem zwiększają znaczenie ubezpieczeń turystycznych. Jak wynika z badań ARC Rynek i Opinia zrealizowanego przy współpracy z Europa Ubezpieczenia, polisę wykupuje 79% Polaków podróżujących turystycznie, z czego 33% robi to przy

każdym wyjeździe, a 46% w zależności od długości lub kierunku podróży. Ubezpieczenie zapewnia nie tylko ochronę finansową, ale także poczucie bezpieczeństwa podczas wyjazdu [14].

Zakup ubezpieczenia turystycznego jest procesem, który obejmuje identyfikację własnych potrzeb, ocenę ryzyka związanego z podróżą oraz porównanie dostępnych ofert, uwzględniając zakres ochrony, cenę, wyłączenia i wiarygodność ubezpieczyciela [11, s. 311-329] [13].

2.1. Rodzaje i zakres ochrony ubezpieczeń turystycznych

Ubezpieczenia turystyczne mogą obejmować różne formy ochrony, dostosowane do potrzeb oraz charakteru podróży. Najczęściej spotykane elementy polisy turystycznej to ubezpieczenie kosztów leczenia, następstw nieszczęśliwych wypadków, odpowiedzialność cywilna, assistance oraz ubezpieczenie bagażu [10, s. 205] [16]. Rodzaje wraz z krótkim wyjaśnieniem zakresu ochrony zostały przedstawione w tabeli 1.

Tabela 1. Elementy ubezpieczenia turystycznego

Table 1. Elements of travel insurance

Rodzaj ubezpieczenia	Zakres ochrony
Koszty leczenia zagranicą (KLZ)	Pokrywa koszty wizyt lekarskich, hospitalizacji, badań, zabiegów i leków w razie nagłego zachorowania lub wypadku za granicą.
Następstw nieszczęśliwych wypadków (NNW)	Zapewnia wypłatę odszkodowania w przypadku trwałego uszczerbku na zdrowiu lub śmierci w wyniku wypadku.
Ochrona cywilna w życiu prywatnym (OC)	Chroni przed finansowymi skutkami szkód wyrządzonych osobom trzecim lub ich mieniu.
Assistance	Obejmuje pomoc organizacyjną i informacyjną, m.in. transport medyczny, pomoc tłumacza i wsparcie w sytuacjach kryzysowych.
Ubezpieczenie bagażu	Chroni przed stratami wynikającymi z utraty, kradzieży lub zniszczenia bagażu.

Źródło: opracowanie własne na podstawie prezentacji „Bezpieczne wakacje- ubezpieczenia turystyczne w pigułce" Mec. Aleksandra Daszewskiego, Rzecznik finansowy [17] (dostęp: 12.02.2026).

Zakres ochrony zależy od wybranego wariantu ubezpieczenia oraz zapisów zawartych w Ogólnych Warunkach Ubezpieczenia. Może być on bardzo szeroki i obejmować zarówno zabezpieczenie zdrowotne, finansowe, jak i organizacyjne. Odpowiedni dobór poszczególnych elementów polisy pozwala na skuteczną ochronę przed ryzykiem związanym z podróżowaniem oraz zwiększenie poczucia bezpieczeństwa podróżnych [17].

2.2. Rynek ubezpieczeń turystycznych w Polsce

Rynek ubezpieczeń turystycznych w Polsce rośnie wraz z liczbą wyjazdów i mobilnością społeczeństwa. Zakłady ubezpieczeń oferują zróżnicowane produkty pod względem zakresu ochrony, sum ubezpieczenia oraz dodatkowych opcji, takich jak assistance czy sportów wysokiego ryzyka. Oferty są coraz częściej dostosowywane do rodzaju podróży, jej długości oraz indywidualnych potrzeb klientów [15].

Istotnym trendem na tym rynku jest dynamiczny rozwój sprzedaży online, który umożliwia łatwe porównywanie ofert i szybki zakup polisy. Zwiększona konkurencja oraz rosnąca świadomość ryzyka związanego z podróżowaniem sprawiają, że coraz więcej osób decyduje się na wykupienie ubezpieczenia, szczególnie przed wyjazdami zagranicznymi. W efekcie rynek ten charakteryzuje się wysoką dostępnością produktów oraz stopniowym wzrostem znaczenia ubezpieczeń turystycznych planowaniu podróży [2, s. 25-39].

2.3. Modele zachowań konsumenta

Funkcjonowanie rynku ubezpieczeń turystycznych oraz zróżnicowanie dostępnych ofert sprawiają, że proces wyboru polisy ma charakter wieloetapowy i wymaga od konsumenta analizy wielu czynników. W tym kontekście użyteczne jest odwołanie się do modelu Engel-Kollat-Blackwell (EKB), który opisuje decyzję zakupową jako ciąg następujących po sobie etapów: rozpoznanie potrzeby, poszukiwanie informacji, ocenę alternatyw, podjęcie decyzji oraz ocenę pozakupową [4, s. 183-190] [8, s. 181-189]. W przypadku ubezpieczeń turystycznych etapy te przejawiają się m.in. w analizie źródeł informacji, porównywaniu zakresu ochrony i ceny oraz w ocenie wpływu posiadania polisy na poczucie bezpieczeństwa. Model ten pozwala zatem interpretować wybór ubezpieczenia jako rezultat świadomego procesu decyzyjnego, opartego na ocenie ryzyka i dostępnych wariantów ochrony [9].

Uzupełnieniem tej perspektywy jest teoria planowanego zachowania (Theory of Planned Behavior - TPB), zgodnie z którą intencja podjęcia określonego działania zależy od postawy wobec danego zachowania, norm społecznych oraz postrzeganej kontroli nad jego realizacją [1, s. 314-324]. W odniesieniu do ubezpieczeń turystycznych postawa wiąże się z przekonaniem o ich funkcji ochronnej, normy społeczne mogą przejawiać się w wpływie opinii rodziny i znajomych, natomiast postrzegana kontrola dotyczy m.in. łatwości i dostępności zakupu online. Takie ujęcie pozwala lepiej wyjaśnić, w jaki sposób czynniki ekonomiczne, społeczne i psychologiczne współkształtują decyzje konsumentów na rynku ubezpieczeń turystycznych [9].

3. Ocena podróżujących pod kątem ich decyzji wyboru ubezpieczenia turystycznego

Przeprowadzone badania empiryczne oparto na rozważaniach teoretycznych dotyczących zachowań konsumentów oraz determinant podejmowania decyzji zakupowych na rynku ubezpieczeń. W tym celu zrealizowano ankietę wśród osób podróżujących. Uzyskane wyniki umożliwiają analizę postaw respondentów oraz ich podejścia do kwestii zabezpieczenia ubezpieczeniowego podczas podróży.

Badanie zastosowało przeprowadzone z wykorzystaniem ankiety internetowej, co umożliwiło szybkie dotarcie do szerokiej grupy respondentów przy ograniczeniu kosztów i czasu realizacji badania. Wybór tej metody był uzasadniony specyfiką badanej populacji, osób aktywnych podróżniczo, które w dużej mierze korzystają z internetu zarówno przy organizacji wyjazdów, jak i zakupie ubezpieczeń. Ankieta online zapewniła anonimowość odpowiedzi, standaryzację pytań oraz porównywalność danych, a także pozwoliła ograniczyć wpływ badacza na udzielane odpowiedzi.

Zastosowano dobór celowy próby, obejmujący osoby podróżujące deklarujące wykupienie ubezpieczenia turystycznego. Kwestionariusz rozpowszechniono za pośrednictwem mediów społecznościowych i komunikatorów internetowych. Badanie przeprowadzono w okresie od stycznia do lutego 2026 roku. W badaniu uczestniczyło 100 osób, z czego 64 spełniły kryterium wykupienia ubezpieczenia i zostały uwzględnione w analizie.

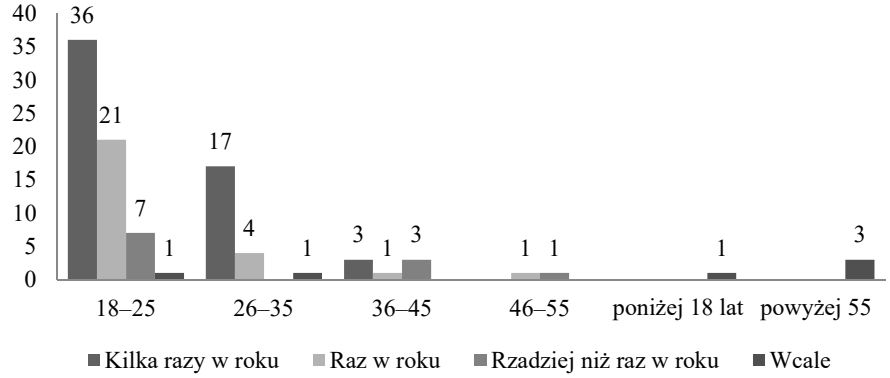
3.1. Charakterystyka próby badawczej

W badaniu wzięło udział 100 respondentów, z czego 60% to kobiety, a 40% stanowili mężczyźni. Najsilniejszą grupę wiekową stanowiły osoby w wieku 18-25 lat (65%), oraz 26-35 lat (22%).

Analiza struktury wykształcenia respondentów wskazuje, że dominowały osoby z wykształceniem wyższym (58%), następnie średnim (31%). Mniejszy odsetek badanych posiadał wykształcenie zawodowe (9%) oraz podstawowe (2%).

Najliczniejszą grupę respondentów stanowiły osoby mieszkające na wsi (38%). Pozostali badani zamieszkiwali głównie miasta powyżej 200 tys. mieszkańców (25%) oraz miasta liczące od 50 do 200 tys. mieszkańców (22%), natomiast najmniejszy odsetek pochodził z miast do 50 tys. mieszkańców (14%).

W pierwszej kolejności w ankiecie zastosowano pytanie filtrujące, którego celem było sprawdzenie, czy respondenci należą do grupy objętej badaniem. Pytanie dotyczyło częstotliwości podróżowania. Osoby, które zadeklarowały, że nie podróżują, były automatycznie kierowane do części metryczkowej ankiety, co pozwoliło na wyłączenie ich odpowiedzi z dalszej analizy merytorycznej. Zastosowanie pytania filtrującego umożliwiło skoncentrowanie badania na osobach, których doświadczenia pozostają zgodne z tematyką ankiety. Pozostało 94 respondentów co zostało zilustrowane wykresem 1.



Wykres 1. Częstotliwość podróży w przekroju wiekowym
Chart 1. Travel frequency by age
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

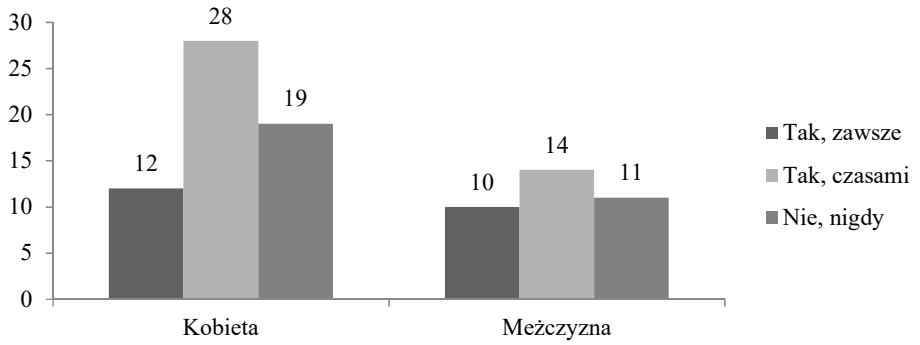
Największą aktywność podróżniczą przedstawioną na wykresie 1 wykazują osoby w wieku 18-25 lat, które najczęściej deklarują wyjazdy kilka razy w roku lub raz w roku. W grupie 26-35 lat również dominuje regularne podróżowanie, jednak liczba takich wskazań jest mniejsza. Wraz z wiekiem widoczny jest spadek częstotliwości wyjazdów oraz większy udział odpowiedzi wskazujących sporadyczne podróże lub ich brak. Należy jednak zaznaczyć, że grupa 18-25 lat stanowiła największy odsetek badanych, co mogło wpłynąć na widoczne różnice w strukturze odpowiedzi.

3.2. Zakup ubezpieczenia turystycznego

Kolejnym etapem analizy było określenie, czy respondenci decydują się na zakup ubezpieczenia turystycznego przed wyjazdem oraz jak często podejmują taką decyzję. Pytanie brzmiało „Czy przed wyjazdem wykupujesz ubezpieczenie turystyczne?". Wyniki zostały przedstawione na wykresie 2.

Wyniki ukazane na wykresie 2. cechuje zróżnicowane podejście do kwestii zakupu ubezpieczenia turystycznego przed wyjazdem w przekroju płciowym. Część badanych traktuje ubezpieczenie jako stały element przygotowań do podróży, podczas gdy inni decydują się na jego zakup jedynie sporadycznie lub wcale. Analiza odpowiedzi również wskazuje na różnice w deklarowanej częstotliwości nabywania polisy między kobietami a mężczyznami.

Analiza odpowiedzi respondentów wskazuje, że 23% badanych zadeklarowało, że zawsze wykupuje ubezpieczenie turystyczne przed podróżą, z czego było to 12 kobiet oraz 10 mężczyzn. 45% badanych deklaruje, iż czasami wykupuje ubezpieczenie turystyczne przed wyjazdem i było to aż 28 kobiet i 14 mężczyzn. Podczas gdy, 32% respondentów przyznało, że nigdy nie decyduje się na zakup tego rodzaju ubezpieczenia, aż 19 kobiet i 11 mężczyzn.



Wykres 2. Decyzję i częstotliwość zakupu ubezpieczenia turystycznego przed wyjazdem w przekroju płciowym

Chart 2. Decision and frequency of purchasing travel insurance before departure by gender

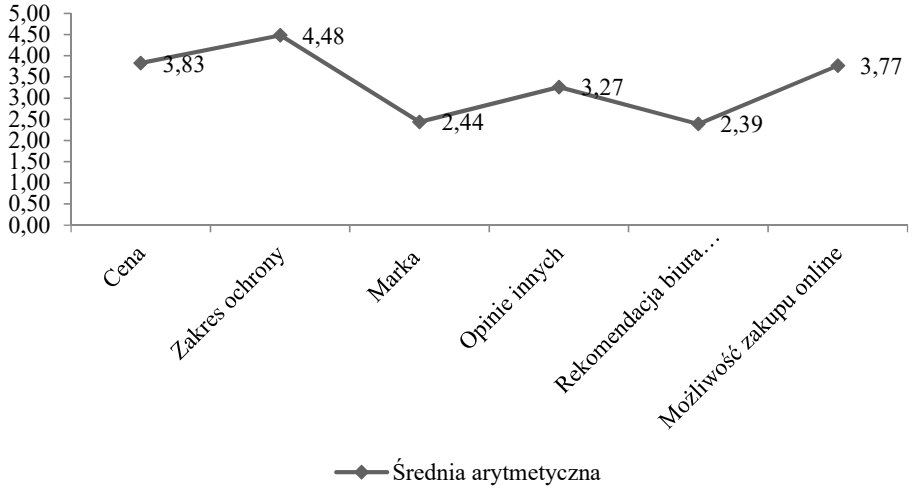
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

Uzyskane wyniki wskazują, że w obu analizowanych grupach dominuje okazjonalne wykupywanie ubezpieczenia turystycznego, przy czym kobiety nieco częściej niż mężczyźni deklarują regularny zakup, co może świadczyć o umiarkowanym poziomie świadomości znaczenia ubezpieczeń turystycznych wśród badanych.

Znaczny odsetek osób rezygnujących z zakupu polisy lub podejmujących taką decyzję jedynie okazjonalnie mogło wskazywać na niedostateczną ocenę ryzyka związanego z podróżowaniem lub przekonanie o braku potrzeby dodatkowego zabezpieczenia. Respondenci, którzy zaznaczyli odpowiedź „Nie, nigdy" w pytaniu o zakup ubezpieczenia, mieli możliwość wskazania powodu swojej decyzji, poprzez wskazanie kilku możliwych opcji. Najczęściej wymieniano posiadanie Europejskiej Karty Ubezpieczenia Zdrowotnego (EKUZ) lub podobnych produktów ubezpieczeniowych (19 osób), brak refleksji nad potrzebą ubezpieczenia („Nie myślałem/-am o tym", 12 osób), przekonanie, że ubezpieczenie nie jest potrzebne (10 osób) oraz zbyt wysoką cenę polisy (5 osób). Uzyskane odpowiedzi potwierdzają, że powody rezygnacji są zróżnicowane i obejmują zarówno aspekty praktyczne, jak i postrzeganie ryzyka czy kosztu ochrony ubezpieczeniowej.

3.3. Determinanty wyboru ubezpieczenia

Najważniejszym elementem badania było określenie, jakie czynniki mają największe znaczenie przy wyborze ubezpieczenia turystycznego. Respondenci zostali poproszeni o ocenę sześciu wskazanych czynników w skali od 1 do 5, gdzie 1 oznaczało czynnik nieważny, natomiast 5 bardzo ważny. Taka forma pytania umożliwiła określenie stopnia istotności poszczególnych czynników w procesie podejmowania decyzji zakupowej. Na wykresie 3. zilustrowano średnie arytmetyczne uzyskanych ocen.

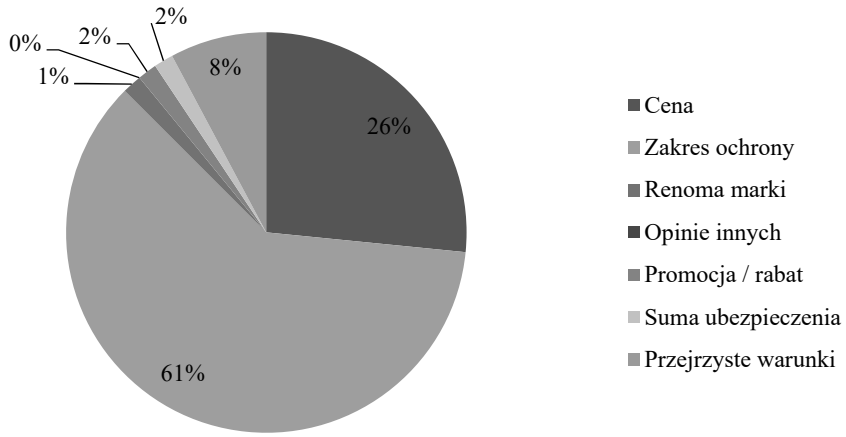


Wykres 3. Ocena znaczenia poszczególnych czynników przy wyborze ubezpieczenia
Chart 3. Assessment of the importance of individual factors when choosing insurance
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

Dane zaprezentowane na wykresie 3. wskazują, że najwyżej ocenionym czynnikiem był zakres ochrony (4,48), co potwierdza jego kluczowe znaczenie w procesie decyzyjnym respondentów. Wysoką wartość uzyskała również cena (3,83) oraz możliwość zakupu online (3,77), co świadczy o istotnej roli zarówno aspektu ekonomicznego, jak i wygody zawarcia umowy. Umiarkowane znaczenie przypisano opiniom innych osób (3,27). Najniżej ocenione zostały natomiast marka towarzystwa ubezpieczeniowego (2,44) oraz rekomendacja biura podróży (2,39), co sugeruje, że czynniki wizerunkowe i instytucjonalne mają mniejsze znaczenie niż realny zakres ochrony i warunki finansowe.

Dla pełniejszego zobrazowania hierarchii preferencji, badani mieli za zadanie odpowiedzieć, który z czynników jest według nich najważniejszy w podjęciu decyzji zakupowej ubezpieczeń turystycznych. Strukturę odpowiedzi respondentów ukazuje wykres 4.

Dane zaprezentowane na wykresie 4. wskazują, że zdecydowanie najważniejszym czynnikiem dla respondentów jest zakres ochrony, odpowiedź tę wskazało 61% badanych. Na drugim miejscu znalazła się cena ubezpieczenia (26%), co potwierdza istotną, choć nie dominującą rolę kosztu w procesie decyzyjnym. Znacznie mniejsze znaczenie przypisywano takim czynnikom jak przejrzyste warunki umowy (8%), renoma firmy (1%), promocja lub rabat (1%) oraz suma ubezpieczeniowa (1%). Żaden z respondentów nie wskazał opinii innych użytkowników jako najważniejszego czynnika (0%).



Wykres 4. Czynniki decydujące o podjęciu decyzji zakupowej
Chart 4. Factors determining the purchase decision
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

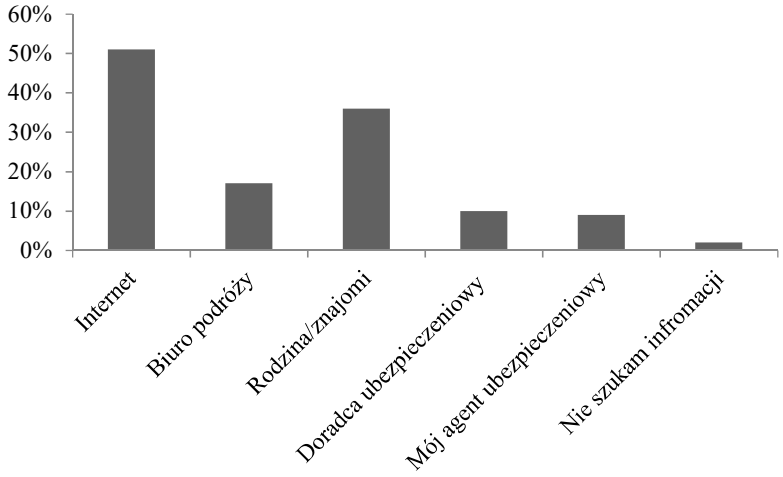
Uzyskane wyniki pokazują, że przy wyborze ubezpieczenia turystycznego podróżni kierują się przede wszystkim realnym zakresem oferowanej ochrony, a dopiero w dalszej kolejności względami finansowymi. Jednocześnie wskazują one, że decyzje zakupowe w tym obszarze mają w dużej mierze charakter racjonalny i oparte są na analizie zakresu ochrony oraz kosztów, przy jednoczesnym rosnącym znaczeniu dostępności zakupu online jako elementu ułatwiającego podjęcie decyzji.

Ponadto respondenci mieli możliwość dodania własnych uwag i wskazania, co według nich powinno zawierać dobre ubezpieczenie turystyczne. Analiza tych odpowiedzi pokazuje wyraźną tendencję do preferowania polis o szerokim zakresie ochrony, obejmujących przede wszystkim wysokie koszty leczenia i transportu medycznego, NNW, OC w życiu prywatnym oraz ubezpieczenie bagażu. Wielu ankietowanych wskazywało również na znaczenie ochrony assistance dostosowanej do kraju i charakteru podróży, a także ubezpieczenia od rezygnacji lub przerwania podróży.

Respondenci podkreślali również istotę przejrzystych i jasnych warunków umowy, uczciwego stosunku jakości do ceny oraz możliwości pokrycia niespodziewanych kosztów związanych z wypadkami, opóźnieniami czy innymi nieprzewidzianymi zdarzeniami. Wiele odpowiedzi wskazywało także na potrzebę dodatkowych świadczeń, takich jak transport zwłok do kraju, hospitalizacja, pokrycie leczenia za granicą czy ubezpieczenie sportowe.

3.4. Źródła informacji i świadomość ubezpieczeniowa

Kolejnym elementem analizy było określenie, z jakich źródeł respondenci czerpią informacje o ubezpieczeniach turystycznych oraz w jakim stopniu zapoznają się z Ogólnymi Warunkami Ubezpieczenia (OWU). Wykres 5. przedstawia sposoby szukania informacji na temat ubezpieczeń.

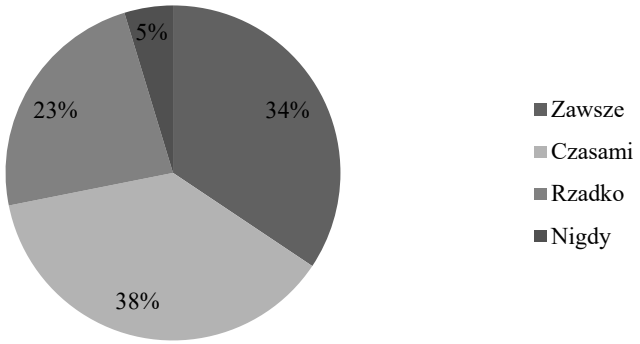


Wykres 5. Źródła pozyskiwania informacji o ubezpieczeniu turystycznym
Chart 5. Sources of information about travel insurance
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

Największy odsetek badanych wskazał Internet jako podstawowe źródło wiedzy o ubezpieczeniach turystycznych (51%), natomiast istotną rolę odgrywają również rodzina i znajomi (36%). Znacznie mniej osób korzystało z informacji przekazywanych przez biuro podróży (17%), doradcę ubezpieczeniowego (10%), czy agenta ubezpieczeniowego (9%). Tylko 2% respondentów przyznało, że w ogóle nie poszukuje informacji przed zakupem polisy.

Równolegle analizowano, w jakim stopniu respondenci zapoznają się z Ogólnymi Warunkami Ubezpieczenia (OWU) swoich polis. Dane zostały przedstawione na wykresie 6.

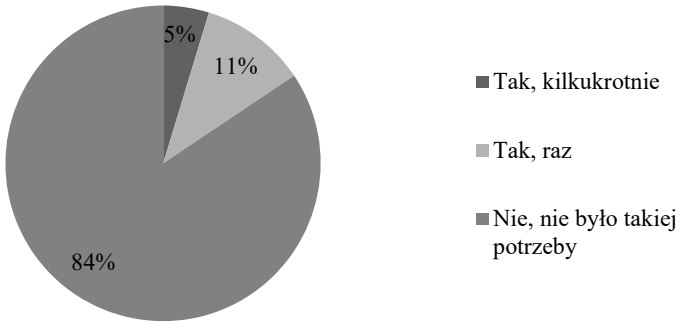
Zaprezentowane na wykresie 6. odpowiedzi wskazują, że 34% badanych czyta je zawsze, 38% czasami, 23% rzadko, a 5% nigdy. Wyniki te pokazują, że chociaż większość respondentów poszukuje informacji o ubezpieczeniach turystycznych, to poziom zaangażowania w lekturę OWU jest zróżnicowany. Może to wskazywać na częściowe niedocenianie znaczenia dokładnego zapoznania się z warunkami polisy lub ograniczenia czasowe i trudności w interpretacji dokumentów ubezpieczeniowych.



Wykres 6. Częstotliwość zapoznawania się z OWU
Chart 6. Frequency of reading the General Terms and Conditions
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

3.5. Doświadczenia z ubezpieczeniem

Kolejnym etapem analizy było zbadanie doświadczeń respondentów związanych z wykupionym ubezpieczeniem turystycznym. Respondenci zostali zapytani, czy podczas podróży mieli okazję skorzystać z ochrony ubezpieczeniowej, co pozwala ocenić praktyczne znaczenie posiadanej polisy. Wyniki przedstawiono na wykresie 7.



Wykres 7. Konieczność skorzystania z ubezpieczenia w trakcie podróży
Chart 7. The need to use insurance while traveling
Źródło: opracowanie własne na podstawie wyników przeprowadzonej własnej ankiety.

Wyniki wskazują, że zdecydowana większość, 84%, nie musiała korzystać z polisy, co oznacza, że w ich przypadku nie wystąpiła sytuacja wymagająca

interwencji ubezpieczeniowej. Jednocześnie 16% badanych skorzystało z ubezpieczenia.

Analiza tych danych pokazuje, że choć faktyczne wykorzystanie polisy jest stosunkowo rzadkie, posiadanie ubezpieczenia nadal odgrywa istotną rolę w zwiększaniu poczucia bezpieczeństwa i komfortu podróży. Dane przedstawione na wykresie 6 potwierdzają, że większość podróżnych traktuje ubezpieczenie jako formę zabezpieczenia na wypadek nieprzewidzianych zdarzeń, nawet jeśli nie wystąpiła potrzeba jego wykorzystania, co wskazuje na znaczącą funkcję prewencyjną i psychologiczną posiadanej ochrony.

W związku z tym ankietowani zostali zapytani, w jakim stopniu posiadanie ubezpieczenia turystycznego zwiększa ich poczucie bezpieczeństwa podczas podróży. Pytanie obejmowało skalę od 1 do 5, gdzie 1 oznaczało „wcale", a 5 „w bardzo dużym stopniu". Uzyskane wyniki wskazują, że większość badanych odczuwała znaczący wpływ posiadanej polisy na poczucie bezpieczeństwa. Najczęściej wskazywano odpowiedzi 4 („w dużym stopniu") i 5 („w bardzo dużym stopniu").

Następnie osoby, które faktycznie wykorzystały ubezpieczenie, zostały poproszone o ocenę swojego doświadczenia. Mieli oni do wyboru odpowiedzi „Bardzo dobrze", „ Dobrze", „Neutralnie", „Źle", „Bardzo źle". Podróżujący, którzy skorzystali z wykupionego wcześniej ubezpieczenia w większości mieli bardzo dobre doświadczenie (7/10 badanych). Dwoje z nich miało dobre, a jeden badany wskazał na neutralny rodzaj tego doświadczenia.

Dodatkowo respondenci zostali również zapytani, czy byliby skłonni dopłacić za większy zakres ochrony ubezpieczeniowej. Wyniki wskazują, że zdecydowana większość, 75% badanych, byłaby gotowa ponieść dodatkowe koszty, natomiast 25% nie wyraziło gotowości dopłaty w celu uzyskania szerszej ochrony.

Wskazuje to, że ubezpieczenie turystyczne pełni nie tylko funkcję ochronną w sensie finansowym, ale również istotnie zwiększa komfort psychiczny podróżnych, dając im poczucie spokoju i pewności podczas wyjazdów. Wyniki te potwierdzają teoretyczne założenia dotyczące psychologicznej roli ubezpieczeń w procesie podejmowania decyzji nabywczych.

4. Wnioski

Analiza wyników ankiety pozwoliła zidentyfikować kluczowe czynniki wpływające na wybór ubezpieczenia turystycznego oraz rolę posiadanej polisy w poczuciu bezpieczeństwa. Najważniejszym determinantem decyzji zakupowej był zakres ochrony, który w ocenie respondentów zdecydowanie przewyższał pozostałe czynniki, takie jak cena, renoma marki towarzystwa ubezpieczeniowego czy opinie innych osób. Cena polisy pełni rolę czynnika wspierającego decyzję, natomiast wygoda zakupu online zyskuje na znaczeniu w kontekście rosnących oczekiwań podróżnych. Choć większość respondentów nie miała potrzeby korzystania

z wykupionej polisy, posiadanie ubezpieczenia znacząco zwiększa poczucie bezpieczeństwa, pełniąc równocześnie funkcję praktyczną i psychologiczną. Analiza źródeł informacji wykazała dominację Internetu oraz opinii rodziny i znajomych, podczas gdy czytanie OWU wciąż nie jest powszechne, co wskazuje na potrzebę zwiększania świadomości dotyczącej warunków polisy. Ogólne wyniki potwierdzają, że decyzje zakupowe w zakresie ubezpieczeń turystycznych mają charakter racjonalny i opierają się przede wszystkim na realnym zakresie ochrony oraz kosztach, przy rosnącym znaczeniu wygody i dostępności zakupów online. Dodatkowo, odpowiedzi udzielone w otwartym pytaniu pozwoliły stwierdzić, że podróżni oczekują od polis nie tylko praktycznej ochrony finansowej, ale również kompleksowego i elastycznego wsparcia w różnych sytuacjach kryzysowych. Z perspektywy użytkowników kluczowe są zarówno realny zakres ochrony, jak i jasność warunków oraz adekwatność ceny do oferowanej ochrony, co podkreśla wagę równoważenia bezpieczeństwa i kosztów przy podejmowaniu decyzji zakupowych.

Literatura

[1] Ajzen I., *The theory of planned behavior: Frequently asked questions*, Human Behavior and Emerging Technologies, 2(3) 2020.

[2] Błaś R., *Analiza rynku ubezpieczeniowego w Polsce pod kątem aplikacji mobilnych*, Acta Universitatis Nicolai Copernici. Zarządzanie, 2016.

[3] Gasińska M., *Ubezpieczenia turystyczne w systemie ubezpieczeń gospodarczych*, Zeszyty Naukowe Uczelni Vistula 32, 2013.

[4] Kotler P., Keller K.L., *Marketing*, Dom Wydawniczy Rebis, Poznań 2015.

[5] Kotler P., *Zarządzanie marketingiem, analiza, planowanie, wdrażanie i kontrola*, Prentice Hall, Englewood Cliffs, New Jersey 1991.

[6] Kowalewski E. *Ubezpieczenia turystyczne*, [w:] Nestorowicz M., red., Odpowiedzialność biur podróży a ochrona klientów w prawie polskim i Unii Europejskiej, Wyższa Szkoła Bankowa w Toruniu, Toruń, 2013.

[7] Kruczek Z., *Obsługa ruchu turystycznego. Teoria i praktyka*, Proksenia, Kraków 2010.

[8] Maciejewski G., *Ryzyko w decyzjach nabywczych konsumenta*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice 2010.

[9] Nowotarska-Romaniak B., *Zachowania klientów indywidualnych w procesie zakupu usługi ubezpieczeniowej*, Oficyna Wolters Kluwer, Warszawa, 2013.

[10] Sawińska A., *Ubezpieczenia w turystyce* [w:] Współczesna Turystyka. Środowisko-człowiek- gospodarka, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2023.

[11] Wang S.-P., Chen L.-C., Chen M.-S., Li M.-J., *Purchasing Factors for Travel Insurance by Asian Consumers*, International Journal of Human Resource Studies, tom 9, nr 1 2019.

[12] Żylak D., Hadzik A., Ryśnik J., Tomik R., *Produkty ubezpieczeniowe jako narzędzia ochrony przed ryzykiem w turystyce sportowej: polskie i chorwackie studia przypadku*, Folia Turistica/Akademia Wychowania Fizycznego im. B.Czecha w Krakowie, nr 43, 2017.

Źródła internetowe

[13] https://www.macrothink.org/journal/index.php/ijhrs/article/view/14171/11358 (dostęp:12.02.2026).

[14] https://media.tueuropa.pl/409997-wyjazdy-za-granice-coraz-czesciej-sklaniaja-do-zakupu-ubezpieczenia (dostęp: 12.02.2026).

[15] https://rynek-turystyczny.pl/artykul/54258/europejski-i-swiatowy-rynek-ubezpieczen-turystycznych/ (dostęp: 12.02.2026).

[16] https://rf.gov.pl/wp-content/uploads/2024/04/Ubezpieczenia-podrozy-29_06_2023.pdf (dostęp: 12.02.2026).

[17] https://rf.gov.pl/wp-content/uploads/2024/04/Prezentacja_Ubezpieczenie-turystyczne-w-pigulce.pdf (dostęp: 12.02.2026).

FACTORS INFLUENCING THE CHOICE OF TRAVEL INSURANCE BY TRAVELERS

Summary

Travel insurance plays a crucial role in protecting travelers against unforeseen costs and risks associated with trips, providing both financial security and psychological comfort. This study aimed to identify the factors influencing the choice of travel insurance and to examine the role of holding a policy in shaping travelers' sense of safety. The most important criterion in decision-making proved to be the scope and level of coverage, which significantly outweighed other considerations such as price, company reputation, or recommendations from acquaintances. Most respondents value the protective function of insurance and are willing to pay extra for extended coverage, although price remains a relevant factor in their decisions. The convenience of online purchase is also gaining importance, reflecting travelers' increasing expectations regarding accessibility and simplicity of services.

Respondents primarily rely on information obtained from the Internet or from family and friends, while regularly reviewing the terms and conditions of policies is still not common. Open-ended responses indicated that travelers expect insurance to provide not only financial protection but also support in a variety of crisis situations, clear and transparent contract terms, and a price appropriate to the level of coverage. Although most participants had not needed to use their insurance in practice, simply having a policy significantly enhances their sense of security, serving both practical and psychological functions. The survey, conducted among 100 travelers, including 64 actual policyholders, confirmed that purchasing decisions are largely rational, based primarily on an assessment of coverage and costs, with the growing importance of convenience and service accessibility.

Keywords: insurance, tourism, travel, scope of coverage, travel protection.

Sebastian Cwynar, Maja Jaszowska, Natalia Klocek,
Julia Ziąbek, Dominik Dworak
Studenckie Koło Naukowe Informatyków „Kod”

dr inż. Bartosz Trybus
opiekun Koła Naukowego

METODYKA PROJEKTOWANIA SYSTEMÓW STEROWANIA ANIMACJĄ SZKIELETOWĄ POSTACI 3D

Streszczenie

W dobie dynamicznego rozwoju grafiki komputerowej, szczególnie w sektorze gier wideo i wirtualnej rzeczywistości, kluczowym wyzwaniem staje się opracowanie metod animacji łączących wysoki realizm z efektywnością produkcji. Celem niniejszego rozdziału było opracowanie i weryfikacja autorskiej metodyki projektowania systemów sterowania (riggingu) postaci 3D, stanowiącej efektywny kompromis pomiędzy wiernością ruchu a nakładem pracy animatora. Przedmiotem analizy uczyniono problem wysokiej czasochłonności oraz ograniczonej ergonomii klasycznych metod animacji. Tradycyjne podejścia, oparte na bezpośredniej manipulacji geometrią siatki lub wyłącznym stosowaniu kinematyki prostej (Forward Kinematics), obarczone są ryzykiem powstawania błędów technicznych, takich jak nienaturalne deformacje w stawach, brak stabilności punktów podparcia (zjawisko "foot sliding") oraz konieczność stosowania żmudnej kontr-animacji. W oparciu o krytyczną analizę ewolucji systemów animacyjnych, zaproponowano zintegrowany workflow realizowany w środowisku Blender. Opracowane rozwiązanie rozszerza standardową animację szkieletową o hybrydowy system sterowania, integrujący łańcuchy kinematyki odwrotnej (Inverse Kinematics), zaawansowane ograniczenia (Constraints) oraz pomocnicze kontrolery orientacji. Zaproponowane podejście stanowi nowoczesną alternatywę dla manualnych transformacji poszczególnych kości, przenosząc ciężar dbałości o poprawność fizyczną z animatora na algorytmy sterujące. Mimo iż wdrożenie tej metodyki wiąże się ze zwiększoną złożonością etapu przygotowawczego (riggingu), pozwala ona na bardziej efektywne animowanie. Automatyzacja kluczowych aspektów motoryki pozwala na znaczącą redukcję liczby niezbędnych operacji korekcyjnych oraz eliminację artefaktów wizualnych. Ostatecznie, zaprezentowany system umożliwia twórcom skupienie się na samej animacji przy jednoczesnym zachowaniu fizycznej poprawności animacji, co przekłada się na wyższą jakość końcowego produktu.

Słowa kluczowe: animacja postaci 3D, animacja szkieletowa, rigowanie, workflow animacji, kinematyka odwrotna, constraints.

1. Wprowadzenie

Dynamiczny rozwój branży rozrywkowej, w szczególności gier wideo oraz produkcji filmowych realizowanych w technice CGI, stawia przed twórcami grafiki trójwymiarowej coraz wyższe wymagania w zakresie wierności odwzorowania rzeczywistości. O ile współczesne silniki renderujące i techniki modelowania pozwalają na osiągnięcie fotorealistycznej jakości statycznych obiektów oraz tekstur, o tyle wiarygodne wprawienie cyfrowej postaci w ruch nadal stanowi jedno z najbardziej złożonych wyzwań technologicznych i artystycznych. Odbiór wirtualnego aktora przez widza jest determinowany nie tylko przez jakość jego modelu, ale przede wszystkim przez naturalność motoryki – nawet najdrobniejsze odstępstwa od zasad biomechaniki czy fizyki ruchu są natychmiast wychwytywane przez ludzkie oko, burząc immersję i wiarygodność przedstawianego świata.

W procesie produkcji animacji 3D kluczowym zagadnieniem staje się zatem znalezienie balansu pomiędzy jakością artystyczną ruchu a efektywnością procesu jego tworzenia. Tradycyjne metody animowania, opierające się na ręcznym definiowaniu transformacji dla każdego elementu szkieletu z osobna, cechują się bardzo wysoką pracochłonnością. Animator zmuszony jest do ciągłego kontrolowania dziesiątek punktów artykulacji, co przy złożonych sekwencjach ruchu często prowadzi do powstawania błędów technicznych, takich jak brak stabilności punktów podparcia (efekt "ślizgania się" stóp po podłożu) czy nienaturalne deformacje siatki w okolicach stawów [2]. Korekta tych niedoskonałości w klasycznym podejściu wymaga żmudnej edycji krzywych animacyjnych, co drastycznie wydłuża czas produkcji.

Niniejszy rozdział koncentruje się na problematyce optymalizacji warsztatu animatora poprzez projektowanie zaawansowanych systemów sterowania (riggingu) w środowisku Blender. Głównym założeniem pracy jest odejście od prostych struktur szkieletowych na rzecz systemów inteligentnych, które automatyzują techniczne aspekty ruchu. Zaprezentowane zostanie podejście polegające na implementacji autorskiego rigu, wykorzystującego algorytmy Kinematyki Odwrotnej (Inverse Kinematics), dedykowane systemy ograniczeń (Constraints) oraz kontrolery pomocnicze.

Celem pracy jest przedstawienie, w jaki sposób odpowiednie przygotowanie techniczne modelu na etapie riggingu może przyczynić się do usprawnienia procesu animacji przy jednoczesnym podniesieniu jej jakości wizualnej. Zaproponowana metodyka zakłada częściowe przeniesienie odpowiedzialności za zachowanie poprawności anatomicznej z animatora na algorytmy sterujące, co umożliwia twórcy większe skupienie się na artystycznym aspekcie animacji. W dalszej części opracowania dokonano przeglądu wybranych technik animacji wraz z omówieniem ich ograniczeń, a następnie przedstawiono proces budowy kompletnego systemu sterowania, stanowiącego praktyczną alternatywę dla klasycznych metod animacji postaci 3D.

2. Znane metody animowania postaci

2.1. Bezpośrednia manipulacja geometrią i interpolacja kształtów (Vertex Animation)

W początkowej fazie rozwoju grafiki trójwymiarowej, zanim standardem stały się systemy szkieletowe, dominującą metodą wprawiania obiektów w ruch była bezpośrednia ingerencja w strukturę siatki geometrycznej. Technika ta, w literaturze anglojęzycznej określana mianem Vertex Animation lub Per-vertex Animation, stanowiła cyfrowy odpowiednik klasycznej animacji poklatkowej. W podejściu tym animator nie dysponował systemem pośredniczącym (takim jak kości), lecz operował bezpośrednio na chmurze punktów definiujących powierzchnię modelu, ręcznie ustalając ich położenie dla każdej klatki kluczowej.

Ruch pomiędzy zdefiniowanymi stanami generowany był przez silnik graficzny w oparciu o interpolację liniową. Oznaczało to, że wierzchołek przemieszczał się z punktu A do punktu B po najkrótszej możliwej drodze – linii prostej. Choć metoda ta zapewniała absolutną kontrolę nad kształtem obiektu w klatkach kluczowych, obarczona była brakiem struktury hierarchicznej. Przesunięcie grupy wierzchołków reprezentujących ramię nie powodowało automatycznego przemieszczenia przedramienia i dłoni. Wymagało to od twórcy ręcznego pozycjonowania każdej części ciała z osobna, co przy złożonych modelach prowadziło do wykładniczego wzrostu nakładu pracy.

Istotnym problemem jakościowym tego podejścia były artefakty wizualne wynikające z natury samej interpolacji. Rick Parent w fundamentalnej pracy "Computer Animation: Algorithms and Techniques" [6] zwraca uwagę na zjawisko tzw. liniowości interpolacji (linear interpolation artifacts). W przypadku ruchu obrotowego – naturalnego dla stawów człowieka – przemieszczanie się wierzchołków po cięciwie zamiast po łuku powoduje niepożądany efekt zapadania się objętości siatki (tzw. volume loss lub collapsing joint). Kończyna w połowie zgięcia nienaturalnie traci swoją grubość, przypominając zgniecioną rurkę, co wymuszało na twórcach sztuczne zagęszczanie klatek kluczowych.

Dodatkową barierą była nieefektywność obliczeniowa wczesnych implementacji (np. format MD2 z lat 90.), wymagających zapisu wektorów pozycji dla każdej klatki, co generowało pliki o znacznej objętości. Współcześnie technika ta, ewoluując w formę Morph Targets (Blend Shapes), znajduje zastosowanie głównie w animacji mimiki twarzy, gdzie deformacje są liniowe. W kontekście motoryki całego ciała została ona jednak całkowicie wyparta przez wydajniejsze systemy szkieletowe.

2.2. Implementacja hierarchicznych struktur szkieletowych i Kinematyka Prosta

Animacja szkieletowa (Skeletal Animation) stanowi obecnie standardowe podejście w procesie przygotowania postaci trójwymiarowych do animacji. Metoda ta opiera się na zastosowaniu struktur hierarchicznych, będących logicznym

odwzorowaniem układu kostnego organizmów żywych. W tym podejściu geometria modelu zostaje powiązana z systemem kości sterujących (bones) za pomocą algorytmów przypisywania wag wierzchołków (tzw. proces skinningu). Fundamentalną cechą tego systemu jest istnienie relacji „rodzic-dziecko" (parent-child relationship), w której transformacja (przesunięcie, rotacja, skalowanie) elementu nadrzędnego jest automatycznie dziedziczona przez wszystkie elementy podrzędne. Przykładowo, rotacja kości ramiennej naturalnie pociąga za sobą zmianę położenia przedramienia i dłoni.

Podstawowym paradygmatem sterowania, bezpośrednio wynikającym z tego typu budowy, jest Kinematyka Prosta (Forward Kinematics – FK). W metodzie tej pozycjonowanie dowolnej kończyny odbywa się poprzez sekwencyjne definiowanie kątów obrotu dla każdego stawu w łańcuchu, postępując od korzenia (root) ku elementowi końcowemu (end effector).

Jak wskazuje Isaac Kerlow w monografii "The Art of 3D Computer Animation and Effects" [4], pomimo intuicyjności samej metody, sterowanie oparte wyłącznie na FK posiada istotne ograniczenia w kontekście interakcji postaci z otoczeniem. Głównym problemem jest brak możliwości stabilizacji punktu końcowego względem globalnego układu współrzędnych. Ponieważ każdy ruch tułowia jest bezwarunkowo propagowany w dół hierarchii, powoduje to niepożądane przemieszczanie się stóp. Wymusza to na animatorze stosowanie tzw. „kontr-animacji" (counter-animation) – ręcznego, żmudnego korygowania rotacji nóg w każdej klatce, aby zniwelować wpływ ruchu bioder i utrzymać stopę na ziemi. Błędy na tym etapie najczęściej objawiają się nienaturalnym „ślizganiem się" postaci po podłożu (foot sliding).

Mimo trudności w stabilizacji kończyn, Kinematyka Prosta pozostaje niezastąpioną i powszechnie stosowaną techniką w określonych scenariuszach ruchowych. Ze względu na swoją naturę – polegającą na bezpośredniej interpolacji kątów obrotu – metoda ta doskonale sprawdza się w animacji swobodnych ruchów kończyn górnych (np. machanie rękami podczas chodu), ogona czy gestykulacji. W takich przypadkach pozwala ona na wygenerowanie naturalnych łuków ruchu (arcs) oraz płynnej bezwładności, które są trudne do uzyskania w innych systemach. Z tego względu w profesjonalnych potokach produkcyjnych metody FK nie traktuje się jako przestarzałej, lecz wykorzystuje komplementarnie z bardziej zaawansowanymi rozwiązaniami, umożliwiając animatorowi płynne przełączanie się między trybami sterowania w zależności od specyfiki ujęcia.

2.3. Zastosowanie znormalizowanego szkieletu oraz odwrotnej kinematyki (IK)

Kluczowym mechanizmem matematycznym i algorytmicznym wykorzystywanym podczas animacji znormalizowanych szkieletów jest odwrotna kinematyka (ang. Inverse Kinematics – IK). W odróżnieniu od klasycznej kinematyki prostej (ang. Forward Kinematics – FK), w której ruch propagowany jest kaskadowo od kości nadrzędnej (rodzica) do podrzędnej (dziecka), algorytmy IK działają

w kierunku przeciwnym. Proces ten opiera się na precyzyjnym określeniu docelowego położenia tzw. efektora końcowego (ang. end effector), którym w przypadku modeli postaci jest zazwyczaj dłoń lub stopa. Na podstawie wektora pozycji tego punktu, system automatycznie oblicza i dostosowuje niezbędne kąty obrotu dla wszystkich powiązanych kości nadrzędnych w danym łańcuchu kinematycznym (np. kości przedramienia i ramienia), aby efektor końcowy znalazł się w zadanym miejscu docelowym [1].

Implementacja systemów odwrotnej kinematyki w oparciu o standardowe szkielety z systemów takich jak Mixamo znacząco optymalizuje proces produkcyjny. Po wygenerowaniu szkieletu następuje proces określania wag (ang. Weight Paint), podczas którego definiuje się, w jakim stopniu dana kość kontroluje poszczególne wierzchołki modelu. Każda kość otrzymuje własny rozkład wag, gdzie maksymalny wpływ (często wizualizowany kolorem czerwonym) oznacza pełne odwzorowanie deformacji obiektu względem ruchu kości, podczas gdy brak wpływu jest całkowicie ignorowany. Zastosowanie kontrolerów IK na tak przygotowanym i zważonym modelu jest niezbędne do uzyskania realistycznych interakcji postaci z otoczeniem. Mechanizm ten gwarantuje m.in. stabilny kontakt stóp postaci z ukształtowaniem terenu, zapobiegając nienaturalnemu zjawisku „ślizgania się" (ang. foot sliding) podczas odtwarzania standardowych cykli chodu w środowiskach silników gier komputerowych.

2.4. Zaawansowana modyfikacja układu kinematycznego

W zaawansowanych procesach animacji 3D standardowy obiekt typu Armature, służący jako system szkieletowy do deformacji siatek, jest poddawany znaczącej rozbudowie w celu poprawy efektywności pracy animatora. Bazowy szkielet modyfikuje się, tworząc system hybrydowy, który łączy zalety kinematyki prostej (FK) oraz odwrotnej (IK). Przełączanie między tymi trybami (tzw. IK/FK switch) realizowane jest za pomocą wprowadzonych do struktury dodatkowych kontrolerów wizualnych. Te niestandardowe obiekty (najczęściej krzywe) tworzą oddzielną warstwę sterującą, co pozwala animatorowi operować wyłącznie na intuicyjnych uchwytach, bez konieczności bezpośredniej ingerencji w bazowe kości deformujące model [8].

Kolejnym istotnym krokiem w optymalizacji działania szkieletu jest zastosowanie więzów kinematycznych (ang. Constraints), które automatyzują określone zachowania układu kości. Klasycznym przykładem takiego rozwiązania jest przypisanie typu Track To do kości szyi i głowy. Skonfigurowanie tego ograniczenia sprawia, że orientacja przestrzenna głowy postaci jest automatycznie determinowana przez pozycję zewnętrznego obiektu docelowego (tzw. targetu). Takie podejście znacząco upraszcza proces animowania punktu skupienia wzroku postaci, eliminując konieczność ręcznego obracania poszczególnych kości szyjnych w każdej klatce kluczowej.

Najwyższy poziom elastyczności zrigowanego modelu osiąga się poprzez implementację sterowników (ang. Drivers). W środowisku programów do grafiki 3D,

takich jak Blender, sterowniki są funkcjami i wyrażeniami matematycznymi, które dynamicznie łączą parametry różnych obiektów. W kontekście układu szkieletowego, Drivers wykorzystuje się do tworzenia złożonych zależności – na przykład zgięcie kości ramienia w trybie IK może, poprzez sterownik, automatycznie aktywować klucze kształtu (ang. Shape Keys) symulujące napięcie mięśni, lub jeden kontroler może płynnie sterować rozwarciem palców dłoni. Integracja systemów IK/FK ze sterownikami oraz więzami pozwala na stworzenie wysoce responsywnego i zoptymalizowanego modelu produkcyjnego.

3. Proces tworzenia standardowego szkieletu przeznaczonego do późniejszej modyfikacji

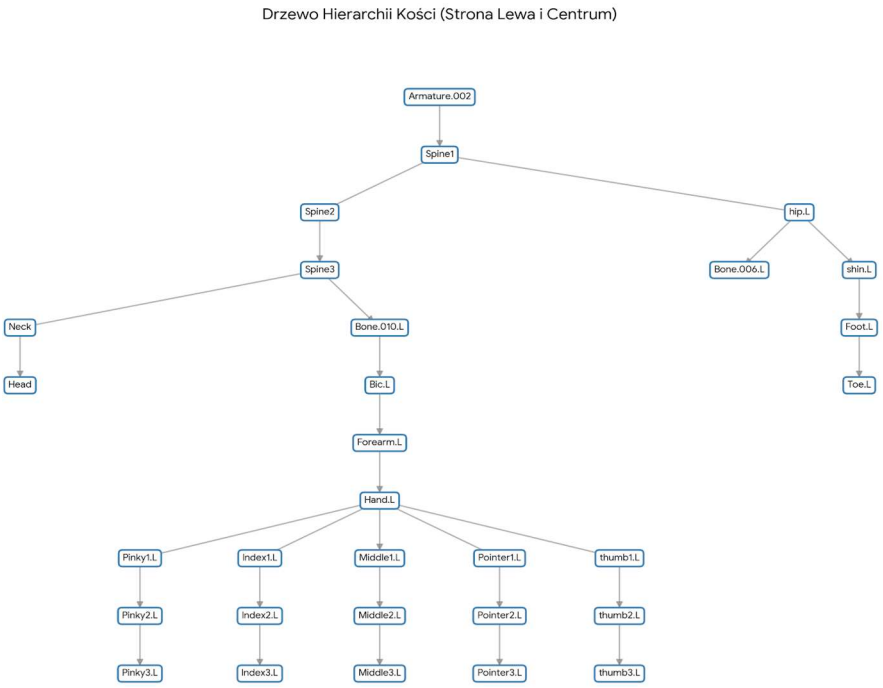
Praktyczny etap prac nad zintegrowanym systemem sterowania rozpoczął się od stworzenia standardowej, hierarchicznej struktury szkieletowej (obiektu typu Armature). Na tym etapie celem było stworzenie poprawnej bazy szkieletu, która w kolejnych fazach projektu posłuży jako fundament pod autorskie modyfikacje i wdrożenie hybrydowego systemu IK/FK.

Kluczowym aspektem tego podejścia było rygorystyczne przestrzeganie znormalizowanej konwencji nazewnictwa. Wszystkim kościom lewostronnym nadano unikalne identyfikatory zakończone sufiksem .L (np. Forearm.L, Hand.L, Shin.L). Zastosowanie tej konkretnej nomenklatury było zabiegiem celowym, warunkującym możliwość wykorzystania zautomatyzowanych narzędzi środowiska Blender [3].

Po ukończeniu i weryfikacji lewej strony szkieletu, zastosowano narzędzie Symmetrize. Algorytm ten, bazując na sufiksach .L, automatycznie wygenerował lustrzane odbicie całej struktury wzdłuż osi X. Proces ten nie tylko powielił geometrię i układ hierarchiczny kości (relacje "rodzic-dziecko"), ale również precyzyjnie skopiował kąty obrotu oraz automatycznie zaktualizował nazewnictwo prawej strony, podmieniając sufiksy na .R (np. Forearm R., Shin R.).

Zastosowanie techniki symetryzacji całkowicie wyeliminowało ryzyko błędów ludzkich związanych z asymetrią szkieletu, znacząco redukując czasochłonność fazy przygotowawczej. Tak przygotowany, szkielet oparty w całości na kinematyce prostej (FK), stanowił bazę wyjściową dla wdrożenia docelowych mechanizmów automatyzujących, które zostaną szczegółowo omówione w kolejnych rozdziałach niniejszej pracy.

Na rysunku 1. przedstawiono zrealizowane drzewo hierarchii kości w szkielecie, gdzie główną kością początkową (bazową) jest Spine1, z której poprowadzone są kolejne powiązania tworzące całość układu kinematycznego. Na rysunku pominięto kości z sufiksem R w celu ograniczenia liczby elementów w drzewie. Z kolei na rysunku 2. zaprezentowano wygląd stworzonego szkieletu w przestrzeni roboczej programu Blender, w postaci zrzutu ekranu wykonanego bezpośrednio z obszaru roboczego 3D.



Rys. 1. Drzewo hierarchii kości bez kości z sufiksem R
Fig. 1. Bone hierarchy tree without bones with the 'R' suffix



Rys. 2. Szkielet bazowy
Fig. 2. Base skeleton

4. Implementacja Kinematyki Odwrotnej (IK)

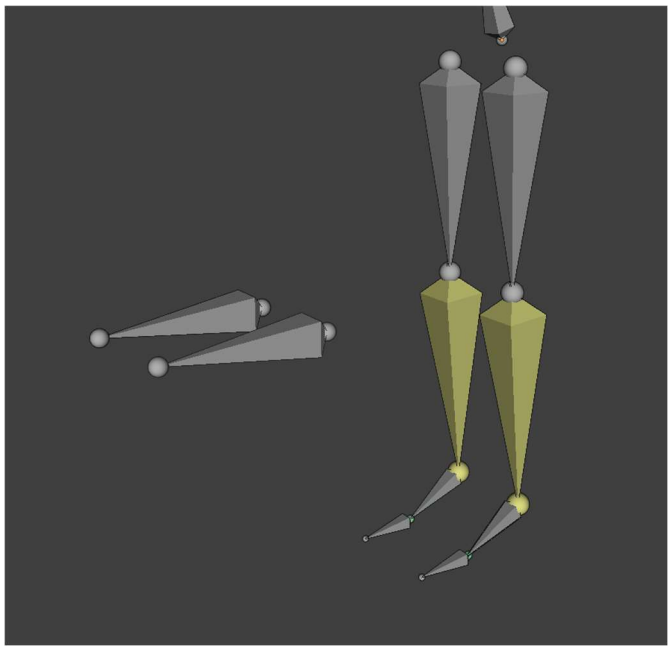
Kolejnym etapem rozbudowy szkieletu było wdrożenie modyfikacji mających na celu automatyzację ruchu kończyn. Podstawową wadą stworzonej wcześniej struktury FK (ang. Forward Kinematics) był brak stabilizacji punktów podparcia względem globalnego układu współrzędnych. Aby rozwiązać ten problem, zaimplementowano hybrydowy system sterowania, integrujący łańcuchy Kinematyki Odwrotnej (ang. Inverse Kinematics – IK) oraz zaawansowane ograniczenia. Zgodnie z przyjętą metodyką opisaną w rozdziale 3., wstępną konfigurację przeprowadzono wyłącznie dla lewej strony szkieletu (wykorzystując sufiksy .L), co umożliwiło późniejsze zduplikowanie kompletnego układu na prawą stronę (.R) za pomocą wbudowanych narzędzi symetryzacji.

4.1. Konfiguracja łańcucha kinematycznego kończyn dolnych

W przypadku kończyn dolnych, główny modyfikator więzów kinematycznych (IK Constraint) przypisano do kości podudzia (Shin.L). Prawidłowe funkcjonowanie tego algorytmu wymagało stworzenia niezależnych, pomocniczych kości sterujących, które zostały wyłączone z bazowej hierarchii deformacji siatki.

W dokładnym położeniu oryginalnej kości stopy (Foot.L) stworzono nową kość pomocniczą o nazwie FootIK.L. Została ona zdefiniowana w ograniczeniu jako główny obiekt docelowy (ang. Target Bone). Rozwiązanie to sprawia, że animator operuje wyłącznie kością FootIK.L, a algorytm matematyczny automatycznie oblicza i dostosowuje kąty zgięcia podudzia i uda, aby utrzymać stopę w pożądanym miejscu. Eliminuje to ryzyko występowania artefaktów wizualnych w postaci utraty stabilności punktu podparcia (foot sliding).

Aby zagwarantować poprawną orientację przestrzenną stawu kolanowego podczas zginania nogi, wykorzystano parametr Pole Target. W tym celu utworzono drugą kość pomocniczą – KneeIK.L. Umiejscowiono ją dokładnie na wysokości kolana, lecz w znacznym oddaleniu przed modelem postaci. Kość ta pełni funkcję wektora kierunkowego, do którego permanentnie podąża zgięcie łańcucha IK. Parametr długości łańcucha (ang. Chain Length) ustawiono na wartość 2, definiując tym samym, że kinematyka odwrotna oddziałuje kaskadowo wyłącznie na dwie kości w górę hierarchii (tj. podudzie Shin.L oraz udo Hip.L). Na rysunku 3. zaprezentowano zrzut ekranu ze środowiska Blender, przedstawiający widok dodanych i odpowiednio spozycjonowanych kości pomocniczych, niezbędnych do obsługi opisanego systemu IK dla kończyny dolnej.



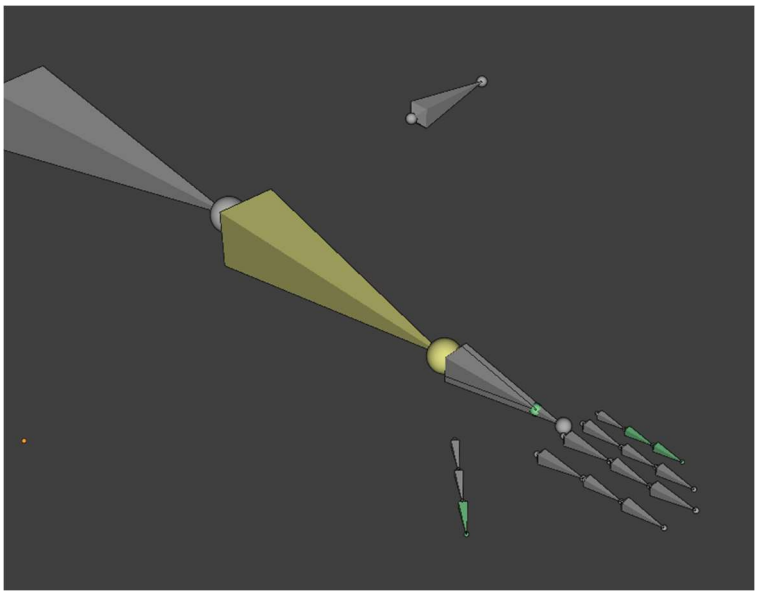
Rys.3. Inwersja kinematyczna kończyn dolnych
Fig. 3. Inverse kinematics of the lower limbs

4.2. Konfiguracja łańcucha kinematycznego kończyn górnych

Analogiczną metodykę zastosowano w celu zwiększenia efektywności sterowania układem kończyn górnych. Ograniczenie Kinematyki Odwrotnej dodano docelowo do kości przedramienia (Forearm.L).

Również i w tym układzie oparto się na dwóch kościach pomocniczych. Jako główny kontroler (obiekt Target Bone) stworzono kość HandIK.L, którą spozycjonowano w miejscu występowania oryginalnej kości dłoni (Hand.L). Translacja tej pojedynczej kości w przestrzeni trójwymiarowej w pełni determinuje aktualne ułożenie całej ręki. Do stabilizacji płaszczyzny zgięcia stawu łokciowego zastosowano kość kierunkową ElbowIK.L pełniącą funkcję Pole Target. Została ona umiejscowiona na wysokości łokcia i celowo odsunięta do tyłu, za plecy postaci, wyznaczając pożądany, naturalny kierunek zginania się ręki.

Podobnie jak w przypadku kończyn dolnych, zmienną Chain Length dla więzu IK przedramienia ograniczono do wartości 2. Dzięki temu algorytm wpływa wyłącznie na kości przedramienia (Forearm.L) oraz ramienia (Bic.L), nie ingerując w orientację obojczyka ani wyższych partii kręgosłupa, co przekłada się na wysoką przewidywalność animacji. Na rysunku 4. przedstawiono analogiczny zrzut ekranu ukazujący właściwe rozmieszczenie kości pomocniczych dla omówionego systemu IK kończyny górnej.



Rys. 4. Inwersja kinematyczna kończyn górnych
Fig. 4. Inverse kinematics of the upper limbs

5. Zaawansowane modyfikacje i automatyzacja motoryki postaci

Po zdefiniowaniu głównych łańcuchów kinematyki odwrotnej (IK) dla kończyn górnych i dolnych, kolejnym etapem prac było wdrożenie szeregu zaawansowanych modyfikacji uszczegóławiających. W profesjonalnych procesach animacji 3D standardowy szkielet poddawany jest znaczącej rozbudowie w celu maksymalnej poprawy efektywności pracy animatora. Aby osiągnąć wysoce responsywny model produkcyjny, strukturę wzbogacono o dodatkowe więzy kinematyczne (ang. Constraints) oraz dedykowane relacje sterujące [9].

Niniejszy rozdział koncentruje się na implementacji autorskiego połączenia mechanizmów automatyzujących specyficzne, często powtarzalne i najbardziej czasochłonne aspekty ruchu. W kolejnych podrozdziałach szczegółowo omówione zostaną następujące rozwiązania techniczne:

- zautomatyzowane śledzenie celu przez głowę postaci,
- system ułatwiający synchroniczne zginanie palców dłoni,
- ograniczniki stopy umożliwiające płynne stawanie na palcach,
- mechanizm wzruszania ramionami (obojczykami),
- implementacja dodatkowej kości pomocniczej ułatwiających animowanie kości odpowiadającym rejonowi kręgosłupa.

Nadrzędnym celem zastosowania powyższych modyfikacji jest dalsze przeniesienie ciężaru dbałości o poprawność anatomiczną z człowieka na algorytmy

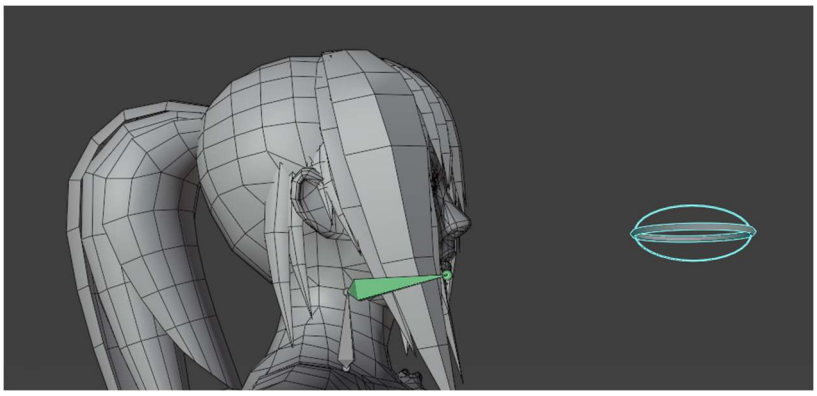
sterujące. Takie przygotowanie techniczne modelu na etapie riggingu pozwala na znaczące przyspieszenie procesu animacji przy jednoczesnym podniesieniu jej jakości wizualnej.

5.1. Implementacja zautomatyzowanego śledzenia celu przez głowę

Jedną z kluczowych modyfikacji ułatwiających proces animacji interakcji postaci z otoczeniem jest automatyzacja orientacji przestrzennej głowy. Klasyczne podejście wymaga ręcznego, dostosowywania obrotu kości szyi i głowy, aby utrzymać wzrok postaci na konkretnym obiekcie. W celu zwiększenia efektywności tego aspektu ruchu, zastosowano więzy kinematyczne automatyzujące śledzenie wyznaczonego punktu skupienia.

Głównym mechanizmem realizującym to założenie jest ograniczenie Track To, który przypisano bezpośrednio do kości głowy (Head). Prawidłowe działanie tego więzu wymagało stworzenie niezależnego punktu odniesienia w przestrzeni. Zdefiniowano w tym celu nową kość pomocniczą o nazwie HeadTarget, którą umiejscowiono horyzontalnie przed twarzą modelu postaci. Obiekt ten został wskazany w ustawieniach więzu jako główny cel śledzenia.

Precyzyjne zdefiniowanie wektorów kierunkowych było niezbędne do zachowania naturalnej orientacji modelu głowy. Parametr osi śledzenia (Track Axis), determinujący wektor, którym kość Head "patrzy" na cel, został przypisany do lokalnej osi Y. Z kolei wektor definiujący orientację pionową względem otoczenia (parametr Up) przypisano do osi Z, co skutecznie zapobiega nienaturalnemu przekrzywianiu się szyi wokół własnej osi podczas skrajnych wychyleń. Dodatkowo aktywowano funkcję Target Z, co stabilizuje rzutowanie wektorów.



Rys. 5. Więzy oraz kość pomocnicza do śledzenia celu przez głowę
Fig. 5. Constraint and helper bone for tracking a target with the head

Dla zapewnienia spójnego zachowania układu niezależnie od transformacji globalnych samej postaci, przestrzenie odniesienia dla obu stron więzu – zarówno dla obiektu śledzącego (Owner Space), jak i dla celu (Target Space) – ustawiono

na globalny układ współrzędnych (World Space). Całkowity wpływ więzu na orientację kości głowy (parametr Influence) ustawiono na maksymalną wartość 1.0. W rezultacie ruch kości HeadTarget w dowolnej osi powoduje płynne i automatyczne obracanie się głowy postaci, precyzyjnie symulując zjawisko podążania wzrokiem za poruszającym się obiektem. Na rysunku 5. przedstawiono zrzut ekranu ze środowiska Blender, ilustrujący widok stworzonej kości pomocniczej, odpowiedzialnej za sterowanie omówionym mechanizmem śledzenia.

5.2. Implementacja zautomatyzowanego systemu zginania palców dłoni

Animacja dłoni stanowi jeden z najbardziej wymagających etapów pracy nad ruchem postaci. Tradycyjne metody animowania, opierające się na ręcznym definiowaniu transformacji dla każdej kości, cechują się one bardzo wysoką pracochłonnością. W przypadku dłoni animator zmuszony jest do ciągłego kontrolowania dziesiątek punktów artykulacji, co przy standardowym układzie kinematyki prostej (FK) wymaga osobnego animowania każdego z trzech segmentów palca. W celu wyeliminowania tego problemu, zaimplementowano system automatyzujący zginanie palców w oparciu o propagację rotacji. Zastosowanie więzów kinematycznych (ang. Constraints) automatyzuje w tym przypadku określone zachowania układu kości, pozwalając na kontrolowanie całej struktury palca za pomocą pojedynczego elementu sterującego [7].

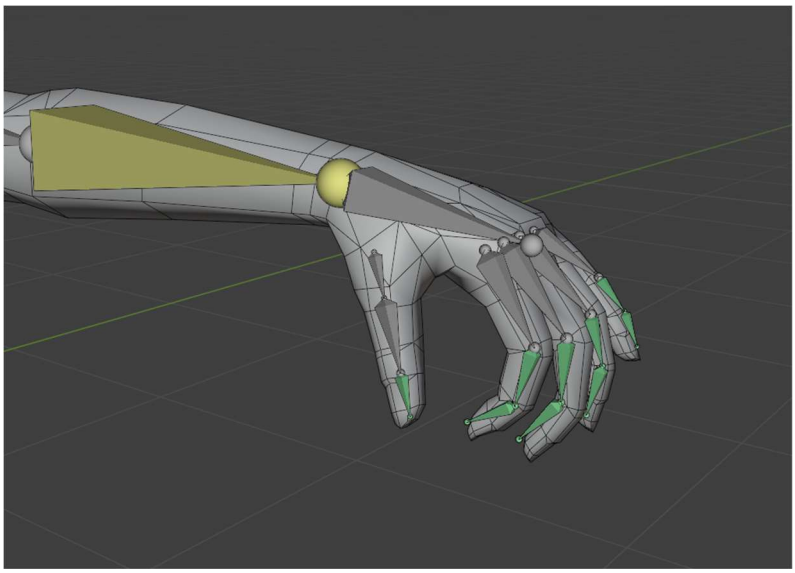
Mechanizm ten oparto na ograniczeniu Copy Rotation, który kaskadowo kopiuje wartości kątowe pomiędzy kolejnymi kośćmi w łańcuchu. Konfigurację przeprowadzono na lewej dłoni postaci. Analizując strukturę na przykładzie palca wskazywanego przez kości z przedrostkiem Pointer, proces rozpoczął się od przypisania ograniczenia do ostatniej kości (koniuszka palca) – Pointer3.L. Jako obiekt docelowy (ang. Target Bone) w ustawieniach więzu wskazano kość środkową – Pointer2.L.

Aby zapewnić prawidłową, anatomiczną deformację, kopiowanie rotacji ograniczono wyłącznie do osi Z (parametr Axis), która w lokalnym układzie współrzędnych odpowiadała za naturalne zginanie stawu. Kluczowym krokiem było ustawienie przestrzeni odniesienia dla obu elementów – zarówno obiektu śledzącego (Owner Space), jak i docelowego (Target Space) – na przestrzeń lokalną (Local Space). Parametr wpływu ograniczenia (Influence) ustawiono na maksymalną wartość 1.0.

Identyczny ciąg operacji zastosowano o stopień niżej w hierarchii. Do kości Pointer2.L dodano ograniczenie Copy Rotation o tożsamych parametrach (oś Z, Local Space, Influence równe 1), wskazując jako cel kość bazową palca – Pointer1.L.

Dzięki takiej konfiguracji uzyskany łańcuch zależności pozwala na sterowanie całym palcem wyłącznie poprzez rotację kości bazowej (Pointer1.L). Jej obrót jest natychmiastowo przekazywany do kości środkowej, która z kolei przekazuje go do kości końcowej, tworząc płynny, naturalny łuk zgięcia. Ten sam schemat

ograniczeń został powielony dla wszystkich pozostałych palców lewej dłoni: Middle, Index, Pinky oraz Thumb, a następnie, poprzez proces symetryzacji, przeniesiony na dłoń prawą. Na rysunku 6. zaprezentowano zrzut ekranu ze środowiska Blender, przedstawiający opisany układ kości dłoni.



Rys. 6. Kości palców z nałożonymi ograniczeniami
Fig. 6. Finger bones with applied contraints

5.3. System artykulacji stopy i mechanizm stanięcia na palcach

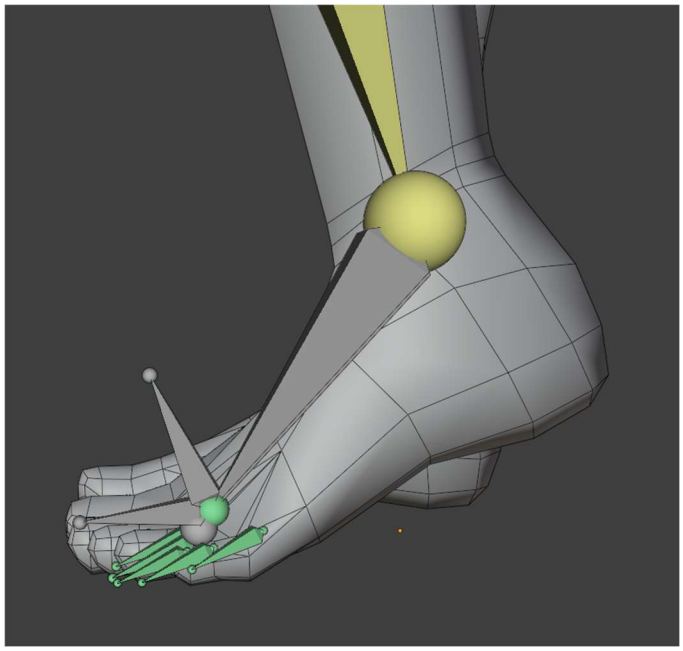
W celu osiągnięcia wysokiego realizmu chodu, standardowy układ kinematyczny stopy został poddany znacznej rozbudowie. Aby umożliwić niezależną i naturalną pracę palców stopy bez konieczności ich żmudnego, manualnego pozycjonowania, stworzono pięć dodatkowych kości sterujących, odpowiadających każdemu z palców. Kości te zostały włączone do układu hierarchicznego poprzez powiązanie ich z główną kością przodu stopy – Toe.L, która w tym układzie pełni funkcję elementu nadrzędnego (rodzica).

Zautomatyzowanie zgięcia zrealizowano poprzez przypisanie do każdej z pięciu nowych kości palców ograniczenia Copy Rotation, w którym jako obiekt docelowy (ang. Target Bone) wskazano bazową kość Toe.L. Aby zachować poprawny kierunek deformacji, kopiowanie rotacji aktywowano dla wszystkich trzech osi (X, Y, Z), jednakże dla osi X oraz Z zastosowano inwersję (parametr Invert). Przestrzenie odniesienia dla obu stron ograniczenia (Target Space oraz Owner Space) ustawiono na przestrzeń lokalną (Local Space).

Kluczowym zabiegiem pozwalającym na uzyskanie asymetrycznego, naturalnego rozkładu palców podczas zgięcia, było zróżnicowanie siły wpływu ograniczenia (parametr Influence). Wartość tę ustawiono na maksymalny poziom 1.0 dla największego palca, a dla każdego kolejnego sukcesywnie redukowano ją o 0.1 (osiągając odpowiednio wartości 0.9, 0.8, 0.7 i 0.6 dla najmniejszego palca). W efekcie, pojedynczy obrót kości Toe.L skutkuje płynnym, zgięciem wszystkich palców stopy.

Ostatnim elementem modyfikacji tego rejonu było wdrożenie mechanizmu umożliwiającego postaci płynne i stabilne stawanie na palcach. Nad stopą wygenerowano dodatkową kość pomocniczą nazwaną ToePivot.L. Aby zintegrować ją z wprowadzonym wcześniej (w rozdziale 4.1) systemem kinematyki odwrotnej, przedefiniowano relacje w łańcuchu. Główny kontroler nogi, kość FootIK.L, został przypisany jako dziecko nowej kości ToePivot.L. Jednocześnie, obiektem nadrzędnym (rodzicem) dla ToePivot.L stała się kość Toe.L.

Dzięki tak skonstruowanej, odwróconej hierarchii zależności, animator zyskuje możliwość wykonania stanięcia na palcach. Wykorzystując rotację pojedynczej kości ToePivot.L, uzyskuje się płynne uniesienie pięty przy jednoczesnym, stabilnym i realistycznym oparciu ciężaru na palcach, które – dzięki wcześniej opisanym ograniczeniom – automatycznie dostosowują swój kąt zgięcia. Na rysunku 7. przedstawiono zrzut ekranu ze środowiska Blender, prezentujący widok kości pomocniczych i sterujących w obrębie stopy.



Rys. 7. Dodatkowe kości dla stóp oraz kości sterujące

Fig. 7. Additional bones for the feet and control bones

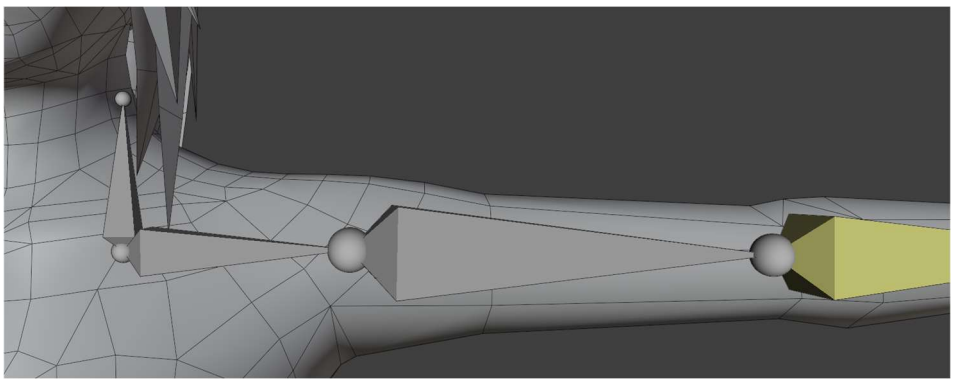
5.4. Modyfikacja układu barkowego i mechanizm wzruszania ramionami

Kolejnym istotnym zagadnieniem w kontekście animacji jest sterowanie barkiem. Naturalne ruchy w tym obszarze, takie jak wzruszanie ramionami czy sięganie w górę, wymagają jednoczesnej aktywacji obojczyka oraz kości ramiennej. W standardowym układzie wymusza to na animatorze niezależne pozycjonowanie obu tych elementów. W celu optymalizacji tego obszaru i zredukowania liczby niezbędnych operacji, przeprowadzono rekonfigurację bazowego łańcucha kinematycznego.

Modyfikacja polegała na implementacji dodatkowego kontrolera nadrzędnego. Nad obojczykiem stworzono nową kość pomocniczą o nazwie Shrug.L, pełniącą funkcję dedykowanego obiektu sterującego dla całego rejonu barku. Aby poprawnie zintegrować ten kontroler z głównym tułowiem postaci, przypisano go do hierarchii centralnej, definiując kość klatki piersiowej (Spine3) jako jego rodzica.

Kluczowym etapem modyfikacji była zmiana przypisań dla istniejących już kości deformujących siatkę. Dotychczasowa kość obojczyka, oznaczona jako Shoulder.L, została odpięta od swojego pierwotnego węzła i zdefiniowana jako element podrzędny (dziecko) nowej kości Shrug.L. Analogiczny zabieg zastosowano wobec kości ramiennej (Bic.L), która również została bezpośrednio podporządkowana kontrolerowi Shrug.L.

Dzięki tak zmodyfikowanej strukturze hierarchicznej, obrót oraz przesunięcie pojedynczej kości pomocniczej Shrug.L pociąga za sobą automatyczną i synchroniczną zmianę położenia zarówno obojczyka, jak i ramienia. Rozwiązanie to pozwala na intuicyjne generowanie ruchu wzruszania ramionami. Na rysunku 8. przedstawiono zrzut ekranu ze środowiska Blender, prezentujący zrekonfigurowany układ kości w obrębie lewego barku postaci.



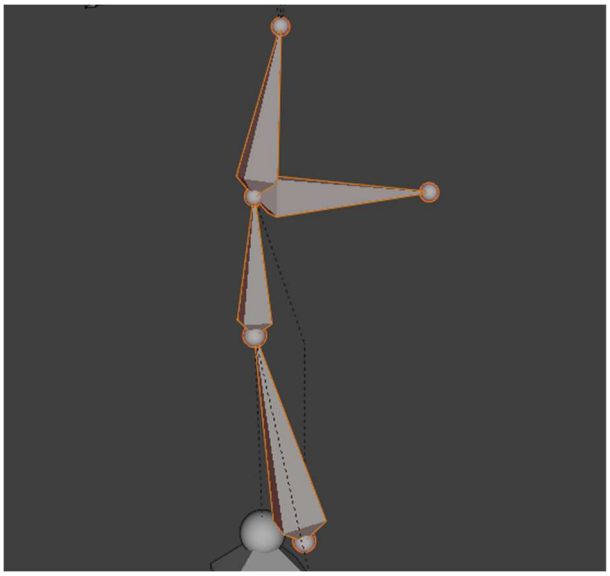
Rys. 8. Kość pomocnicza do wzruszania ramionami oraz jej powiązanie z ręką
Fig. 8. Helper bone for shrugging the shoulders and its connection to the arm.

5.5. Stabilizacja i sterowanie układem kręgosłupa

Ostatnim elementem omawianego zintegrowanego układu była modyfikacja centralnego łańcucha kinematycznego postaci, czyli kręgosłupa. Klasyczne operowanie zginaniem tułowia wyłącznie za pomocą kinematyki prostej (FK) wymaga ciągłego korygowania środka ciężkości. Aby zwiększyć efektywność tego procesu, wdrożono rozwiązanie wykorzystujące algorytmy Kinematyki Odwrotnej (Inverse Kinematics), dedykowane systemy ograniczeń (Constraints) oraz kontrolery pomocnicze.

Pierwszym krokiem było wygenerowanie dedykowanej kości sterującej o nazwie CoreIK. Umiejscowiono ją dokładnie na wysokości zakończenia kości Spine2. Zgodnie z przyjętymi założeniami projektowymi, wyłączono dla niej bezpośredni wpływ na deformację siatki geometrycznej modelu, nadając jej funkcję wyłącznie kontrolną. W strukturze hierarchicznej szkieletu kość CoreIK została zdefiniowana jako element podrzędny (dziecko) względem wyższej partii kręgosłupa, tj. kości Spine3.

Następnie przystąpiono do konfiguracji więzów kinematycznych, warunkujących poprawne zachowanie mechaniki tułowia. Do kości Spine2 przypisano ograniczenie Kinematyki Odwrotnej (Inverse Kinematics). Jako obiekt docelowy (ang. Target Bone) wskazano nowo utworzoną kość CoreIK. Parametr długości łańcucha (Chain Length) celowo ograniczono do wartości 1, definiując tym samym, że algorytm oddziałuje bezpośrednio i wyłącznie na kość Spine2. Siłę wpływu więzu (Influence) ustawiono na maksymalną wartość 1.0.



Rys. 9. Zmodyfikowany układ kości szkieletu

Fig. 9. Modified skeleton bone structure

Aby zapobiec nienaturalnemu rozdzielaniu się segmentów szkieletu podczas skrajnych transformacji i zachować integralność całego tułowia, wprowadzono dodatkowe ograniczenie wiążące. Na kość Spine3 nałożono ograniczenie Copy Location, wskazując element Spine2 jako jego cel. Kopiowanie pozycji aktywowano bezwzględnie we wszystkich trzech płaszczyznach przestrzennych (parametr Axis ustawiony na osie X, Y oraz Z), a współczynnik wpływu (Influence) również zdefiniowano jako 1.0.

Zastosowanie powyższej konfiguracji tworzy precyzyjne sprzężenie zwrotne pomiędzy środkową a górną partią tułowia postaci. Operowanie kontrolerem CoreIK wymusza zautomatyzowane ułożenie kręgosłupa. Wizualizację opisanego układu kości sterujących i ograniczeń w obszarze roboczym programu Blender przedstawiono na rysunku 9.

6. Podsumowanie i wnioski

W niniejszym rozdziale przeanalizowano problematykę optymalizacji procesu animacji postaci trójwymiarowych, ze szczególnym uwzględnieniem redukcji czasu pracy animatora oraz eliminacji najczęstszych błędów technicznych. Na tle klasycznych metod sterowania podejście oparte wyłącznie na kinematyce prostej (FK) bywa wysoce nieefektywne, wymuszając żmudną kontr-animację i często prowadząc do nienaturalnych artefaktów, takich jak zjawisko braku stabilności punktów podparcia (tzw. foot sliding).

Jako alternatywę zaproponowano i zaimplementowano zintegrowany workflow realizowany w środowisku Blender, polegający na głębokiej rozbudowie znormalizowanego szkieletu postaci. Zaprezentowane autorskie podejście do riggingu rozszerzyło bazową strukturę o system hybrydowy, integrujący łańcuchy Kinematyki Odwrotnej (IK) z zaawansowanymi więzami kinematycznymi (Constraints) oraz hierarchią kości pomocniczych.

Wdrożenie szczegółowo opisanych modyfikacji – w tym zautomatyzowanego śledzenia celu przez głowę (head tracking), mechanizmu kaskadowego zginania palców dłoni i stóp, systemu foot roll oraz intuicyjnego sterowania układem barkowym i kręgosłupem – diametralnie zmieniło ergonomię pracy z modelem. Liczba pojedynczych operacji (rotacji i translacji), które animator musi wykonać w celu uzyskania naturalnej pozy postaci, ulega drastycznej redukcji po wprowadzeniu tych modyfikacji. Odpowiednie przypisanie relacji nadrzędno-podrzędnych sprawia, że skomplikowane ruchy wielostawowe realizowane są za pomocą pojedynczych kontrolerów.

Mimo iż proces przygotowania tak zaawansowanego układu sterowania wymaga znacznego nakładu pracy, dużej precyzji w zachowaniu nazewnictwa oraz specjalistycznej wiedzy technicznej na etapie przedprodukcyjnym, wnioski z przeprowadzonych prac jednoznacznie wskazują na zasadność zastosowania tej techniki. Przeniesienie ciężaru dbałości o fizyczną i anatomiczną poprawność układu

z człowieka na algorytmy oprogramowania stanowi niezwykle efektywny kompromis. Zastosowana metodyka jest wysoce rekomendowana w profesjonalnych produkcjach (takich jak gry wideo czy animacje CGI), w których szybkość iteracji ruchu postaci musi iść w parze z jej bezbłędnym, fotorealistycznym odwzorowaniem [5].

W ujęciu praktycznym, przygotowania standardowego, 30-klatkowego cyklu chodu postaci. Przy wykorzystaniu bazowego, znormalizowanego szkieletu opartego wyłącznie na kinematyce prostej (FK), animator musi ręcznie korygować pozycję stóp w każdej klatce kluczowej, aby uniknąć efektu poślizgu (foot sliding) , a także niezależnie animować każdy ze stawów palców. Proces ten, wraz z późniejszym wygładzaniem krzywych animacyjnych, zajmuje średnio od 4 do 6 godzin pracy doświadczonego artysty. W przypadku zastosowania opisanego w rozdziale, autorskiego rigu z wdrożonymi systemami IK stóp, automatyzacją zgięcia palców (w tym systemem foot roll) oraz zintegrowanym ruchem kręgosłupa i głowy, czas wykonania tej samej sekwencji, przy zachowaniu wyższego poziomu realizmu biomechanicznego, skraca się do zaledwie 1,5-2 godzin. Eliminuje to konieczność żmudnej kontr-animacji , co wskazuje na potencjalnie wyższą opłacalność produkcyjną wdrożenia zaproponowanego rozwiązania. Przedstawione wartości czasowe mają charakter orientacyjny i wynikają z obserwacji autorów przeprowadzonych podczas testowej realizacji cyklu chodu na opracowanym modelu.

Literatura

[1] Aristidou A., Lasenby J.: FABRIK: A fast, iterative solver for the Inverse Kinematics problem, Graphical Models, vol. 73, no. 5, 2011, s. 243–260.

[2] Beane A.: 3D Animation Essentials, John Wiley & Sons, Indianapolis 2012.

[3] Blain J.M.: The Complete Guide to Blender Graphics: Computer Modeling & Animation, CRC Press, Boca Raton 2019.

[4] Kerlow I.V.: The Art of 3D Computer Animation and Effects, John Wiley & Sons, Hoboken 2004.

[5] Kuhn C.: Mind-Melding Unity and Blender for 3D Game Development, Apress, Nowy Jork 2015.

[6] Parent R.: Computer Animation: Algorithms and Techniques, Morgan Kaufmann, Burlington 2012.

[7] Valenza E.: Blender 3D Cookbook, Packt Publishing, Birmingham 2015.

[8] Villar O.: Learning Blender: A Hands-On Guide to Creating 3D Animated Characters, Addison-Wesley Professional, Boston 2017.

[9] Welman C.: Inverse Kinematics and Geometric Constraints for Articulated Figure Manipulation, Simon Fraser University, Burnaby 1993.

[10] Dokumentacja techniczna oprogramowania Blender: Blender 5.0 Reference Manual https://docs.blender.org/manual/ (dostęp: 25 lutego 2026 r.).

A METHODOLOGY FOR THE DESIGN OF CONTROL SYSTEMS IN 3D CHARACTER SKELETAL ANIMATION

Summary

In the era of rapid development of computer graphics, particularly in the video game and virtual reality sectors, one of the key challenges is the development of animation methods that combine high realism with efficient production processes. The aim of this paper was to develop and present an original methodology for designing control systems (rigging) for 3D characters, representing an effective compromise between motion fidelity and the animator's workload. The analysis focuses on the problem of the high time consumption and limited ergonomics of classical animation techniques. Traditional approaches based on direct manipulation of mesh geometry or the exclusive use of Forward Kinematics are often associated with technical issues such as unnatural joint deformations, instability of support points (the "foot sliding" phenomenon), and the need for time-consuming counter-animation.

Based on a critical analysis of the evolution of animation systems, an integrated workflow implemented in Blender was proposed. The developed solution extends standard skeletal animation with a hybrid control system integrating Inverse Kinematics chains, advanced constraints, and auxiliary orientation controllers. This approach represents a modern alternative to manual bone transformations, transferring the responsibility for maintaining physical correctness from the animator to control algorithms. Although the implementation of this methodology increases the complexity of the rigging stage, it significantly improves animation efficiency. Automation of key motion aspects reduces the number of corrective operations and helps eliminate visual artifacts, enabling creators to focus more on the artistic quality of the animation.

Keywords: 3D character animation, skeletal animation, rigging, animation workflow, inverse kinematics, constraints.

Maja Jaszowska, Natalia Klocek, Sebastian Cwynar,
Julia Ziąbek, Dominik Dworak
Koło Naukowe Elektroniki i Technologii Informacyjnych

dr inż. Bartosz Pawłowicz
opiekun Koła Naukowego

BEZPIECZEŃSTWO IOT W SYSTEMACH KRYTYCZNEJ INFRASTRUKTURY, SIEĆ LOKALNA

Streszczenie

Celem rozdziału jest praktyczna analiza bezpieczeństwa domowych urządzeń Internetu Rzeczy (IoT) oraz ocena ich komunikacji w sieci lokalnej pod kątem podatności na nieautoryzowany dostęp. Obiektem badań uczyniono popularny mikrokontroler ESP8266 z oprogramowaniem open-source WLED, wykorzystywany do sterowania oświetleniem LED. Metoda badawcza opiera się na przeprowadzeniu aktywnego rekonesansu programem Nmap oraz wykonaniu ataku ARP Spoofing (Man-in-the-Middle). Wykorzystano do tego narzędzia Ettercap i Wireshark w celu przechwycenia i analizy ruchu sieciowego. W toku przeprowadzonych badań zidentyfikowano szereg krytycznych luk bezpieczeństwa, w tym całkowity brak uwierzytelniania w panelu sterowania oraz przesyłanie danych sterujących i poświadczeń (loginów i haseł do brokera MQTT) jawnym tekstem. Wykazano, że komunikacja oparta na niezabezpieczonym protokole WebSocket umożliwia łatwe odczytanie struktury poleceń JSON. Wyniki eksperymentu potwierdziły możliwość pełnego, nieautoryzowanego przejęcia kontroli nad urządzeniem poprzez replikację ataku w programie Postman, a skanowanie Nmap ujawniło dodatkową podatność związaną z błędną konfiguracją nagłówków CORS. Głównym wnioskiem płynącym z pracy jest stwierdzenie, że urządzenia IoT bez wdrożonego szyfrowania WSS/HTTPS stanowią realne zagrożenie dla bezpieczeństwa sieci. Jako skuteczne metody mitygacji zagrożeń zaproponowano modyfikację kodu źródłowego firmware'u w celu implementacji protokołów szyfrowanych, usunięcie niebezpiecznych nagłówków oraz zastosowanie logicznej izolacji urządzeń poprzez segmentację sieci VLAN i konfigurację reguł zapory sieciowej.

Słowa kluczowe: Internet Rzeczy, ESP8266, cyberbezpieczeństwo, ARP Spoofing, WebSocket.

1. Wprowadzenie

Współczesny świat w coraz większym stopniu opiera się na technologiach informacyjnych. Najważniejszą koncepcją, która wyznacza kierunek rozwoju branż w XXI wieku, jest Internet Rzeczy (IoT) (ang. Internet of Things).

IoT można zdefiniować jako globalną infrastrukturę dla społeczeństwa informacyjnego, która umożliwia dostęp do zaawansowanych usług poprzez połączenie fizycznych lub wirtualnych przedmiotów, bazujące na istniejących i rozwijanych technologiach teleinformatycznych (ICT) [7]. Koncepcja ta polega na przekształcaniu codziennych przedmiotów w „inteligentne rzeczy", które potrafią reagować na środowisko, przetwarzać, przechowywać i przesyłać informacje cyfrowe do innych obiektów za pośrednictwem protokołów internetowych [1].

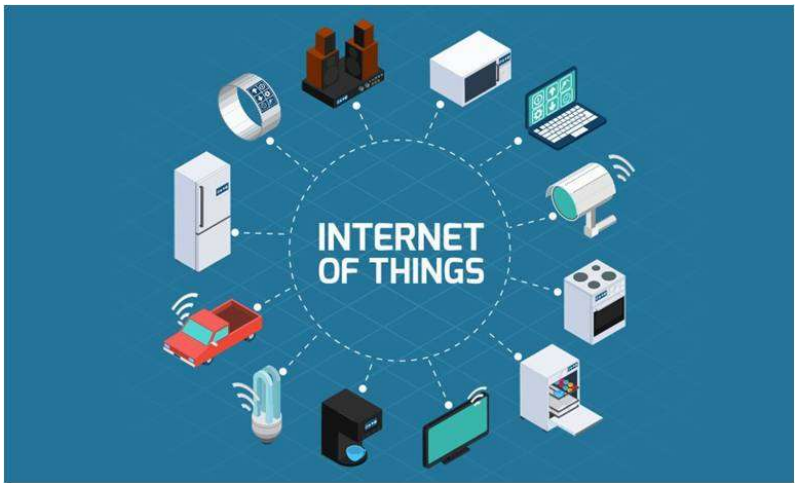
Dynamiczny rozwój IoT zrewolucjonizował sposób, w jaki ludzie ułatwiają sobie rutynowe czynności. Urządzenia IoT (np. inteligentne gniazdka, czujniki) są powszechne w domach, zapewniając wygodę i automatyzację. Pomimo korzyści, szybka adopcja tych technologii powiększa ilość potencjalnych wektorów ataku[6]. Każde urządzenie w sieci może być celem hakerów. Wykorzystując luki, mogą oni przejąć kontrolę, kraść dane lub przeprowadzać ataki typu DDoS. Systemy IoT są często projektowane z myślą o niskich kosztach produkcji, a nie o zabezpieczeniach przed atakującymi.

Celem pracy jest praktyczna analiza bezpieczeństwa domowych urządzeń Internetu Rzeczy (IoT). Obejmuje to ocenę, jak w rzeczywistości wygląda ich komunikacja w sieci oraz czy jest ona wystarczająco zabezpieczona, ze szczególnym uwzględnieniem zagrożeń wynikających z wykorzystania niezabezpieczonych protokołów komunikacyjnych.

2. Internet Rzeczy – IoT

Internet Rzeczy (IoT) jest siecią fizycznych urządzeń, które mogą komunikować się między sobą, systemami chmurowymi oraz innymi urządzeniami w danej sieci. Internet jest wykorzystywany w każdym aspekcie życia, a urządzenia smart pomagają w ułatwianiu codziennych czynności. Takimi urządzeniami można zdalnie sterować i dostosowywać ich działanie.

W odróżnieniu od tradycyjnego Internetu, którego użytkownikami byli głównie ludzie, IoT tworzy środowisko, w którym maszyny komunikują się z maszynami (M2M). Z odpowiednim oprogramowaniem są one w stanie samoistnie podejmować działania na podstawie zebranych informacji [6]. Rysunek 1. przedstawia urządzenia które są określane jako Internet Rzeczy. IoT to przede wszystkim danie cyfrowej tożsamości i możliwości komunikacji technologiom – od zwykłej żarówki, przez maszyny w przemyśle, po pojazdy. Świat fizyczny staje się inteligentny i pomaga użytkownikom w codzienności.



Rys. 1. Urządzenia określane jako Internet Rzeczy
Fig. 1. Devices known as the Internet of Things[3]
Źródło: https://www.startertutorials.com/blog/introduction-to-iot.html

2.1. Zastosowanie i Wyzwania IoT

Potencjał technologii IoT jest ogromny i jest stosowany w niemal każdej dziedzinie życia codziennego. Głównym celem jest poprawa jakości usług świadczonych ludziom.

Obszary zastosowania urządzeń Internetu Rzeczy:

1. Przemysł i Produkcja: Optymalizacja procesów i zwiększanie wydajności[7]. Systemy diagnostyki maszyn technologicznych, które identyfikują nieprawidłową pracę [1].
2. Inteligentne Miasta (Smart City): Rozwiązania infrastruktury publicznej. Monitorowanie sieci energetycznych i wodno-kanalizacyjnych w celu efektywniejszego wykorzystania zasobów [7].
3. Opieka Zdrowotna: Telemedycyna i zdalne monitorowanie stanu pacjentów [6].
4. Inteligentny Dom: Możliwość podłączenia elementów codziennego życia (np. pralki, oświetlenia) do globalnej sieci w celu zdalnej kontroli [6].

Internet Rzeczy (IoT) łączy wiele urządzeń, umożliwiając gromadzenie i wymianę danych. Choć zastosowania IoT poprawiają jakość usług i redukują koszty operacyjne, niosą ze sobą poważne wyzwania związane z bezpieczeństwem i prywatnością[1]. Eksperci uznają bezpieczeństwo za najważniejsze wyzwanie dla dynamicznego rozwoju IoT [7]. Problemy z bezpieczeństwem IT z przeszłości powracają w nowych urządzeniach, dając hakerom wiele możliwości cyberataków [6].

Główne przyczyny trudności w zarządzaniu bezpieczeństwem IoT:

1. Ograniczone Zasoby: Wiele urządzeń ma ograniczone zasoby (moc, pamięć), co utrudnia wdrożenie zaawansowanych systemów ochrony. Tradycyjne metody są często niewystarczające [6].
2. Słabe Zabezpieczenia Domyślne: Producenci stosują łatwe uwierzytelnianie, a dane logowania są często na stałe w oprogramowaniu (hardcoded).
3. Brak Aktualizacji: Część urządzeń IoT nie może być w ogóle aktualizowana [2].
4. Brak Szyfrowania: Urządzenia nie szyfrują komunikacji z Internetem i sieciami lokalnymi. Komunikacja na sieciach bezprzewodowych umożliwia podsłuch przesyłanych danych [6].
5. Zagrożenia dodatkowe: Wynikają z braku monitorowania urządzeń oraz niskiego poziomu świadomości cyberbezpieczeństwa [4].

Ogólna architektura systemu IoT jest dzielona na trzy główne warstwy: warstwę percepcji, warstwę sieci i warstwę aplikacji. Poprawa bezpieczeństwa wymaga uwzględnienia luk i potencjalnych ataków we wszystkich tych warstwach.

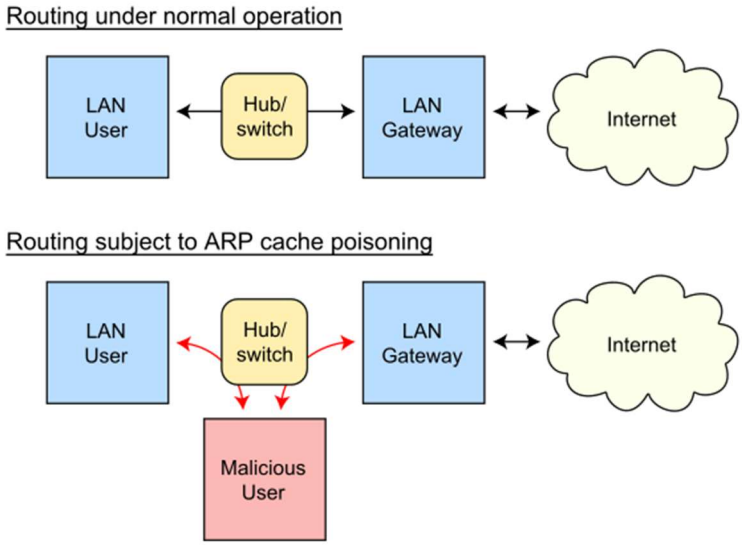
3. ARP Spoofing

Niniejszy rozdział przedstawia mechanizm ataku ARP Spoofing, który stanowi podstawę metody badawczej zastosowanej w pracy. Omówiono w nim zasadę działania protokołu ARP oraz sposób jego wykorzystania do przechwycenia ruchu w sieci lokalnej.

ARP Spoofing, znany również jako zatruwanie pamięci podręcznej ARP (ARP Cache Poisoning) [8], jest jedną z najbardziej podstawowych, ale jednocześnie niebezpiecznych metod ataków w lokalnych sieciach LAN [3]. Z powodu fundamentalnych słabości protokołu ARP, atak ten jest często punktem wyjścia do bardziej zaawansowanych działań, takich jak ataki typu Man-in-the-Middle (MiM) lub Denial of Service (DoS) [5]. Protokół ARP służy do rozwiązywania adresu MAC hosta (adresu warstwy łącza danych) na podstawie znanego adresu IP (adresu warstwy sieciowej) [3]. Rysunek 2. przedstawia uproszczony schemat ARP Spoofing.

Gdy host chce komunikować się z innym hostem, którego adresu MAC nie zna, 13 Zakład Systemów Złożonych wysyła pakiet żądania ARP (ARP request), który jest rozgłaszany (broadcast) w całej sieci. Host posiadający ten adres IP odpowiada pakietem odpowiedzi ARP (ARP reply), zawierającym swój adres MAC, który jest wysyłany bezpośrednio do nadawcy (unicast) [3].

Wszystkie hosty utrzymują pamięć podręczną ARP (ARP cache), czyli tabelę przechowującą mapowania adresów IP na odpowiadające im adresy MAC, których nauczono się dynamicznie z sieci lub skonfigurowano statycznie. Wpisy dynamiczne wygasają po ustalonym czasie, różnym w zależności od systemu operacyjnego, i są usuwane [5].



Rys. 2. Skuteczny atak ARP Spoofing umożliwia przeprowadzenie ataku typu man-in-the-middle
Fig. 2. A successful ARP Spoofing attack enables a man-in-the-middle attack[5]
Źródło: https://en.wikipedia.org/wiki/ARP_spoofing

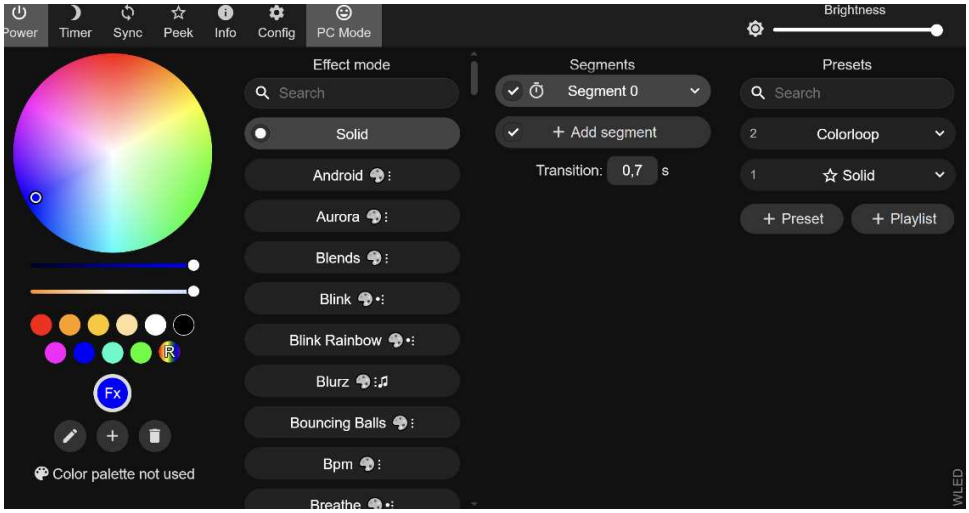
ARP spoofing to proces fałszowania pakietów ARP w celu podszycia się pod innego hosta w sieci. Atak polega na złośliwym wprowadzeniu fałszywego mapowania adresu IP na adres MAC do pamięci podręcznej ARP hosta docelowego [8].

4. Badania

Mikrokontroler ESP8266 jest urządzeniem znajdującym się w lokalnej sieci IoT. To dość popularny sprzęt wśród osób, pasjonujących się robotyką i technologią IoT. W badanym przypadku, służy do sterowania oświetleniem w jednym z pokoi. Oświetlenie stanowi długi pasek diod LED RGB.

ESP hostuje serwer webowy. Będąc w tej samej sieci lokalnej, wystarczy wprowadzić w przeglądarce adres http:[adres_IP_ESP], po to aby dostać dostęp do panelu sterowania (UI). Serwer wyświetla stronę na której można ustawiać kolor (w przestrzeni RGB), jasność i układ kolorystyczny oświetlenia. Na mikrokontrolerze działa oprogramowanie open-source WLED.

Na rysunku 3. znajduje się wspomniany wcześniej interfejs webowy, za pomocą którego można sterować oświetleniem podłączonym do ESP8266. Głównym celem badania jest pokazanie ryzyka związanym z użytkowaniem niezabezpieczonych urządzeń IoT w zaufanej sieci domowej.



Rys. 3. Panel sterowania (UI) oświetleniem WLED hostowany na ESP8266

Fig. 3. WLED lighting control interface (UI) hosted on an ESP8266 microcontroller

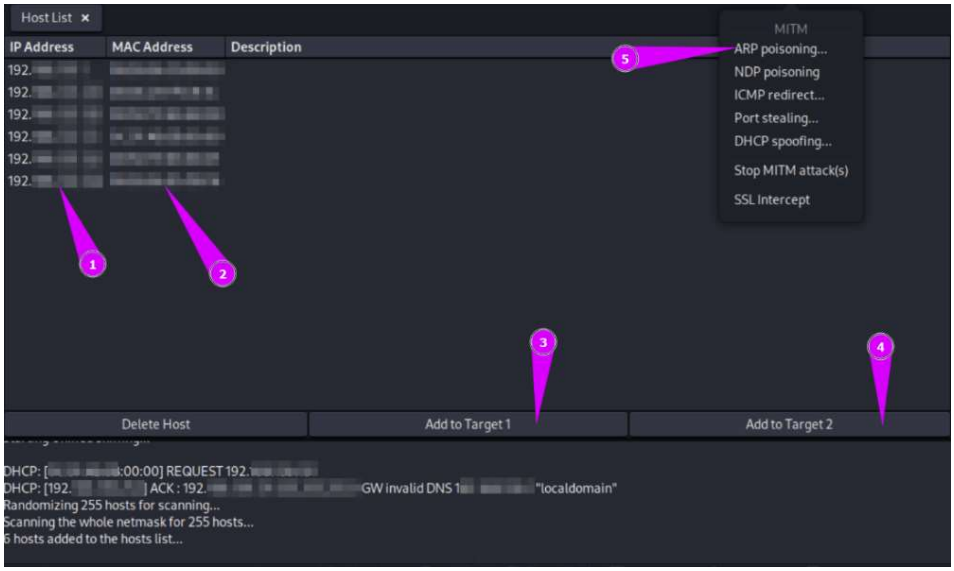
4.1. Scenariusz Ataku: ARP Spoofing (Man-in-the-Middle)

Scenariusz: Nieświadomy użytkownik (ofiara) steruje oświetleniem LED za pomocą swojego komputera. Do komunikacji używa ESP8266 przez UI. Atakujący, znajdujący się w tej samej sieci, wykorzystuje atak typu ARP Spoofing, czyli ataku Man-in-the-Middle (MITM). Celem jest przechwycenie i analiza tej komunikacji, a następnie jej odtworzenie w celu zdalnego przejęcia kontroli nad ESP:

1. Ofiara (Użytkownik): Laptop z systemem Windows 10.
2. Cel (Urządzenie IoT): Mikrokontroler ESP8266.
3. Atakujący: Laptop z systemem Kali Linux.

Do przeprowadzenia ataku wybrano oprogramowanie Ettercap. Oprogramowanie jest darmowe i proste w obsłudze. Pozwala na przeprowadzenie zautomatyzowanego ataku ARP bez potrzeby specjalistycznej konfiguracji. Jest w stanie aktywnie skanować sieć, aby identyfikować hosty i automatycznie rozpocząć atak ARP.

Na rysunku 4. przedstawiono wyniki skanowania sieci przez Ettercap. Program wyświetla listę wykrytych hostów, ich adresy IP (1) oraz ich adresy fizyczne MAC (2). Przed całym atakiem, konieczne jest ustawienie celów (wektory ataku). Target 1 (3): Adres IP routera. Target 2 (4): Reszta hostów w sieci (czyli ofiary lub po prostu inne urządzenia). Po kliknięciu "Start"(5), Ettercap rozpoczyna wysyłanie pakietów ARP. Więc zarówno router, jak i pozostałe urządzenia, zostają przekonane, że adres MAC atakującego należy do wszystkich innych hostów. Cały ruch sieciowy z routera będzie kierowany na Kali Linuxa i podsłuchiwany przez atakującego. Do analizy przechwyconych pakietów wykorzystano program Wireshark.



Rys. 4. Skanowanie sieci i konfiguracja celów ataku w Ettercap

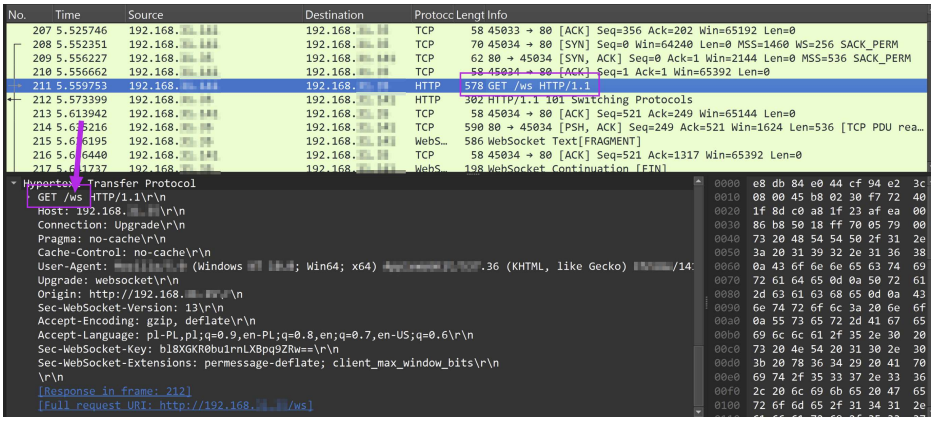
Fig. 4. Network scanning and configuring attack vectors in Ettercap

4.2. Analiza Luki i Przechwyconego Ruch

Pierwszą luką systemu jest brak uwierzytelniania dostępu do panelu sterowania ESP8266. Urządzenie nie wymaga loginu, hasła, ani żadnego tokenu dostępu. Dostęp opiera się wyłącznie na zasadzie "security by obscurity"(bezpieczeństwo przez ukrycie). Opiera się ona na założeniu, że nikt niepowołany nie zna adresu IP ESP oraz nie posiada dostępu do sieci lokalnej. Każda osoba podłączona do tej samej sieci, może zeskanować sieć, znaleźć adres IP i zdobyć kontrolę nad LEDami.

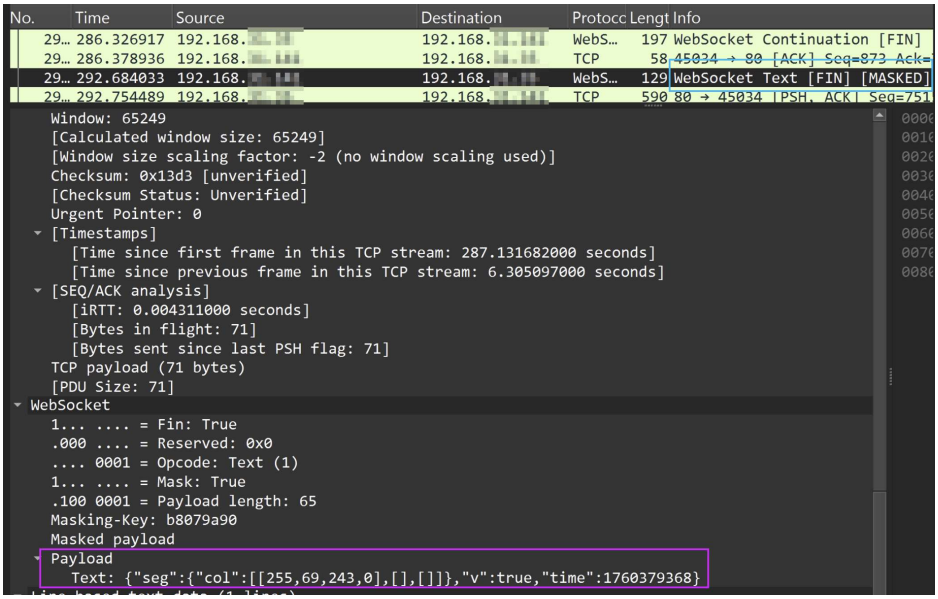
Drugą luka jest realizacja komunikacji między UI, a ESP za pomocą protokołu WebSocket. Zamiast bezpiecznego, szyfrowanego odpowiednika Secure WebSocket. Protokół WebSocket jest używany do utrzymania stałego połączenia między przeglądarką, a serwerem. Pozwala to na szybkie przesyłanie poleceń. Ponieważ komunikacja odbywa się w formie jawnego tekstu (plaintext), atakujący w pozycji Man-in-the-Middle może odczytać całą zawartość przesyłanych pakietów.

Najważniejsza informacja została znaleziona w pakiecie inicjującym połączenie Websocket Text. Jak widać na rysunku 5., przeglądarka użytkownika (klient), wysyła standardowe żądanie HTTP GET z nagłówkiem Upgrade: websocket. Jest to mechanizm handshake protokołu WebSocket. Żądanie to jest kierowane na konkretną ścieżkę na serwerze: http:IP_ESP/ws. Eendpoint usługi WebSocketjest pod adresem /ws. Bez tej informacji niemożliwe byłoby nawiązanie połączenia z serwerem WebSocket. Przy każdej próbie połączenia pojawiałby się błąd.



Rys. 5. Analiza pakietu HTTP GET inicjującego połączenie WebSocket

Fig. 5. Analysis of the HTTP GET packet initiating the WebSocket connection



Rys. 6. Analiza zawartości (payload) pakietu WebSocket z danymi koloru (JSON)

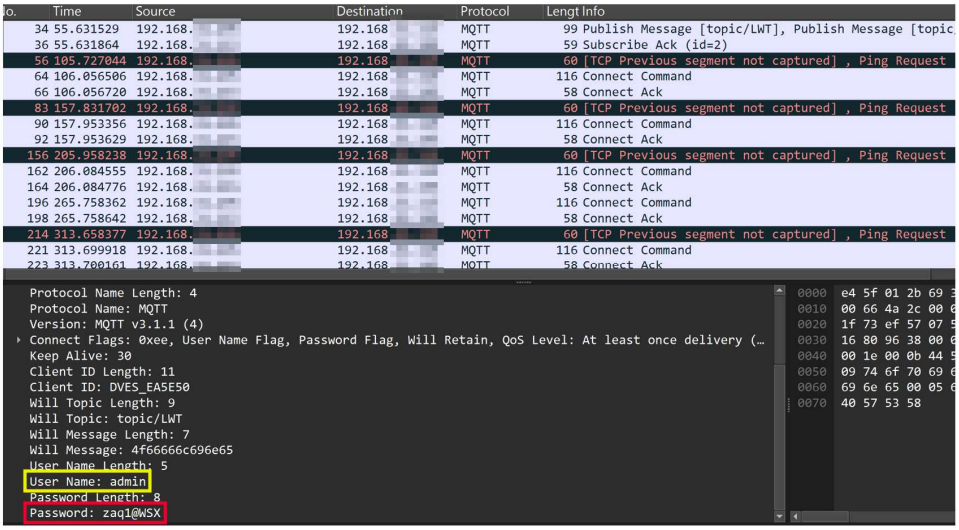
Fig. 6. Analysis of the WebSocket packet payload containing color data (JSON)

ESP odbiera wiele informacji dotyczących parametrów, które mają zostać ustawione dla oświetlenia. Nie tylko kolor, ale także nasycenie i moc światła monochromatycznego. Jak widać na obrazku, pierwsza wiadomość przesyłana przez WebSocket zawiera dane o kolorze. Można to wywnioskować z faktu, że przesyłana jest tablica z trzema liczbami, co bardzo przypomina sposób definiowania kolorów w modelu RGB.

Analiza struktury tego JSON-a pozwala na dokładne zrozumienie poleceń:

- "seg": Obiekt definiujący parametry dla segmentu paska LED,
  - "col": Tablica definiująca kolory.
    - [255, 69, 243]: Tablica wartości RGB (Red=255, Green=69, Blue=243).
    - 0: Wartość dla kanału White.
    - [][][]: Puste tablice zarezerwowane na dodatkowe definicje efektów.
- "v": true: Flaga "Visible", oznaczająca, że światła mają być włączone (true) lub wyłączone (false).
- "time": 17600379368: Znacznik czasowy wysłania polecenia.

Wiele urządzeń IoT wykorzystuje protokół MQTT (Message Queuing Telemetry Transport). Analiza ruchu w Wireshark bardzo szybko ujawniła pakiety tego protokołu.



Rys. 7. Przechwycony pakiet MQTT CONNECT z danymi uwierzytelniającymi plain text

Fig. 7. Captured MQTT CONNECT packet with plain text credentials

Jak widać na rysunku 7., pakiet MQTT CONNECT (służący do nawiązania sesji z brokerem), przesyłany jest jako plain text. Nie ma tutaj szyfrowania, W nim widoczne są krytyczne dane uwierzytelniające: login oraz hasło do brokera MQTT.

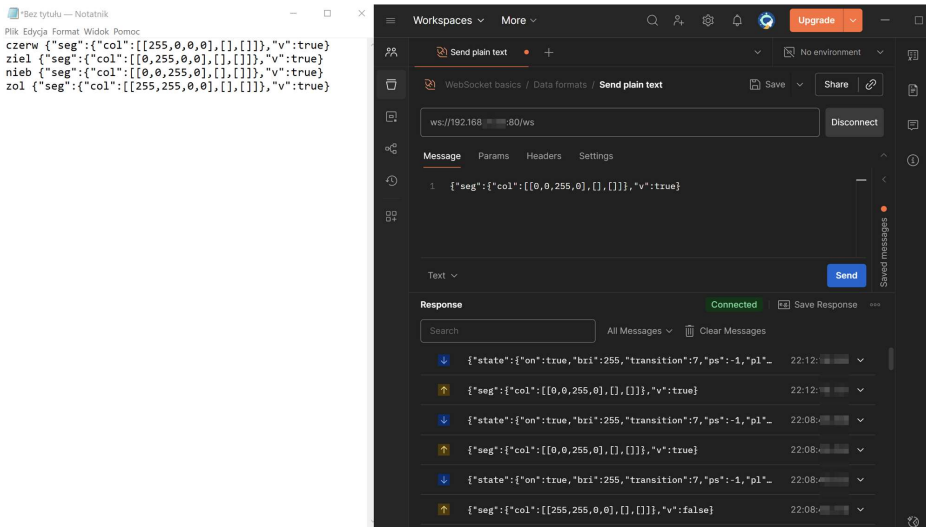
Broker jest centralnym punktem systemu, który odpowiada za wymianę wszystkich informacji (statusów, poleceń) między urządzeniami. Po dłuższej obserwacji pakiet nie pojawia się tylko przy pierwszym połączeniu, ale jest wysyłany cyklicznie, co pewien czas. Prawdopodobnie w celu odnowienia wygasającej sesji.

Ujawnienie tych danych stanowi krytyczną lukę bezpieczeństwa. Atakujący, który je przechwyci, zyskuje możliwość podłączenia się do brokera, subskrybowania wszystkich wiadomości przesyłanych w systemie, a także publikowania własnych, sfałszowanych poleceń. Natomiast samo hasło i login nie jest wymagane przy sterowaniu światłami. Dzieje się tak ponieważ, samo sterowanie odbywa się po web serwerze bez autentykacji. Czyli każda osoba z sieci jest w stanie sterować światłami.

4.3. Replikacja Ataku

Analiza pakietów pokazała, że oprócz podstawowych danych o kolorze, przez WebSocket przesyłane są również bardziej złożone obiekty JSON. Informują one ESP8266 o parametrach takich jak nasycenie (sat), jasność (bri) oraz identyfikatory (fx) i parametry (sx, ix) efektów specjalnych (np. "fade", "blink"). Dane są "kosmetyczne" i nie są niezbędne do samego włączenia i sterowaniem LED.

Po wpisaniu poprawnie skonfigurowanego adresu docelowego. Po kliknięciu "Connect", Postman nawiązuje i podtrzymuje sesję WebSocket z mikrokontrolerem. W tej konfiguracji możliwe jest wpisywanie ładunku (payload) w polu "Message" i wysyłanie ich do ESP.



Rys. 8. Wysyłanie spreparowanych ładunków JSON przejmujących kontrolę nad oświetleniem

Fig. 8. Sending crafted JSON payloads to take control of the lighting

Rysunek 8. przedstawia interfejs wysyłania wiadomości w Postman. Na zrzucie ekranu widoczne są przykładowe ładunki (payloady), które po kliknięciu "Send" zostały pomyślnie wykonane przez ESP8266. Chociaż przedstawienie praktycznego, fizycznego rezultatu (zmiany światła) nie jest możliwe w formie

niniejszego rozdziału, przeprowadzone badanie wykazało, że jest możliwe przejęcie zdalnej i nieautoryzowanej kontroli nad urządzeniem IoT.

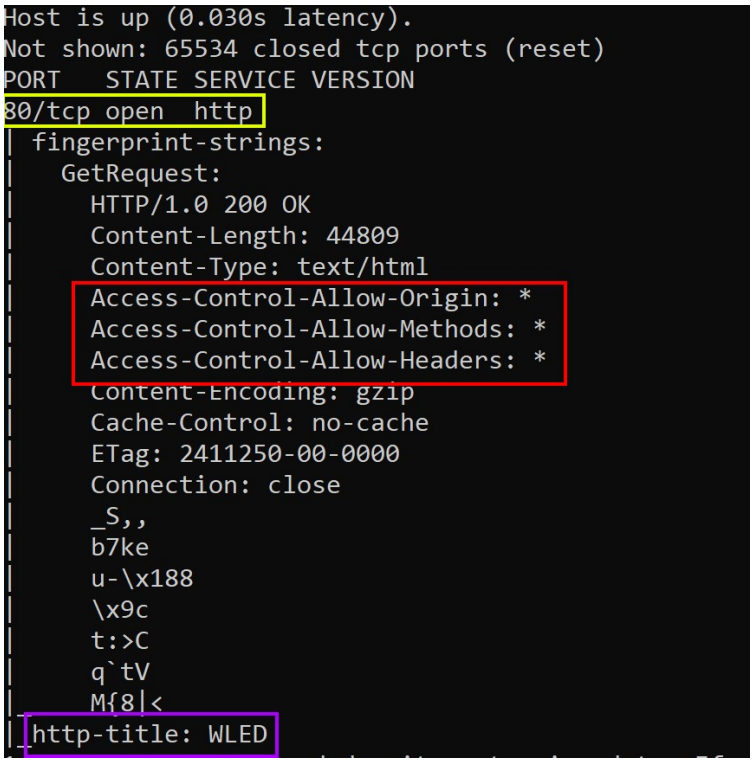
5. Skanowanie NMAP

Po analizie pasywnej (podsłuchu ruchu) przeprowadzono dodatkowy rekonesans za pomocą programu Nmap. Celem było zebranie szczegółowych informacji o urządzeniach. Przede wszystkim o otwartych portach, działających oraz wersjach oprogramowania. Byłby to ważny krok ze strony atakującego, ponieważ pozwala on na mapowanie powierzchni ataku.

Użyto polecenia: nmap -A -p- -T4 [ADRES_IP]

Znaczenie flag:

- -A: włącza "tryb agresywny", który obejmuje wykrywanie systemu operacyjnego (-O), wykrywanie wersji usług (-sV), skanowanie skryptowe (-sC) oraz traceroute.
- -p-: skanuje wszystkie 65 535 portów TCP.
- -T4: ustawia szablon czasowy, aby przyspieszyć skanowanie.



Rys. 9. Wynik skanowania NMAP dla ESP8266

Fig. 9. NMAP scan result for ESP8266

Informacje uzyskane ze skanowania:

1. Otwarty Port: Skan wykazał, że host jest aktywny i ma otwarty port 80/tcp (HTTP).
2. Identyfikacja Stosu Sieciowego: Nmap zidentyfikował system operacyjny jako lwIP 2.1.0 - 2.2.0. lwIP (lightweight IP) to popularny, lekki stos TCP/IP.
3. W odpowiedzi serwera Nmap przechwycił nagłówek: Access-Control-AllowOrigin: *. Ta konfiguracja CORS (Cross-Origin Resource Sharing) informuje przeglądarki internetowe, że dowolna strona internetowa z dowolnej domeny (oznaczonej *) ma prawo wysyłać żądania do tego urządzenia.
4. Adres MAC został zidentyfikowany jako Espressif, co potwierdza producenta chipu.

Skan Nmap potwierdził oprogramowanie WLED, co uwiarygodniło całą poprzednią analizę. Co więcej, skan ten odkrył zupełnie nową podatność (błędną konfigurację CORS), która nie była widoczna podczas pasywnej analizy, otwierając tym samym nowy wektor ataku.

6. Propozycje Zabezpieczeń

Analiza wykazała wiele luk w zabezpieczeniach IoT. Poniżej przedstawiono praktyczne metody i konfiguracje, które można zastosować, aby wyeliminować lub znacząco zminimalizować zidentyfikowane ryzyka.

Zidentyfikowano trzy główne problemy:

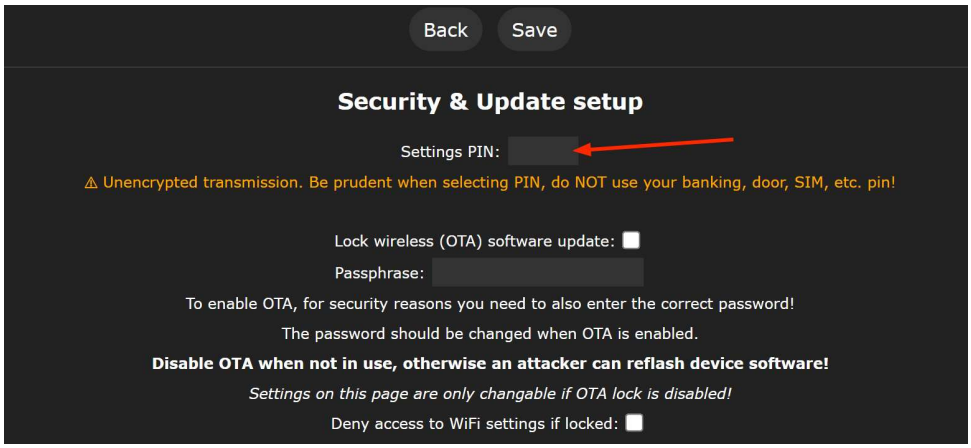
1. Brak uwierzytelniania: Każda osoba posiadająca adres IP urządzenia może sterować oświetleniem.
2. Nieszyfrowana komunikacja: Sterowanie odbywa się poprzez nieszyfrowany protokół WebSocket (‘ws://‘).
3. Podatność na ataki międzywitrynowe: wykryta podczas skanowania Nmap, pozwala innej stronie internetowej na sterowanie mikrokontrolerem.

W panelu sterowania oprogramowania WLED można ustawić kod PIN w celu zabezpieczenia dostępu. Niestety, jest to jedynie powierzchowne zabezpieczenie skuteczne przeciwko osobom nietechnicznym, a nie przeciwko atakującemu.

Rysunek 10. przedstawia panel ustawień, w którym można aktywować PIN. Ze względu na brak szyfrowania komunikacji kod PIN może być łatwo przechwycony w taki sam sposób, w jaki przechwycono inne pakiety (metodą Man-in-the-Middle).

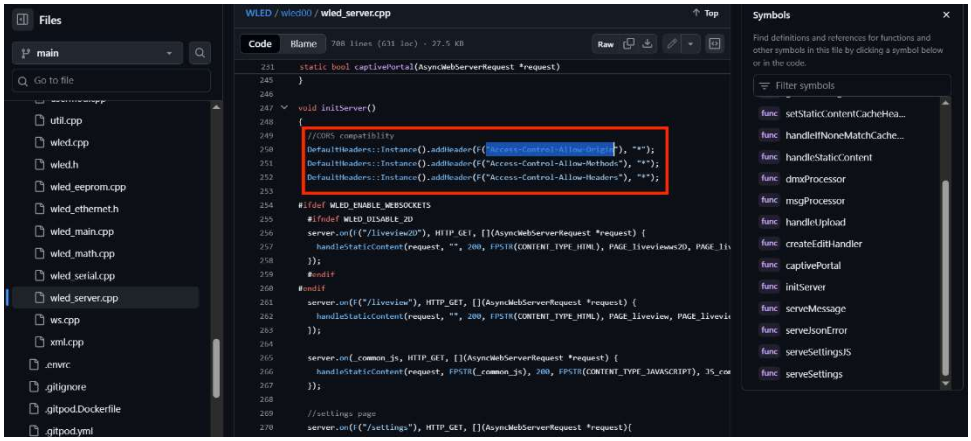
Oprogramowanie WLED, podobnie jak Tasmota, jest projektem open-source. Więc jest możliwe pobranie kodu źródłowego i stworzenie własnego, zmodyfikowanego firmware’u. Repozytorium projektu WLED na GitHubie, pozwala na pobranie kodu źródłowy. Po modyfikacji odpowiednich plików kodu, możliwe jest wdrożenie komunikacji HTTPS i WSS. WSS to szyfrowana wersja protokołu WebSocket, działająca w oparciu o TLS (podobnie jak HTTPS). Zmiana prefiksu

połączenia z ws:// na wss:// uniemożliwiłaby przechwycenie danych w ataku Man-in-the-Middle. Dopiero po tym, ustawienie PIN z szyfrowaną transmisją WSS załatałoby lukę bezpieczeństwa. Ostatnią luką jest błędna konfiguracja, pozwalająca na dostęp do urządzenia z poziomu innej domeny internetowej.



Rys. 10. Panel sterowania dla ESP8266

Fig. 10. Control panel for ESP8266



Rys. 11. Kod źródłowy WLED

Fig. 11. WLED source code

Na rysunku 11. widoczny jest fragment kodu źródłowego WLED (wled_server.cpp). Na czerwono zaznaczono fragmenty odpowiedzialne za ustawianie nagłówków AccessControl-Allow-Origin: *, które są przyczyną luki. Podczas modyfikacji firmware'u w celu implementacji WSS (WebSocket Secure), można przy okazji usunąć te linie kodu. W ten sposób zostaje wyeliminowany kolejny wektor

ataku. Ważne żeby pamiętać, że podobne nagłówki są w innych plikach oprogramowania.

Skuteczną metodą zabezpieczenia jest segmentacja sieci. Można wydzielić specjalną sieć VLAN przeznaczoną wyłącznie dla urządzeń IoT. Za pomocą takiego VLANu można logicznie odizolować urządzenie od reszty hostów w sieci. Daje to możliwość skonfigurowania reguł zapory sieciowej (firewall), dopuszczając ruch tylko z zaufanych adresów IP.

7. Wnioski

Przeprowadzone testy pozwoliły na osiągnięcie celu pracy. Przeanalizowano bezpieczeństwo urządzenia IoT w sieci domowej i oceniono zabezpieczenia ESP. Badania przeprowadzone na mikrokontrolerze ESP8266, pokazały istotne luki bezpieczeństwa, potwierdziło to postawioną tezę, że urządzenia IoT nie mają odpowiednich zabezpieczeń. Całkowity brak mechanizmów uwierzytelniania dostępu do panelów sterowania jest krytyczną luką bezpieczeństwa. Jedynym zabezpieczeniem była nieznajomość adresu IP urządzenia przez potencjalnego atakującego. Ponadto, urządzenie realizowało komunikację w sposób niezaszyfrowany. W przypadku ESP8266 sterowanie odbywało się za pomocą protokołu WebSocket, przesyłając polecenia w formie plaintext. Umożliwiło to przechwycenie i analizę ruchu sieciowego. W ataku Man-in-the-Middle na ESP8266, udało się zidentyfikować adres IP i strukturę poleceń w formacie JSON. Zebrane informacje pozwoliły na replikację ataków za pomocą programu Postman, wysyłając odpowiednio spreparowane ładunki JSON do ESP. Nmap ujawniło nowe wektory ataku. Wykryto niebezpieczną konfigurację CORS, otwierając możliwość ataku z poziomu złośliwej strony internetowej. Odnalezione luki są możliwe do wyeliminowania Wymaga to zmiany w kodzie, a nie jedynie edycji ustawień w panelu, dodanie hasła dostępu lub kodu PIN nie jest skutecznym rozwiązaniem. Dzieje się tak, ponieważ pakiety te są przesyłane w formie jawnego tekstu przez nieszyfrowane protokoły. Pozwala to na ich łatwe przechwycenie, tak jak przechwycono dane sterujące. Dla ESP8266 należy zastąpić protokół WebSocket na szyfrowaną wersję WSS. Do tego, trzeba też usunąć nagłówki Access-ControlAllow-Origin: *. Dopiero po wdrożeniu szyfrowania i ustawieniu hasła dostępu, można powiedzieć o skutecznym zabezpieczeniu. Innym rozwiązaniem jest wydzielenie osobnej podsieci VLAN wyłącznie dla urządzeń IoT. Za pomocą reguł zapory sieciowej (firewall), można logicznie odizolować urządzenie, zezwalając na komunikację z nim tylko z jednego, zaufanego adresu IP. Podsumowując, badanie wykazało, że nawet w zaufanej sieci domowej, urządzenia IoT stanowią realne zagrożenie. Brak szyfrowania i uwierzytelniania sprawia, że dowolna osoba posiadająca dostęp do tej samej sieci lokalnej może przejąć kontrolę nad urządzeniami. Poprzez luki bezpieczeństwa może też pozyskać krytyczne dane, takie jak poświadczenia do systemów MQTT. Konieczne jest więc poprawne zabezpieczenie sprzętu przez atakami, ponieważ urządzenia IoT mogą być wektorem ataku.

Dalsze prace nad tym zagadnieniem mogą obejmować rozszerzenie analizy o inne popularne urządzenia oraz protokoły komunikacyjne IoT, takie jak Zigbee, Z-Wave czy Bluetooth Low Energy. Wartościowym kierunkiem byłoby praktyczne wdrożenie i przetestowanie zaproponowanych zabezpieczeń, w tym zmodyfikowanego firmware'u z obsługą protokołu WSS oraz segmentacji sieci VLAN. Uzupełnieniem mogłoby być również opracowanie metod automatycznego wykrywania opisanych podatności w sieciach domowych.

Literatura

[1] Abdullah A., Hamad R., Abdulrahman M., Moala H., Elkhediri S.: Cybersecurity: a review of internet of things (iot) security issues, chal-lenges and techniques. In 2019 2nd International Conference on Computer Applica-tions & Information Security (ICCAIS). IEEE, 2019.

[2] Dąbrowski R.: Przeciwdziałanie zagrożeniom bezpieczeństwa iot i ot związa-nym z cyfrową transformacją za pomocą technik „deception"/oszukiwania. Nowa Energia, 2020.

[3] Gao J., Xia K.: Arp spoofing detection algorithm using icmp protocol. In 2013 international conference on computer communication and informatics. IEEE, 2013.

[5] https://en.wikipedia.org/wiki/ARP_spoofing (dostęp 18.12.2025).

[3] https://www.startertutorials.com/blog/introduction-to-iot.html (dostęp 18.12.2025).

[4] Lee I.: Internet of things (iot) cybersecurity: Literature review and iot cyber risk management. Future internet, 2020.

[5] Ramachandran V., Nandi S.: Detecting arp spoofing: An active tech-nique. In International conference on information systems security. Springer, 2005.

[6] Rot A., Blaicke B.: Bezpieczeństwo internetu rzeczy. Wybrane zagro-żenia i sposoby zabezpieczeń na przykładzie systemów produkcyjnych. Zeszyty Naukowe Politechniki Częstochowskiej. Zarządzanie, 2017.

[7] Rot A., Sobińska M.: Skalowalność, bezpieczeństwo i interopera-cyjność jako kluczowe wyzwania dla projektowania systemów internetu rzeczy. Zeszyty Naukowe Politechniki Częstochowskiej. Zarządzanie, 2018.

[8] Trabelsi Z., El-Hajj W.: On investigating arp spoofing security solu-tions. International Journal of Internet Protocol Technology, 2010.

IOT SECURITY IN CRITICAL INFRASTRUCTURE SYSTEMS, LOCAL NETWORK

Summary

The aim of this article is a practical security analysis of domestic Internet of Things (IoT) devices and an assessment of their communication within a local network regarding vulnerability to unauthorized access. The object of the study is the popular ESP8266 microcontroller with open-source WLED software, used for controlling LED lighting. The research methodology is based on conducting active reconnaissance using Nmap and performing an ARP Spoofing (Man-in-the-Middle) attack. For this purpose, Ettercap and Wireshark tools were utilized to capture and analyze network traffic. Throughout the research, a number of critical

security vulnerabilities were identified, including a complete lack of authentication in the control panel and the transmission of control data and credentials (logins and passwords for the MQTT broker) in plain text. It was demonstrated that communication based on the unsecured WebSocket protocol allows for easy reading of the JSON command structure. The experimental results confirmed the possibility of full, unauthorized takeover of device control through attack replication in Postman, while the Nmap scan revealed an additional vulnerability related to the misconfiguration of CORS headers. The main conclusion of the work is the statement that IoT devices without implemented WSS/HTTPS encryption pose a real threat to network security. As effective threat mitigation methods, modifying the firmware source code to implement encrypted protocols, removing dangerous headers, and applying logical device isolation through VLAN segmentation and firewall rule configuration were proposed.

Keywords: Internet of Things, ESP8266, cybersecurity, ARP Spoofing, WebSocket.

Eryk Kopciuch
Studenckie Koło Naukowe Informatyków „Kod”

PROCEDURALNA GENERACJA TERENU Z WYKORZYSTANIEM WIELOWARSTWOWEGO SZUMU PERLINA W SILNIKU UNITY

Streszczenie

Poniższy tekst opisuje skrypt generujący siatkę terenu w silniku Unity. Skrypt wykorzystuje algorytm szumu Perlina, który pozwala na uzyskanie wartości skorelowanych przestrzennie, co umożliwia tworzenie naturalnie wyglądających form terenu, takich jak doliny i wzniesienia. Parametryzacja skryptu pozwala na modyfikowanie właściwości generowanej siatki w zależności od potrzeb użytkownika. Zastosowany został system biomów definiowanych na podstawie wysokości, pozwalający na wizualizację ukształtowania powierzchni i zwiększający realizm generowanego terenu. Rozdział opisuje korzyści wynikające z nakładania na siebie wielu warstw szumu gradientowego oraz wpływ poszczególnych parametrów na wydajność. Opisany został sposób określania sumarycznej wysokości punktu. Autor proponuje możliwe ścieżki rozwoju generatora i przedstawia wpływ parametrów na proces generacji. Wykorzystanie silnika Unity pozwala na modyfikację parametrów z poziomu wbudowanego inspektora, a możliwość określenia ziarna (seed) pozwala na replikację błędów i zapewnia powtarzalność wyników.

Słowa kluczowe: generowanie proceduralne, Unity, gamedev, Szum Perlina, szum fraktalny, grafika komputerowa.

1. Wstęp

1.1. Rola generacji proceduralnej we współczesnych symulacjach i grach wideo

Proceduralna generacja treści jest kluczowa dla zwiększania skali gier komputerowych, poprzez zwiększanie regrywalności (replayabilty) oraz efektywności produkcji. Pozwala na zmniejszenie kosztów i oszczędność pamięci dyskowej, jest też sposobem na osiągnięcie ogromnych, unikalnych wirtualnych światów, których stworzenie byłoby praktycznie niemożliwe w przypadku “ręcznego” projektowania otoczenia. Tereny generowane w ten sposób mogą też służyć do różnego rodzaju symulacji, przykładowo w pracy “Emergence of Locomotion Behaviours in Rich Environments” [1], autorzy pracy wykorzystali proceduralnie generowany

teren w uczeniu ze wzmocnieniem, poprzez wykonywanie symulacji na zmiennym terenie, wymusiło na agentach wykształcenie ogólnych zasad lokomocji, zamiast dopasowania do statycznej trasy.

1.2. Zalety zastosowania ziarna (seed).

Zastosowanie ziarna pozwala na otrzymanie tego samego wyniku generacji, mimo generacji losowej w celu uzyskania tego samego środowiska, odpowiadającego oczekiwaniom użytkownika. Inną zaletą tego rozwiązania jest możliwość replikacji błędów zgłaszanych przez użytkowników na podstawie ziarna, które zostało użyte przez nich podczas generacji.

2. Matematyczne podstawy generowania terenu

2.1. Szum Perlina

Opracowany przez Kena Perlina w 1985 r. [2] algorytm tworzący szum gradientowy powstał jako narzędzie zwiększające realizm w grafice komputerowej. W przeciwieństwie do szumu białego, który opiera się na wartościach całkowicie losowych i nieskorelowanych, szum Perlina bazuje na wektorach nachylenia (gradientach) wyznaczanych w punktach kratowych. Dzięki interpolacji tych wektorów algorytm generuje płynne, skorelowane w przestrzeni przejścia. To właśnie ta metoda sprawia, że teren nie jest „poszarpany", lecz tworzy naturalnie wyglądające „pagórki" i „doliny" (rys. 2.).

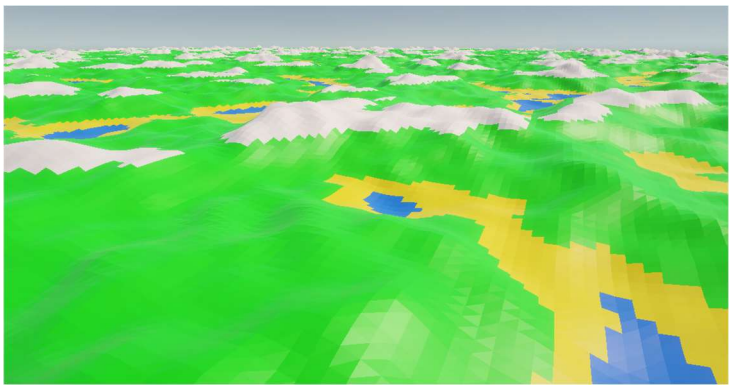
2.2. Zjawisko fraktalności i wykorzystanie oktaw

W celu poprawienia wyniku generacji i uzyskania bardziej realistycznego wyniku, konieczne jest nałożenie na siebie kilku warstw szumu o różnych częstotliwościach i amplitudach nazywanych oktawami. Wzór na wynikową wysokość punktu (x, y) można wyrazić wzorem:

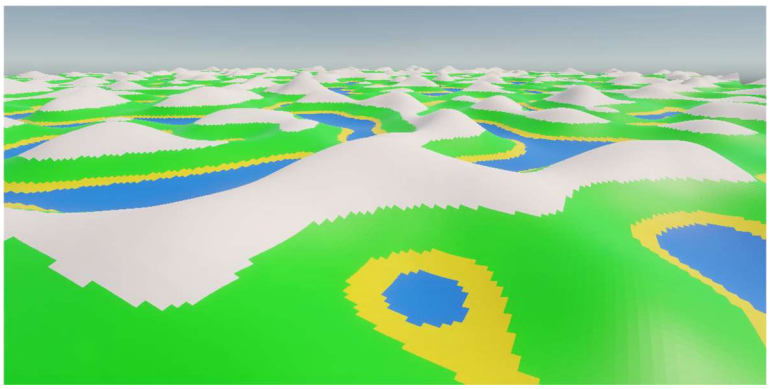
$$H(x,y) = \sum \left(A_i \cdot P(f_i \cdot x,\ f_i \cdot z)\right) \tag{1}$$

gdzie: H – sumaryczna wysokość,
n – liczba oktaw,
A_i – amplituda i-tej warstwy,
f_i – częstotliwość i-tej warstwy,
P – funkcja Perlina

Skrypt korzysta ze zmiennych persistence i lacunarity do modyfikowania wpływu kolejnych oktaw na generowaną siatkę. Persistence (trwałość) charakteryzuje wpływ każdej kolejnej oktawy na końcowy kształt. Lacunarity (lakunarność) określa tempo wzrostu szczegółowości (częstotliwości) w kolejnych warstwach. Modyfikacja tych dwóch parametrów ma znaczący wpływ na wynik końcowy.



Rys. 1. Przykładowa generacja terenu (4 oktawy)
Fig. 1. Example of generated terrain (4 octaves)



Rys. 2. Przykładowa generacja terenu (1 oktawa)
Fig. 2. Example of generated terrain (1 octave)

2.3. Normalizacja i redystrybucja wartości

W związku z tym, że wynik sumowania oktaw może wykroczyć poza standardowy zakres, skrypt (Listing 1) dzieli sumy przez maksymalną wartość co pozwala na utrzymanie wartości w zakresie [0,1]. Następnie dokonuje redystrybucji poprzez podniesienie wartości do kwadratu. Dzięki temu uzyskujemy teren bardziej realistyczny, którego zmiana jest nieliniowa.

```
Vector3 GetHeightWithOctaves(int x, int z)
{
  float noiseHeight = 0;
  float amplitude = 1;
  float frequency = 1;
  float maxValue = 0; // Potrzebne do normalizacji
```

```
    for (int i = 0; i < octaves; i++)
    {
        float xCoord = (float)(x + offsetX) / scale * frequency;
        float zCoord = (float)(z + offsetZ) / scale * frequency;

        float perlinValue = Mathf.PerlinNoise(xCoord, zCoord);
        noiseHeight += perlinValue * amplitude;

        maxValue += amplitude;
        amplitude *= persistence;
        frequency *= lacunarity;
    }

    // Normalizacja
    float normalizedHeight = noiseHeight / maxValue;

    // Redystrybucja
    float finalHeight = Mathf.Pow(normalizedHeight, 2.0f);

    return new Vector3(x, finalHeight * meshHeightMultiplier, z);
}
```

Listing 1. Funkcja zwracająca sumaryczną wysokość punktu

Listing 1. Function returning summarized point height.

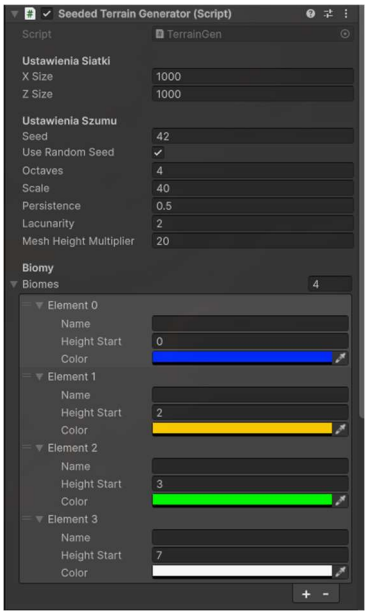
3. Architektura algorytmu w środowisku Unity

3.1. Parametryzacja generatora

W celu umożliwienia użytkownikowi dopasowania generatora do jego wymagań kluczowe zmienne wpływające na wynik są modyfikowalne z poziomu Inspektora Unity (rys. 3.). xSize, zSize to zmienne określające rozdzielczość siatki w płaszczyźnie poziomej. Seed to zmienna określająca ziarno dla generatora liczb losowych.

3.2. Budowa siatki

Proces tworzenia siatki (Mesh) [3], charakteryzuje się tworzeniem kafelków złożonych z 2 trójkątów co pozwala na łatwy port do innych frameworków/silników np. OpenGL w zależności od potrzeb. Kolor kafelka jest określany na podstawie wysokości punktów, z których się składa, dzięki czemu możliwe jest uzyskanie "biomów" ułatwiających wizualizację (rys. 1.). W związku z tym, że domyślne indeksowanie 16-bitowe ogranicza siatkę do 65 535 wierzchołków, przy większych terenach należy wymusić indeksowanie typem o większym rozmiarze (IndexFormat.UInt32).



Rys. 3. Widok zmiennych edytowalnych z poziomu inspektora Unity

Fig. 3. Variables editable from Unity inspector

4. Analiza wydajności

4.1. Złożoność obliczeniowa

Głównymi czynnikami wpływającymi na wydajność algorytmu są liczba oktaw i liczba wierzchołków określana przez rozmiar siatki (xSize * ySize). Wyznaczanie wysokości jest operacją o złożoności O (V * n), gdzie V to liczba wierzchołków, a n to liczba oktaw, co przy bardzo gęstych lub rozległych siatkach może powodować problemy wydajnościowe.

5. Podsumowanie

5.1. Wnioski

W wyniku połączenia sumowania oktaw z nieliniową redystrybucją wysokości możliwe jest generowanie zróżnicowanych krajobrazów, zmiennych w zależności od określonych parametrów. Generacja uzależniona od podanych przez użytkownika parametrów pozwala na dostosowanie do własnych potrzeb.

5.2. Kierunki rozwoju

Przyszłe prace powinny być skupione na modyfikacji algorytmu w taki sposób, aby generował on płaty (chunki) terenu wokół gracza/użytkownika. W ten

sposób możliwe będzie osiągnięcie nieskończonej mapy powiększającej się wraz z eksploracją przez użytkownika. Możliwe jest też rozszerzenie mechaniki "biomów", aby stworzyć tereny o zmiennym krajobrazie i specyfice, takie jak pustynie czy wysokie pasma górskie.

Literatura

[1] Heess, Nicolas, et al. "Emergence of locomotion behaviours in rich environments." arXiv preprint arXiv:1707.02286 (2017).

[2] Perlin K.: An Image Synthesizer, SIGGRAPH Computer Graphics, 1985.

[3] Dokumentacja klasy Mesh, dostępnej w silniku Unity https://docs.unity3d.com/6000.3/Documentation/ScriptReference/Mesh.html

PROCEDURAL TERRAIN GENERATION USING MULTI-LAYERED PERLIN NOISE IN THE UNITY ENGINE

Summary

The following text describes a script for generating terrain meshes within the Unity engine. The script utilizes the Perlin noise algorithm, which provides spatially correlated values, enabling the creation of natural-looking terrain features such as valleys and ridges. Script parameterization allows for modifying the properties of the generated mesh according to user requirements. A height-based biome system has been implemented, facilitating the visualization of surface topography and enhancing the realism of the generated terrain. The article discusses the benefits of layering multiple gradient noise octaves and the impact of specific parameters on performance. The method for determining the total height of a point is also described. The author proposes potential development paths for the generator and demonstrates the influence of various parameters on the generation process. Utilizing the Unity engine enables parameter modification via the built-in inspector, while the ability to specify a seed allows for bug replication and ensures result reproducibility.

Keywords: procedural generation, Unity, gamedev, Perlin noise, fractal noise, computer graphics.

Maciej Karczmarz
Studenckie Koło Naukowe Informatyków „Kod”

ZINTEGROWANY POTOK PRODUKCYJNY PROCEDURALNEGO GENEROWANIA ŚRODOWISK TRÓJWYMIAROWYCH Z WYKORZYSTANIEM SILNIKA UNREAL ENGINE 5 ORAZ NARZĘDZI GAEA I HOUDINI

Streszczenie

Głównym celem niniejszych badań jest ewaluacja zintegrowanego potoku proceduralnego służącego do budowy wysoce realistycznych środowisk trójwymiarowych czasu rzeczywistego. Badanie koncentruje się na synergii technicznej między silnikiem Unreal Engine 5 – wykorzystującym wirtualizowaną geometrię Nanite oraz globalne oświetlenie Lumen – a wyspecjalizowanymi narzędziami proceduralnymi: Gaea do zaawansowanych symulacji geologicznych terenu oraz Houdini do parametrycznego generowania roślinności, enkapsulowanego w formie Houdini Digital Assets (HDA). Metodyka opierała się na czteroetapowym przepływie pracy: generowaniu bazowej mapy wysokości, proceduralnym tworzeniu zasobów za pośrednictwem Houdini Engine oraz finalnej kompozycji sceny z wykorzystaniem fotogrametrii Quixel Megascans i środowiska Procedural Content Generation (PCG) Framework. Scena demonstracyjna oferuje trzy dynamicznie wymienne warianty biomów (leśny, pustynny, zimowy), sterowane w sposób niedestrukcyjny poprzez Master Material Instances. Telemetria wydajnościowa wykazała, że zastosowanie wektorowego proceduralnego maskowania śniegu oraz technologii Runtime Virtual Texture (RVT) do płynnego mieszania siatek pozwoliło utrzymać 30–36 FPS w rozdzielczości 1440p. Wynik ten osiągnięto przy użyciu technologii NVIDIA DLSS (Deep Learning Super Sampling) na karcie graficznej RTX 4060 Ti. Mimo imponującej skalowalności systemu i skrócenia czasu iteracji dla artystów technicznych, rezultaty ujawniają krytyczne ograniczenie przepustowości pamięci VRAM. Telemetria zarejestrowała zużycie na poziomie około 7,5 GB, wyczerpując 94% dostępnego limitu 8 GB. Tak znaczne zapotrzebowanie na VRAM stanowi spore wyzwanie dla wielu maszyn obecnych w naszym posiadaniu.

Słowa kluczowe: generowanie proceduralne, fotogrametria, Unreal Engine 5, projektowanie poziomów, optymalizacja renderowania.

1. Wprowadzenie

Dynamiczny rozwój grafiki czasu rzeczywistego stawia przed twórcami gier i wizualizacji coraz wyższe wymagania w zakresie realizmu oraz skali tworzonych światów. Odbiorcy oczekują dziś światów nie tylko uderzająco realistycznych, ale także rozległych i wypełnionych detalami. Próba sprostania tym wymaganiom przy użyciu wyłącznie tradycyjnych metod, opartych na ręcznym modelowaniu i teksturowaniu każdego zasobu (ang. asset), staje się poważnym ograniczeniem produkcyjnym ze względu na wysokie koszty i znaczące wydłużenie czasu produkcji.

W tradycyjnym podejściu proces produkcji opierał się na rygorystycznych procesach optymalizacyjnych, mających na celu zrównoważenie jakości wizualnej z ograniczoną wydajnością sprzętową. Kluczowym elementem tego potoku produkcyjnego było tzw. wypalanie detali (ang. baking). Proces ten polegał na przenoszeniu skomplikowanych detali geometrycznych z modelu o wysokiej liczbie poligonów (high-poly) na model uproszczony (low-poly) przy wykorzystaniu specjalnych tekstur, zwanych mapami normalnymi (ang. normal maps). Dodatkowo standardem było ręczne tworzenie poziomów szczegółowości (ang. Levels of Detail, LODs), co polegało na przygotowaniu kilku wersji tego samego modelu o stopniowo redukowanej liczbie trójkątów, podmienianych dynamicznie przez silnik w zależności od odległości od kamery. Wymagało to od grafików ogromnego nakładu pracy oraz generowało dodatkowe koszty pamięciowe.

Odpowiedzią na te wyzwania są nowoczesne silniki gier, takie jak Unreal Engine 5, który zrewolucjonizował zarządzanie geometrią dzięki technologii Nanite oraz dynamicznemu oświetleniu Lumen. Technologie te całkowicie odwróciły tradycyjny proces produkcji zasobów cyfrowych. Samo posiadanie dostępu do potężnych narzędzi nie rozwiązuje jednak problemu skalowania produkcji. Prawdziwym wyzwaniem inżynierskim staje się ich efektywna integracja i opracowanie metod precyzyjnego sterowania algorytmami proceduralnymi. Celem niniejszego rozdziału jest dogłębna prezentacja i ewaluacja autorskiego potoku produkcyjnego (ang. pipeline), integrującego silnik Unreal Engine 5 z oprogramowaniem Gaea, Houdini oraz bibliotekami fotogrametrycznymi Quixel Megascans w celu stworzenia realistycznego środowiska naturalnego.

2. Przegląd technologii i metodyka badawcza

Współczesne techniki renderowania w czasie rzeczywistym dążą do fizycznie poprawnego odwzorowania światła (PBR). Nowoczesny proces produkcyjny odwraca dawny paradygmat manualnego rzeźbienia przestrzeni.

2.1. Generowanie proceduralne i fotogrametria

Generowanie proceduralne to metoda tworzenia treści za pomocą algorytmów i zestawów zdefiniowanych reguł, zamiast ręcznego ich projektowania. Zamiast

manualnie modelować każdą skałę lub rzeźbić każdy fragment terenu, twórca definiuje procedurę, która jest w stanie wygenerować pożądany zasób. Fundamentem tej techniki jest kontrolowana losowość, często oparta na wartości początkowej zwanej ziarnem (ang. seed) [4]. Pozwala to na nieliniowość i iteracyjność – zmiana jednego z początkowych parametrów automatycznie modyfikuje wszystkie kolejne etapy. Główną zaletą tej techniki jest niezwykła skalowalność i efektywność pamięciowa, jednak wiąże się ona z problemem trudności w utrzymaniu ścisłej kontroli artystycznej nad ostatecznym wynikiem oraz stromą krzywą uczenia się.

Z kolei fotogrametria to technika polegająca na odtwarzaniu kształtów, rozmiarów i położenia obiektów w trójwymiarze na podstawie analizy serii obrazów. Efektem jest wygenerowanie siatki 3D o wysokiej gęstości oraz zestawu tekstur idealnie odwzorowujących rzeczywisty materiał. Biblioteka Quixel Megascans, wykorzystywana w niniejszej pracy, jest w całości oparta na tej technologii. Pozwala to na przeniesienie detali ze świata rzeczywistego (np. mikroporów w skale) do wirtualnego, co eliminuje potrzebę artystycznej interpretacji i zgadywania, jak dany materiał reaguje na światło.

2.2. Nowoczesne techniki renderowania

Wprowadzenie silnika Unreal Engine 5 zrewolucjonizowało podejście do renderowania za sprawą dwóch flagowych systemów [1]:

- Nanite: Jest to system wirtualizowanej geometrii mikropoligonów, który eliminuje konieczność ręcznego upraszczania modeli do wersji low-poly. System ten poprzez analizę siatki 3D inteligentnie generuje i strumieniuje tylko te mikropoligony, które są aktualnie potrzebne do pokrycia pikseli na ekranie, pozwalając na bezpośrednie wykorzystanie skomplikowanych siatek o jakości filmowej bez utraty wydajności i bez potrzeby tworzenia LOD-ów [2].
- Lumen: System w pełni dynamicznego globalnego oświetlenia (ang. Global Illumination, GI) oraz odbić. Pozwala on na natychmiastową aktualizację oświetlenia przy dowolnej zmianie geometrii lub źródła światła, w odróżnieniu od starszych metod wymagających wielogodzinnego wypalania statycznych map świetlnych (lightmaps).

3. Charakterystyka wykorzystanych narzędzi

Realizacja projektu opierała się na zintegrowanym potoku produkcyjnym, wymagającym synchronizacji kilku specjalistycznych aplikacji.

- Unreal Engine 5: Zaawansowany silnik do tworzenia gier (wersja 5.6.1), łączący wysoką jakość oprawy wizualnej z bogatym zestawem narzędzi deweloperskich (np. system Blueprints).
- Houdini: Profesjonalne środowisko do proceduralnego generowania grafiki 3D (wersja 21.0.512) oparte na węzłach. Integracja z Unreal Engine za

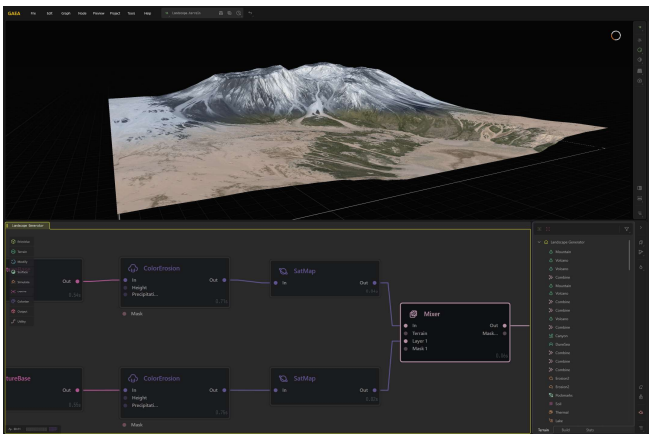
pomocą wtyczki Houdini Engine pozwala na dynamiczny import i modyfikację proceduralnych zasobów w czasie rzeczywistym.

- Gaea: Specjalistyczne oprogramowanie (wersja 2.0) dedykowane tworzeniu realistycznych terenów 3D poprzez zaawansowaną symulację zjawisk geologicznych (erozja, sedymentacja).
- Quixel Megascans i Bridge: Narzędzie integrujące potężną bibliotekę fotogrametryczną z silnikiem Unreal Engine, udostępniane bezpłatnie dla użytkowników tego silnika.
- GitLab: Platforma wykorzystana jako centralne repozytorium (system kontroli wersji Git) zapewniająca bezpieczeństwo i synchronizację danych.

4. Metodyka badawcza i implementacja potoku

Proces budowania środowiska wirtualnego został ustrukturyzowany w cztery logicznie następujące po sobie etapy.

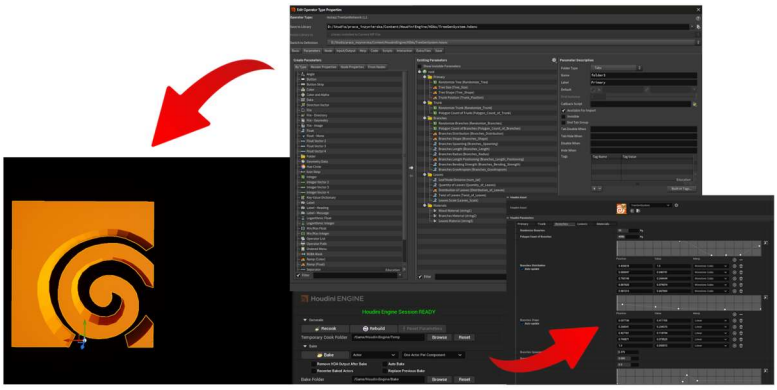
Etap 1: Generowanie fundamentu terenu. Podstawą środowiska jest mapa wysokości wygenerowana w aplikacji Gaea [3]. Oprogramowanie to, dzięki interfejsowi opartemu na węzłach (node-based), specjalizuje się w realistycznej symulacji zjawisk geologicznych. Proces rozpoczął się od zdefiniowania podstawowych kształtów za pomocą węzłów Mountain lub Ridge [rys. 1.]. Następnie pierwotna forma została poddana serii modyfikacji urealniających. Kluczową rolę odegrał węzeł Erosion, symulujący procesy erozji hydraulicznej (tworzenie żłobień i osadów), węzeł Thermal odpowiadający za erozję termiczną (osuwanie materiału skalnego) oraz węzeł Slope do selektywnego maskowania nachylenia. Finalnym krokiem był eksport z głównego łańcucha węzłów map wysokości w formacie 8-bitowego pliku PNG w rozdzielczości 1009x1009 pikseli i ich import poprzez narzędzie Landscape w silniku Unreal Engine.



Rys. 1. Proceduralne generowanie mapy wysokości w oprogramowaniu opartym na węzłach

Fig. 1. Procedural heightmap generation in node-based software

Etap 2: Tworzenie zasobów proceduralnych. Do generowania bardziej złożonych zasobów, w szczególności roślinności (drzew), wybrano środowisko Houdini [5]. Zamiast modelować pojedyncze drzewo, zbudowano proceduralny system zdolny do wygenerowania całej populacji unikalnych, spójnych stylistycznie drzew. Centralną koncepcją tego procesu jest stworzenie zasobu Houdini Digital Asset (HDA). Jest to skapsułkowana sieć węzłów posiadająca własny interfejs parametryczny [rys. 2.].



Rys. 2. Interfejs parametryczny cyfrowego zasobu (HDA) osadzony bezpośrednio w silniku

Fig. 2. Parametric interface of a Digital Asset (HDA) embedded directly in the engine

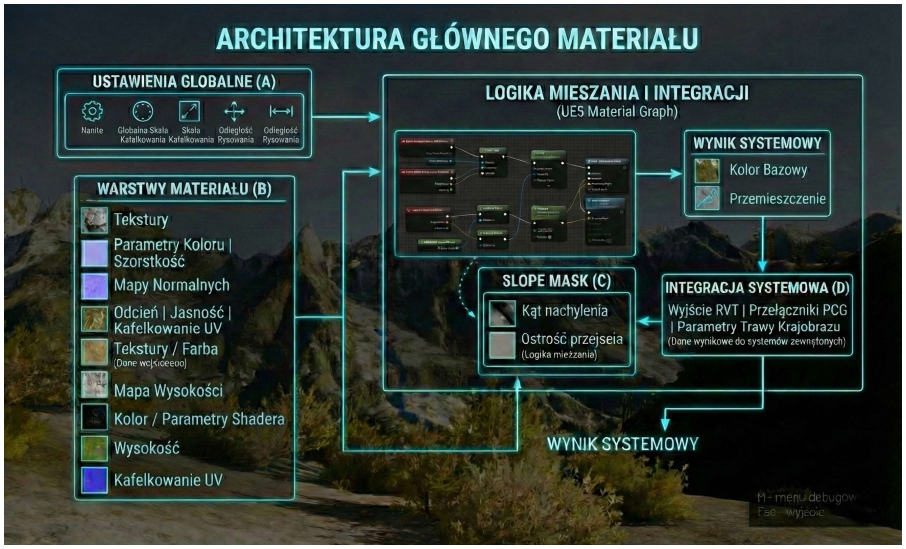
Integracja z Unreal Engine odbywała się za pośrednictwem wtyczki Houdini Engine. Zamiast eksportować model statyczny (np. FBX), plik HDA był importowany bezpośrednio do edytora silnika, gdzie wtyczka przetwarzała go w czasie rzeczywistym. Dało to autorowi bezpośredni dostęp do zdefiniowanych parametrów (suwaków odpowiadających np. za wysokość drzewa, krzywiznę pnia) w panelu szczegółów Unreal Engine, drastycznie przyspieszając iterację poziomów bez powrotu do aplikacji zewnętrznej.

Etap 3 i 4: Integracja fotogrametrii i kompozycja. Równolegle z generowaniem terenu i roślinności, zgromadzono bibliotekę wysokiej jakości zasobów fotogrametrycznych. Do importu użyto narzędzia Quixel Bridge, służącego jako pomost do biblioteki Megascans. Pobrane stamtąd setki wysokiej jakości modeli (Static Meshes) – takich jak formacje skalne czy poszycie leśne – wykorzystano do wypełnienia sceny detalami. Warto odnotować trwającą integrację tych systemów w platformę Fab przez firmę Epic Games, niemniej w omawianym projekcie bazowano na dojrzałym środowisku Quixel. Elementy te zostały finalnie rozmieszczone na terenie przy pomocy systemów generowania proceduralnego wewnątrz silnika (m.in. PCG Framework).

5. Architektura proceduralnego materiału terenu (Studium Przypadku)

Aby zachować wysoką wydajność, elastyczność oraz spójność wizualną, zaimplementowano zaawansowany system materiału terenu (Auto Blend Material). Został on oparty na rygorystycznej trójpoziomowej architekturze, zgodnej z dobrymi praktykami inżynierskimi:

1. Funkcje Materiałowe (Material Functions): Hermetyczne bloki logiki realizujące konkretne zadania matematyczne.
2. Materiał Główny (Master Material): Główny węzeł logiczny organizujący wywoływanie i łączenie funkcji.
3. Instancje Materiałowe (Material Instances): Parametryczny panel sterowania pozwalający na modyfikacje bez rekompilacji shadera.



Rys. 3. Logika węzłów realizująca próbkowanie i mieszanie wirtualnej tekstury

Fig. 3. Node logic executing sampling and blending of a virtual texture

5.1. Wtapianie obiektów w teren (MF_RVT_Blend)

Kluczowym problemem inżynierskim w projektowaniu środowisk jest zniwelowanie nienaturalnych, ostrych krawędzi na styku obiektów statycznych z gruntem. Wykorzystano do tego technikę Runtime Virtual Texture (RVT) poprzez autorską funkcję MF_RVT_Blend. Logikę podzielono na trzy etapy [rys. 3.]:

- Próbkowanie danych: Pobranie informacji o wektorze normalnym, chropowatości i kolorze terenu z wirtualnej tekstury (węzeł RVT Sample).
- Generowanie maski przenikania: System porównuje bezwzględną wysokość (Z) piksela obiektu z wysokością terenu zapisaną w RVT. Różnica ta

służy do wygenerowania gradientu decydującego o tym, w którym miejscu skała przejmuje właściwości mchu czy piasku.
- Finalna interpolacja: Matematyczne wymieszanie atrybutów materiałów za pomocą węzłów interpolacji liniowej (Lerp).

5.2. Matematyka proceduralna maski śnieżnej (MF_Snow_Mask)

Kolejnym innowacyjnym elementem, kluczowym dla zimowego wariantu biomu, była funkcja MF_Snow_Mask, odpowiadająca za proceduralne generowanie pokrywy śnieżnej. Algorytm nie opierał się na statycznych maskach bitmapowych, lecz na dynamicznych obliczeniach wektorowych. Wykorzystano operację iloczynu skalarnego pomiędzy wektorem normalnym powierzchni w danym pikselu (N) a wektorem idealnego pionu reprezentującym grawitację (Up = [0, 0, 1]).

Zależność tę opisuje równanie:

$$N \cdot Up = \cos\theta$$

co dla (U_P=[0,0,1]) jest równoważne zależności:

$$N_z = \cos\theta$$

Gdzie θ jest kątem pomiędzy wektorem normalnym terenu a wektorem pionu. Wynik zbliżony do 1.0 oznacza powierzchnię płaską (kąt bliski 0 stopni), sprzyjającą naturalnej akumulacji śniegu. Wynik zbliżony do 0.0 oznacza powierzchnię pionową (kąt bliski 90 stopni), z której śnieg uległby zsunięciu. Otrzymany gradient jest następnie poddawany operacji potęgowania, co pozwala na precyzyjną manipulację kontrastem maski i ostrością krawędzi śniegu.

5.3. Agregacja i struktura kaskadowa

Aby wyeliminować powtarzalność tekstur na rozległych terenach (ang. tiling), opracowano funkcję MF_Tiling_Variation. Z kolei funkcja MF_Create_Layer_Landscape posłużyła jako konstruktor danych, pakujący wszystkie sygnały bazowe (Albedo, Normal, Height) w ustandaryzowaną strukturę atrybutów, znacząco poprawiając czytelność kodu. Dodatkowo wykorzystano mniejsze bloki, takie jak MF_Depth_Fade do miękkiego wygaszania materiałów przy kolizjach obiektów oraz MF_Color_Variation do wprowadzania losowości barw. W Materiale Głównym ustrukturyzowano to w architekturę stosu warstw (kaskadowe mieszanie od dołu do góry), co pozwoliło na zachowanie logiki przy ogromnej złożoności całego systemu.

6. Wyniki i dyskusja wydajnościowa

Efektem wdrożenia omówionego, zintegrowanego potoku proceduralnego jest w pełni funkcjonujące środowisko demonstracyjne. Zastosowanie architektury

opartej na Instancjach Materiałowych pozwoliło na wykreowanie trzech radykalnie odmiennych wariantów środowiska na tej samej siatce terenu, bez konieczności kosztownej rekompilacji shaderów. Stworzono biom leśny (scena główna), biom pustynny oraz biom zimowy [rys. 4.].



Rys. 4. Fragment finalnej sceny demonstracyjnej w wariancie biomu leśnego
Fig. 4. Fragment of the final demonstration scene in the forest biome variant

Obiektywna analiza wydajności sceny stanowi kluczowy element badawczy. Pomiary zostały przeprowadzone na platformie testowej wyposażonej w procesor AMD Ryzen 5 5600X, 32 GB pamięci RAM DDR4 3600 MHz oraz układ graficzny NVIDIA GeForce RTX 4060 Ti posiadający 8 GB pamięci VRAM.

W trakcie testów dynamicznego biomu leśnego, przy renderowaniu w rozdzielczości 2560x1440 pikseli (1440p) na bardzo wysokich ustawieniach detali graficznych, płynność utrzymywała się na poziomie około 30-36 FPS (klatek na sekundę).

Osiągnięcie tej płynności wymagało kluczowego wsparcia sprzętowego w postaci implementacji technologii NVIDIA DLSS (ang. Deep Learning Super Sampling). Jest to zaawansowana technika rekonstrukcji obrazu oparta na sztucznej inteligencji. Jej mechanizm polega na renderowaniu trójwymiarowej sceny w znacznie niższej, wewnętrznej rozdzielczości (np. 1080p), co drastycznie obniża czas pracy procesora graficznego. Następnie, wyspecjalizowany model sieci neuronowej uruchamiany na rdzeniach Tensor rekonstruuje klatkę do docelowej, wysokiej rozdzielczości z wykorzystaniem wektorów ruchu z poprzednich klatek, zapewniając jakość nierzadko przewyższającą rendering natywny.

Analiza danych telemetrycznych pozwoliła na zidentyfikowanie najpoważniejszego wąskiego gardła (ang. bottleneck) technologicznego, którym okazało się radykalne zużycie pamięci wideo (VRAM). Wykorzystanie dedykowanej pamięci dla renderowanej sceny wynosiło około 7.5 GB z dostępnego budżetu wynoszącego 8 GB. Stanowi to wykorzystanie na poziomie niemal 94% całkowitej pojemności karty. Obserwowano, że wartość 8 GB stanowi brutalną granicę pojemnościową. Przekroczenie tego limitu skutkuje przymusowym transferem danych tekstur do powolniejszej pamięci operacyjnej komputera (system RAM), co wywołuje nieprzewidziane opóźnienia, zacinanie obrazu i degradację wydajności.

Przyczyną tak ekstremalnego obciążenia sprzętu była kombinacja kilku czynników charakterystycznych dla najnowszych silników. Mimo dobrodziejstw wirtualizacji Nanite dla obiektów skalnych, renderowanie gęstej, transparentnej roślinności, która poddaje się deformacjom wiatru, stanowi ogromny koszt obliczeniowy. Kolejnym czynnikiem było agregowanie wielowarstwowych materiałów proceduralnych z teksturami PBR w najwyższej rozdzielczości, dostarczanymi bezpośrednio z systemów fotogrametrii z bibliotek Quixel.

7. Wnioski i kierunki dalszych prac

Przeprowadzone analizy badawcze i eksperymentalne potwierdzają, że integracja nowoczesnych narzędzi proceduralnych (takich jak Gaea do symulacji terenu i Houdini do HDA) z silnikiem Unreal Engine 5 jest kierunkiem optymalnym z punktu widzenia efektywności. Zastosowanie modułowego frameworka opublikowanego jako instancje materiałowe oraz dynamiczne zasoby parametryczne znacząco przyspiesza proces iteracyjny twórców. Pozwala to jednoosobowym deweloperom i małym studiom na generowanie rozbudowanych światów, które w tradycyjnym, manualnym systemie wymagałyby całych zespołów artystycznych.

Z drugiej jednak strony, eksperyment ujawnił surowe wymogi sprzętowe nowego potoku. Rekomenduje się, by środowiskiem roboczym do płynnego budowania proceduralnych poziomów w jakości fotogrametrycznej PBR były stacje robocze wyposażone w karty graficzne z minimum 12 GB, a najlepiej 16 GB dedykowanej pamięci VRAM.

Biorąc pod uwagę zaprezentowane wyniki, proponuje się następujące kierunki rozwoju badawczego opisanego projektu:

1. Optymalizacja roślinności: Implementacja rygorystycznych systemów HLOD (Hierarchical Level of Detail) lub testowanie w pełni zgeometryzowanych liści obsługiwanych natywnie przez Nanite, co zredukowałoby koszt zjawiska tzw. nadmiernego rysowania pikseli (overdraw).
2. Zastosowanie Virtual Texturingu (RVT): Rozszerzenie wdrożonej już techniki przenikania skał na cały krajobraz, co pozwoliłoby na strumieniowanie zespolonych tekstur wprost do pamięci buforowej, znacząco odciążając budżet VRAM.

3. Wdrożenie mechanik rozgrywki: Rozbudowa biernego poziomu wizualnego o logikę interakcji gracza, na przykład autorskie systemy wspinaczkowe bazujące na algorytmicznie analizowanej topografii zboczy wygenerowanej przez oprogramowanie Gaea.

Literatura

[1] Akenine-Möller T., Haines E., Hoffman N.: Real-Time Rendering, Fourth Edition. AK Peters/CRC Press, 2018.

[2] Epic Games: Nanite Virtualized Geometry Overview, Unreal Engine Documentation, https://dev.epicgames.com (dostęp: luty 2026 r.).

[3] Quadspinner: Gaea 2.0 Documentation, https://docs.quadspinner.com/ (dostęp: luty 2026 r.).

[4] Shaker N., Togelius J., Nelson M. J.: Procedural Content Generation in Games, Springer International Publishing, 2016.

[5] SideFX: Houdini 21.0 Documentation, https://www.sidefx.com/docs/houdini/ (dostęp: luty 2026 r.).

INTEGRATED PRODUCTION PIPELINE FOR PROCEDURAL GENERATION OF 3D ENVIRONMENTS USING UNREAL ENGINE 5, GAEA AND HOUDINI

Summary

The primary objective of this research is to evaluate an integrated procedural pipeline for constructing highly realistic, real-time 3D environments. The study focuses on the technical synergy between Unreal Engine 5 – utilizing Nanite virtualized geometry and Lumen global illumination – and specialized procedural tools: Gaea for advanced geological terrain simulations and Houdini for parametric vegetation generation encapsulated as Houdini Digital Assets (HDA). The methodology followed a four-stage workflow: base heightmap generation, procedural asset development via the Houdini Engine, and final scene composition utilizing Quixel Megascans photogrammetry alongside the Procedural Content Generation (PCG) Framework. The demonstration scene features three dynamically swappable biome variants (forest, desert, winter) controlled non-destructively via Master Material Instances. Performance telemetry indicated that applying vector-based procedural snow masking and Runtime Virtual Texture (RVT) technology for seamless mesh blending sustained 30–36 FPS at 1440p resolution. This was achieved using NVIDIA DLSS (Deep Learning Super Sampling) on an RTX 4060 Ti graphics card. While demonstrating immense scalability and optimizing iteration times for technical artists, the results highlight a critical VRAM bottleneck. Telemetry recorded memory usage of approximately 7.5 GB, exhausting 94% of the 8 GB limit. Such substantial VRAM demands pose a significant challenge for a large portion of the available hardware.

Keywords: procedural generation, photogrammetry, Unreal Engine 5, level design, rendering optimization.

Przemysław Lib, Karolina Kielar, Michał Meszko, Kacper Szlachta, Bartosz Dusza, Filip Mordka, Kacper Barański
Koło Naukowe PRz Polycomp

Dr hab. inż. Wiesław FRĄCZ, Prof. PRz
Dr inż. Grzegorz JANOWSKI
Dr inż. Łukasz BĄK
opiekunowie naukowi

WYKORZYSTANIE RECYKLATÓW POLIMEROWYCH W TECHNOLOGII DRUKU 3D – PRZEGLĄD WŁAŚCIWOŚCI I WYZWAŃ TECHNOLOGICZNYCH

Streszczenie

W pracy przeanalizowano rolę technologii przyrostowych w gospodarce o obiegu zamkniętym, wskazując na potencjał zagospodarowania odpadów polimerowych jako źródła surowca do produkcji filamentów. Omówiono przydatność popularnych tworzyw termoplastycznych do ponownego przetworzenia, zwracając uwagę na istotne różnice między zanieczyszczonymi odpadami poużytkowymi a czystszymi odpadami poprodukcyjnymi. Przedstawiono technologie wytwarzania filamentu, podkreślając konieczność odpowiedniego przygotowania materiału, w szczególności jego oczyszczenia i suszenia w celu ograniczenia wad strukturalnych. Wskazano, że wielokrotne przetwarzanie polimerów prowadzi do degradacji termiczno-oksydacyjnej oraz skracania łańcuchów makrocząsteczkowych, co skutkuje pogorszeniem właściwości mechanicznych i obniżeniem adhezji międzywarstwowej wydruków. Jednocześnie omówiono metody ograniczania tych efektów, takie jak stosowanie stabilizatorów, mieszanie recyklatu z surowcem pierwotnym oraz optymalizacja parametrów druku. Analiza literatury wskazuje, że przy zachowaniu odpowiedniego reżimu technologicznego materiały z odzysku mogą stanowić wartościową alternatywę dla tworzyw pierwotnych.

Słowa kluczowe: recykling, druk 3D, filament, polimery, gospodarka o obiegu zamkniętym.

1. Wprowadzenie

Druk 3D stał się istotnym procesem wytwarzania wyrobów polimerowych na świecie. Technologia ta znajduje zastosowanie w szybkim prototypowaniu,

produkcji wyrobów spersonalizowanych oraz elementów o złożonej geometrii, trudnej do uzyskania metodami konwencjonalnymi [1].

Gospodarka o obiegu zamkniętym (GOZ) to zrównoważony model gospodarczy zakładający minimalizację zużycia surowców i ilości generowanych odpadów poprzez praktyki takie jak: recykling, ponowne wykorzystywanie, naprawa uszkodzonych produktów i efektywne wykorzystywanie zasobów. Druk 3D może generować duże ilości odpadów ze względu na czynniki związane z koniecznością stabilizacji procesu takie jak: trudności w kalibracji drukarki 3D i ustawieniu parametrów wydruku, nieudane wydruki, konieczność stosowania podpór [2]. Ponadto proces ten skłania inżynierów do tworzenia dużej liczby prototypów, które po sprawdzeniu funkcjonalności danego wyrobu poddawane są utylizacji. Mimo rosnącego zainteresowania recyklingiem w druku 3D, brakuje podsumowania aktualnych osiągnięć i wyzwań w tym obszarze. Niniejszy przegląd stanowi pewne uzupełnienie i zebranie informacji w temacie możliwości wykorzystania recyklatów w druku 3D.

W 2022 roku każdy mieszkaniec Unii Europejskiej wytworzył średnio 36,1 kg odpadów opakowaniowych z tworzyw sztucznych. Całkowita masa odpadów z tworzyw sztucznych wytworzonych w UE w 2022 r. wyniosła 16,16 miliona ton, z czego 6,58 miliona ton, czyli 40,7%, zostało poddanych recyklingowi. W Polsce poziom recyklingu tworzyw sztucznych wyniósł w 2022 r. 21,2%, obejmując 542 tys. ton odpadów tworzyw [4].

Według Europejskiej Agencji Środowiska unijna gospodarka tworzywami sztucznymi wygenerowała 193 mln ton emisji CO_2e, w tym 63% pochodzi z produkcji materiałów polimerowych, 22% z przetwarzania, a 15% z utylizacji (przede wszystkim spalania). Ponadto z szacunków wynika, że każdego roku od 19 do 23 mln ton tworzyw sztucznych trafia do środowiska naturalnego [3, 4].

Tworzywa sztuczne mają w wybranych zastosowaniach ciekawe właściwości fizyczne (odporność chemiczna, niska gęstość, dobre właściwości mechaniczne, łatwość przetwarzania i formowania, niska cena). Cechy te sprawiają, że tworzywa sztuczne są chętnie wybieranymi materiałami przez producentów i inżynierów. Zapotrzebowanie na nie rozszerza i upowszechnia się wraz z rosnącą konsumpcją wynikającą z poprawy jakości życia społeczeństwa. Istotna część emisji gazów cieplarnianych związanych z łańcuchem wartości tworzyw sztucznych przypada na etap produkcji polimerów.

FDM (z ang. *Fused Deposition Modeling*) oraz FFF (z ang. *Fused Filament Fabrication*) to technologie przyrostowe polegające na warstwowym osadzaniu uplastycznionego materiału termoplastycznego w postaci filamentu. Filament jest podawany do nagrzanej głowicy, w której ulega stopieniu, a następnie wytłaczany przez dyszę zgodnie z trajektorią wyznaczoną na podstawie modelu CAD i pliku sterującego (G-code). Materiał odkładany jest warstwa po warstwie na platformę roboczą aż do uzyskania docelowej geometrii elementu. W zależności od konstrukcji urządzenia możliwe jest przetwarzanie jednego lub kilku materiałów jednocześnie [5].

Popularną domową inicjatywą związaną z recyklingiem jest wytwarzanie filamentu z butelek PET. Urządzenie do produkcji takiego filamentu może być w większości wykonane metodą druku 3D. Zawiera ono silnik podający taśmę przez rozgrzaną dyszę, szpulę mocującą filament, przekładnię, nadmuch chłodzący gotowy filament, kontroler prędkości, kontroler temperatury oraz podajnik. Pierwszy etap produkcji filamentu to przygotowanie taśmy z butelki PET obejmujące usunięcie etykiety i resztek kleju z butelki, wygładzenie butelki za pomocą opalarki, odcięcie dna butelki, osuszenie jej wnętrza oraz wycięcie wstęgi [6].

W przemyśle produkcję filamentu wykorzystującego tworzywa z recyklingu przeprowadza się na większą skalę. Coraz więcej przedsiębiorstw dostrzega potencjał w regranulacji termoplastycznych surowców wtórnych oraz tworzeniu linii do przetwarzania granulatu w filament służący do druku 3D. Plastik pochodzący z odpadów jest przerabiany na regranulat, a następnie podawany do ogrzewanego bębna wytłaczarki za pośrednictwem leja zasypowego i wytłaczany na filament do druku 3D w procesie ciągłym. Wytłoczony filament jest schładzany i cięty na zadaną długość, a następnie nawijany na szpulę. W celu modyfikacji właściwości tworzywa przed wytłaczaniem regranulat miesza się z granulatem materiału pierwotnego i dodatkami [7].

Pomimo rosnącej liczby badań dotyczących zastosowania recyklatów w druku 3D, brakuje kompleksowego opracowania integrującego aspekty materiałowe, technologiczne i środowiskowe tego zagadnienia. Celem niniejszej pracy jest krytyczna analiza dostępnej literatury oraz identyfikacja kluczowych ograniczeń i kierunków dalszych badań w zakresie wykorzystania tworzyw wtórnych w technologii FDM/FFF.

2. Źródła i charakterystyka recyklatów polimerowych

2.1. Rodzaje tworzyw stosowanych w druku 3D

Druk 3D w technologii FDM/FFF charakteryzuje się wykorzystaniem szerokiej gamy polimerów termoplastycznych obecnych w gospodarce. Jednak ze względu na stosunkowo niskie wymagania przetwórcze najczęściej stosowanymi materiałami są:

ABS (terpolimer akrylonitrylo-butadieno-styrenowy) jest jednym z najwcześniej i najczęściej stosowanych materiałów w technologii FDM/FFF. Swoją popularność zawdzięcza dobrej udarności, stosunkowo wysokiej odporności cieplnej oraz podatności chemicznej na aceton, co umożliwia wygładzanie i łączenie detali. ABS jest polimerem amorficznym o temperaturze zeszklenia Tg wynoszącej około 100–110°C, natomiast temperatura ugięcia pod obciążeniem (HDT) zależy od gatunku i warunków badania i zwykle mieści się w zakresie 85–100°C. Typowe temperatury przetwórstwa ABS w druku 3D mieszczą się w zakresie 230–260°C [7,8]. Ze względu na stosunkowo wysoki skurcz przetwórczy i podatność na wypaczanie większe elementy wymagają stosowania drukarek z zamkniętą, ogrzewaną komorą

roboczą (ok. 40–60°C) [9]. Przetwarzanie ABS wiąże się z emisją lotnych związków organicznych (VOC) oraz cząstek ultradrobnych, dlatego zaleca się stosowanie odpowiedniej wentylacji lub filtracji powietrza [10]. ABS, jako materiał pochodzenia petrochemicznego, nie ulega biodegradacji w warunkach naturalnych.

PLA (kwas polimlekowy) jest alifatycznym poliestrem otrzymywanym z surowców odnawialnych, takich jak skrobia kukurydziana, trzcina cukrowa czy inne biomasy zawierające cukry fermentowalne [11]. W porównaniu z polimerami pochodzenia petrochemicznego charakteryzuje się częściowo biopochodnym łańcuchem dostaw oraz możliwością biodegradacji w warunkach przemysłowego kompostowania. W technologii FDM/FFF PLA jest jednym z najczęściej stosowanych materiałów ze względu na dobrą przetwarzalność, stosunkowo niski skurcz przetwórczy oraz stabilność wymiarową wydruków [11]. W wielu przypadkach możliwe jest drukowanie bez stosowania zamkniętej komory roboczej, a wymagane temperatury dyszy są niższe niż w przypadku ABS i zwykle mieszczą się w zakresie 190–230°C [12]. Jednocześnie PLA wykazuje ograniczoną odporność cieplną — temperatura ugięcia pod obciążeniem (HDT) wynosi zazwyczaj około 50––60°C, co ogranicza jego zastosowania w elementach pracujących w podwyższonej temperaturze [13]. Parametry druku, w tym prędkość, powinny być dobierane z uwzględnieniem właściwości konkretnego gatunku filamentu oraz warunków procesu, aby zapewnić stabilność ekstruzji i odpowiednią adhezję międzywarstwową [14].

PETG (politereftalan etylenu modyfikowany glikolem) jest kopoliestrem powstałym w wyniku modyfikacji klasycznego PET poprzez wprowadzenie dodatkowego składnika glikolowego, co ogranicza zdolność materiału do krystalizacji i prowadzi do uzyskania struktury w dużej mierze amorficznej [15]. W efekcie PETG charakteryzuje się dobrą stabilnością wymiarową, wysoką przezroczystością oraz stosunkowo dobrą udarnością przy zachowaniu łatwości przetwarzania w technologii FDM/FFF [16]. Odporność cieplna tego materiału, wyrażana temperaturą ugięcia pod obciążeniem (HDT), wynosi zwykle około 65–75°C (w zależności od gatunku i warunków badania) [17, 18].

PA (poliamidy) to grupa materiałów, gdzie najczęściej stosowanymi w technologii druku polimerami z tej grupy są PA6 oraz PA12. Do największych zalet tych materiałów należy niski współczynnik tarcia, doskonała wytrzymałość mechaniczna oraz wysoka odporność temperaturowa, osiągająca nawet 120–160°C (dla wybranych gatunków PA) [19]. Druk z poliamidów wiąże się jednak z istotnymi wyzwaniami: materiały te są silnie higroskopijne, co wymusza rygorystyczne osuszanie filamentu przed procesem, aby uniknąć degradacji polimeru [20]. Ponadto, podobnie jak ABS, PA wykazuje znaczny skurcz przetwórczy i podatność na wypaczanie, co wymaga stabilnych warunków temperaturowych. Ta wysoka chłonność wilgoci jest kluczowa również w kontekście recyklingu, ponieważ zawartość wody w recyklacie bezpośrednio wpływa na pogorszenie właściwości reologicznych i wytrzymałościowych wtórnie przetworzonego surowca [21].

PC (poliwęglany) – materiały z tej grupy należą do najbardziej wymagających, a zarazem najbardziej wytrzymałych termoplastów stosowanych w technologii FDM. Do największych zalet poliwęglanu należy wysoka udarność oraz sztywność i stabilność termiczna, której temperatura ugięcia pod obciążeniem (HDT) może wynosić od 110°C do nawet 140°C [22]. Jednak drukz poliwęglanu wiąże się z istotnymi wyzwaniami technicznymi: materiał ten charakteryzuje się wysoką temperaturą przetwarzania (zwykle 280–310°C) oraz znacznym skurczem przetwórczym w warunkach druku FDM. Wymusza to stosowanie aktywnie grzanych komór roboczych, aby zapobiec rozwarstwianiu się warstw (delaminacji) i deformacjom modelu [23]. Podobnie jak PA, poliwęglan jest silnie higroskopijny. Nawet niewielka ilość wilgoci w filamencie prowadzi do zjawiska hydrolizy podczas druku, co objawia się pęcherzykami gazu w strukturze i drastycznym spadkiem wytrzymałości mechanicznej [23]. Ta wrażliwość na wodę jest krytyczna w procesie recyklingu – obecność wilgoci w regranulacie PC podczas ponownego przetwórstwa termicznego powoduje nieodwracalną degradację łańcuchów polimerowych, co skutkuje gwałtownym wzrostem kruchości i utratą kluczowych właściwości konstrukcyjnych wtórnego surowca.

2.2. Źródła odpadów: poużytkowe a poprodukcyjne

Źródła tworzyw sztucznych na potrzeby recyklingu dzielą się na dwie zasadnicze grupy: odpady poużytkowe (pokonsumenckie) oraz poprodukcyjne. Każda z nich ma odmienną specyfikę.

Odpady poużytkowe to materiały które zostały wykorzystane zgodnie ze swoim przeznaczeniem i stały się odpadem w gospodarstwach domowych, sektorze usług lub handlu. W 2022 r. tylko mieszkańcy Unii Europejskiej wytworzyli ponad 16 milionów ton odpadów z tworzyw sztucznych, z czego 40,7% zostało ponownie przetworzonych, jak wspomniano we wprowadzeniu. Pod względem wolumenu stanowią bardzo istotne źródło recyklatu, jednakże ich specyfika sprawia, że są one bardzo trudnym w przetwórstwie materiałem. Odpady poużytkowe składają się z bardzo wielu typów polimerów wykorzystywanych każdego dnia. Są to worki foliowe wykonane z polietylenu, opakowania po kosmetykach wytworzone z polipropylenu lub PET, czy też elementy artykułów gospodarstwa domowego wykonane z ABS lub poliamidu często wypełnionego włóknem szklanym [24]. Kolejną przeszkodą w przetwórstwie tych tworzyw jest wysoki stopień zanieczyszczenia organicznego (resztki żywności) i nieorganicznego (etykiety, kleje, smary, metale) a także ich potencjalna degradacja spowodowana ekspozycją na promieniowanie UV, ciepło i utlenianie podczas użytkowania. Proces recyklingu takich tworzyw wymaga wieloetapowego sortowania, oczyszczania, suszenia oraz granulacji, co zwiększa koszty przetwarzania i zwykle prowadzi do uzyskania materiału o parametrach niższych niż w przypadku surowca pierwotnego [24].

Odpady poprodukcyjne powstające w trakcie procesu przetwórstwa polimerów, nie mają kontaktu z zanieczyszczeniami w takim stopniu jak tworzywa

z wcześniej opisanej grupy, co znacznie upraszcza proces ich ponownego wykorzystania. Określenie składu takiego surowca jest dość proste, ponieważ każda fabryka prowadzi skrupulatną dokumentację przetwarzanych tworzyw, co pozwala na uzyskanie recyklatu o wysokiej czystości i parametrach zbliżonych do surowca bazowego. Potencjał ten dostrzega coraz więcej przetwórców polimerów, uzupełniając swoje zakłady w działy regranulacji zajmujące się przywróceniem odpadów poprodukcyjnych do strumienia produkcji [25].

2.3. Wpływ liczby cykli przetwórstwa na degradację łańcuchów polimerowych

Zjawiskiem towarzyszącym recyklingowi tworzyw termoplastycznych jest skracanie łańcuchów polimerowych powodujące pogorszenie właściwości mechanicznych (na rozciąganie, zginanie, ściskanie) oraz zmęczeniowej, odporności na uszkodzenia, elastyczności, zdolności do odkształcania. Może dojść do zmian temperatur charakterystycznych dla danego polimeru (np. temperatury zeszklenia Tg lub temperatury topnienia Tm – w zależności od rodzaju tworzywa) oraz spadku jego stabilności termicznej. Ponadto zwiększa się kruchość i ryzyko powstawania mikropęknięć. W procesie recyklingu degradacji mogą ulec dodatki chroniące przed promieniowaniem UV, przez co recyklat szybciej starzeje się pod wpływem światła słonecznego. Tworzywa przezroczyste, takie jak PET, mogą stracić swoją przejrzystość, stać się matowe, żółtawe lub nabrać odcień zanieczyszczeń takich jak resztki barwników. W niektórych przypadkach recyklaty mogą mieć zmienioną zdolność do przewodzenia lub izolowania elektryczności oraz równomiernego rozprowadzania ciepła [26].

Badania polipropylenu Moplen HP500N wykazały, że wielokrotne cykle przetwórstwa prowadzą do obniżenia lepkości stopu, co w procesie formowania wtryskowego skutkuje zmniejszeniem ciśnienia w gnieździe formy. Zjawisko to związane jest ze skracaniem łańcuchów polimerowych i wzrostem masowego wskaźnika płynięcia (MFR). W analizowanym przypadku MFR pięciokrotnie przetworzonego materiału wzrósł z około 5,7 do 10,8 g/10 min, natomiast naprężenia ścinające uległy obniżeniu o około 67% w stosunku do materiału pierwotnego [27]. Choć przytoczony przykład dotyczy przetwórstwa wtryskowego, analogiczne mechanizmy zmian reologicznych zachodzą w procesie wytłaczania filamentów do druku 3D.

Zmniejszona lepkość ułatwia proces wtrysku lub wytłaczania, szczególnie w przypadku skomplikowanych form wymagających szybkiego wypełniania wnęk. Ponadto możliwe jest dostosowywanie właściwości materiału do konkretnych wymagań lub polepszenie parametrów przez dodawanie stabilizatorów, plastyfikatorów, wypełniaczy lub innych dodatków [28].

2.4. Modyfikacje i dodatki kompensujące degradację

Ciągle rosnące wykorzystanie polimerów we wszystkich gałęziach przemysłu doprowadziło do szerokiego rozwoju wypełniaczy i modyfikatorów dodawanych

w procesie przetwórstwa tych materiałów. Kluczowe parametry użytkowe takie jak sztywność, odporność na ścieranie czy odporność termiczna mogą zostać podniesione poprzez uzupełnienie kompozycji dodatkami o różnym wpływie na otrzymywany ostatecznie materiał.

Dodatki stabilizujące, ich rolą podczas przetwórstwa lub recyklingu polimerów jest zapobieganie degradacji termicznej oraz radiacyjnej, która z czasem wpływa na jakość produktu polimerowego. Wraz z upływem czasu skuteczność stabilizatorów spada, dlatego w procesie recyklingu niezbędna jest restabilizacja pozwalająca, w pewnym stopniu, na przywrócenie recyklatowi parametrów materiału bazowego. W przemyśle często stosowane są stabilizatory HALS [29] lub antyoksydanty fenolowe, w przypadku recyklingu PET dodawane są antyoksydanty fosforanowe [30].

Wypełniacze włókniste mogą poprawić niektóre właściwości mechaniczne, takie jak moduł sprężystości i wytrzymałość na rozciąganie, ale pogarszają przetwarzalność oraz wartość wydłużenia przy zerwaniu. Jednocześnie, w wielu przypadkach, ich użycie przynosi korzystny wpływ na właściwości mechaniczne także w warunkach pracy o podniesionej temperaturze. Dostępne badania wskazują, że dodanie włókien szklanych do niekompatybilnych mieszanek PET/HDPE powoduje znaczną poprawę modułu sprężystości, wytrzymałości na rozciąganie i udarności materiału przy niewielkiej zmianie wydłużenia przy zerwaniu [31].

Modyfikatory, w przeciwieństwie do wypełniaczy, poprawiają wydłużenie przy zerwaniu oraz udarność, które to parametry są bardziej wrażliwe na degradację oraz niekompatybilność tworzyw wykorzystywanych do produkcji regranulatu. Można to osiągnąć poprzez zastosowanie niewielkich ilości elastomerów.

Środki sprzęgające (kompatybilizatory) stosowane są w przypadku tworzyw zawierających wypełniacze, ich zadaniem jest zwiększenie interakcji pomiędzy polimerem, a materiałem wypełniającym. Tworzą one wiązania chemiczne między cząsteczkami w celu poprawy adhezji między materiałami. Ponieważ zwiększają one wiązania pomiędzy fazami, mogą być wykorzystywane do łączenia materiałów które bez ich udziału są niekompatybilne co ma kluczowe zastosowanie w dziedzinie recyklingu polimerów [31]. Przykładowo, w przypadku przetwórstwa PP często stosowanymi kompatybilizatorami są silany (związki krzemu) oraz bezwodnik maleinowy [32].

3. Technologia wytwarzania filamentów z recyklatów

Rozwój technologii przyrostowych, zwłaszcza druku 3D metodą FDM, spowodował zwiększone zainteresowanie użyciem surowców wtórnych jako zamiennika materiałów pierwotnych. Wytwarzanie filamentów z materiałów recyklingowych jest kluczowym aspektem gospodarki o obiegu zamkniętym, mającym na celu zmniejszenie odpadów i ograniczenie zużycia energii. Przetwarzanie polega na przygotowaniu regranulatu, jego wytłaczaniu, ocenie jakości oraz badaniu cech

końcowych filamentów. W tym rozdziale skupiono się na przedstawieniu technologii wytwarzania filamentów z tworzyw pochodzących z recyklingu, z uwzględnieniem różnic między skalą laboratoryjną a przemysłową oraz parametrów procesu ekstruzji, metod kontroli jakości oraz porównania właściwości mechanicznych filamentów z surowców pierwotnych i wtórnych [33].

3.1. Linie do wytłaczania

Proces wytłaczania filamentów z materiałów zrecyklingowanych może być przeprowadzany zarówno w warunkach laboratoryjnych, jak i przemysłowych. W ramach badań laboratoryjnych wykorzystuje się zintegrowane, małe linie ekstruzyjne z jednoślimakowymi wytłaczarkami o niskiej wydajności. Urządzenia tego typu umożliwiają szczegółową analizę wpływu parametrów procesu, takich jak temperatura stref grzewczych, prędkość ślimaka i chłodzenia, na cechy wytwarzanego filamentu. W warunkach przemysłowych używa się wytłaczarek dwuślimakowych, które zapewniają lepszą jednorodność stopu i bardziej precyzyjne mieszanie składników, zwłaszcza w przypadku filamentów modyfikowanych dodatkami. Zautomatyzowane systemy chłodzenia, kontrola średnicy w czasie rzeczywistym (realizowana z użyciem bramek laserowych) i automatyczne zwijanie na szpule to kluczowe elementy linii przemysłowej. W każdym z tych przypadków istotnym etapem jest przygotowanie regranulatu poprzez oczyszczenie, suszenie i rozdrobnienie. Aby zapobiec hydrolizie i powstawaniu pęcherzy powietrza w strukturze filamentu, polimery takie jak PLA czy PETG wymagają suszenia w temperaturze 60–70°C przez 4–6 godzin [34].

3.2. Parametry ekstruzji a jakość filamentu

Jakość filamentu oceniana jest przede wszystkim na podstawie stabilności średnicy, owalności, chropowatości powierzchni oraz obecności pęcherzy gazowych. Temperatury stref grzewczych, prędkość obrotowa ślimaka, ciśnienie stopu oraz szybkość chłodzenia należą do kluczowych parametrów procesu determinujących jakość filamentu. Zbyt wysoka temperatura może prowadzić do degradacji termicznej polimeru, natomiast zbyt niska – do niepełnego uplastycznienia regranulatu i powstawania defektów powierzchniowych. Poziom naprężeń wewnętrznych oraz charakter wykończenia powierzchni są silnie uzależnione od szybkości chłodzenia. Intensywne chłodzenie sprzyja uzyskaniu powierzchni o wysokim połysku, jednak może prowadzić do utrwalania naprężeń wewnętrznych, które w czasie eksploatacji mogą powodować relaksację struktury lub zmiany wymiarowe. Wolniejsze chłodzenie umożliwia częściową relaksację naprężeń, lecz zwykle skutkuje bardziej matowym wykończeniem powierzchni [35].

W badaniach nad recyklingiem PETG stwierdzono, że wielokrotne przetwarzanie może początkowo poprawić stabilność mechaniczną materiału podnosząc jego wytrzymałość na rozciąganie i zginanie aż do trzeciego lub czwartego cyklu recyklingu w wyniku zjawiska usztywnienia. Dopiero w kolejnych cyklach materiał traci swoje właściwości, dochodzi do pogorszenia spójności między

warstwami wydruku, powstawania widocznych szczelin i deformacji oraz znacznego spadku płynności (przepływu) polimeru, co wiąże się ze zmianami jego lepkości wynikającymi z sieciowania. Warto również zauważyć, że w przeciwieństwie do wytrzymałości na rozciąganie, udarność materiału spada z każdym kolejnym cyklem ze względu na narastającą kruchość polimeru poddanego obróbce termomechanicznej [36].

3.3. Metody kontroli jakości

Kontrola jakości filamentów polega na przeprowadzaniu pomiarów geometrycznych oraz analizie ich właściwości mechanicznych i termicznych. W laboratoriach badawczych używa się czujników laserowych do monitorowania średnicy filamentu w czasie rzeczywistym, co umożliwia natychmiastową korektę parametrów ekstruzji.

Badania mechaniczne wykonuje się zgodnie z normami ISO 527 – statyczna próba rozciągania oraz ISO 178 – statyczna próba zginania, natomiast analiza termiczna obejmuje techniki skaningowej kalorymetrii różnicowej (DSC) oraz termograwimetrii (TGA). Zastosowanie mikroskopii SEM umożliwia ocenę struktury powierzchni i porowatości filamentów. W ramach działań dodatkowych realizuje się testy druku próbnych próbek 3D, mające na celu ocenę jakości adhezji warstw oraz precyzji wymiarowej. Badania wykazują, że stabilność wymiarowa filamentów z recyklingu zależy od jednorodności regranulatu i stopnia degradacji polimeru [37].

3.4. Przykłady udanych mieszanin materiałowych

W ostatnich latach opracowano liczne kompozyty bazujące na materiałach z recyklingu, które charakteryzują się łatwością obróbki i korzystnymi właściwościami użytkowymi. Recyklingowany PLA zmodyfikowany włóknami konopnymi stanowi przykład, wykazując poprawę modułu sprężystości i odporności na kruche pękanie [38]. Inne badania dotyczyły mieszanin PET z PLA lub TPU, które poprawiają elastyczność i udarność wydruków 3D. Stwierdzono, że nawet niewielka ilość termoplastycznych elastomerów może znacznie poprawić odporność na zginanie i uderzenia [39]. W wielu przypadkach wykorzystywane są wypełniacze mineralne lub włókna naturalne, które poprawiają sztywność bez istotnego wzrostu masy.

3.5. Porównanie właściwości filamentów z surowców pierwotnych i z recyklingu

Analiza właściwości mechanicznych filamentów wykonanych z surowców pierwotnych i z recyklingu pokazuje, że materiały z recyklingu zazwyczaj cechują się nieznacznie mniejszą wytrzymałością na rozciąganie i niższą odpornością na wysokie temperatury [40]. Jednocześnie zauważono, że odpowiednie mieszanie recyklatu z surowcem pierwotnym pozwala uzyskać filament o właściwościach

zbliżonych do materiału bazowego. W przypadku PLA wielokrotne cykle przetwarzania prowadzą do skracania łańcuchów polimerowych, co obniża wytrzymałość mechaniczną, ale można częściowo odzyskać właściwości, stosując dodatki reologiczne lub stabilizatory.

4. Właściwości i jakość wydruków 3D z recyklatów

4.1. Wpływ degradacji materiału na właściwości wydruków

Proces recyklingu polimerów stosowanych w druku 3D, takich jak PLA, PETG czy ABS, prowadzi do degradacji łańcuchów makrocząsteczkowych w wyniku działania temperatury, tlenu i naprężeń mechanicznych w czasie wytłaczania. Skracanie łańcuchów polimerowych powoduje obniżenie masy cząsteczkowej, co bezpośrednio wpływa na spadek lepkości materiału i pogorszenie spójności międzywarstwowej.

W konsekwencji w niektórych badaniach odnotowano zmniejszenie wytrzymałości na rozciąganie (nawet o 15–30% po jednym cyklu recyklingu) oraz modułu sprężystości (modułu Younga). Degradacja materiału oraz zmiany jego właściwości reologicznych sprzyjają powstawaniu nieciągłości strukturalnych, takich jak zwiększona porowatość i mikropęknięcia, co prowadzi do pogorszenia adhezji międzywarstwowej i obniżenia integralności mechanicznej wydruku [33].

Degradacja wpływa także na dokładność wymiarową wydruków. Zaobserwowano większe odchylenia wymiarów i tendencję do wypaczeń w elementach wykonanych z recyklatu, co tłumaczy się niejednorodnością lepkości stopionego materiału oraz podwyższoną zawartością wilgoci [41].

4.2. Analiza mikrostruktury i defektów

Badania mikrostruktury wydruków z recyklatu, prowadzone z wykorzystaniem mikroskopii skaningowej (SEM), wskazują na zwiększoną liczbę pustek międzywarstwowych, nieregularny przebieg granic między ścieżkami materiału oraz lokalne obszary nadtopienia. W wybranych badaniach raportowano, że wydruki z recyklatu PLA i PETG mogą wykazywać porowatość na poziomie około 8–12%, podczas gdy dla materiałów pierwotnych wartości te były zazwyczaj niższe (często poniżej 5%) [42]. Różnice te tłumaczone są obniżeniem masy cząsteczkowej i zmianą właściwości reologicznych polimeru, co ogranicza zdolność do dyfuzji łańcuchów na granicy warstw i prowadzi do osłabienia adhezji międzywarstwowej.

Występujące defekty, takie jak mikropęknięcia, niepełne połączenia warstw i wtrącenia zanieczyszczeń (np. z wcześniejszych cykli druku), znacząco pogarszają adhezję międzywarstwową, a tym samym zmniejszają wytrzymałość zmęczeniową wydruku. Zjawiska te są szczególnie nasilone przy zbyt niskiej temperaturze druku lub przy niedostatecznym chłodzeniu warstw [43].

4.3. Porównanie właściwości mechanicznych wydruków

W wielu pracach raportowano, że właściwości mechaniczne wydruków wykonanych z recyklatu są niższe w porównaniu z elementami wytworzonymi z filamentu pierwotnego. W zależności od rodzaju polimeru, liczby cykli przetwórstwa oraz parametrów druku, spadek wytrzymałości na rozciąganie wynosił od kilkunastu do około 40%, modułu sprężystości od kilku do około 25%, natomiast udarność mogła ulec obniżeniu nawet o 50% [44].

Jednocześnie wykazano, że odpowiednie modyfikacje materiałowe oraz kontrola procesu mogą istotnie ograniczyć te różnice. W przypadku mieszanin rPLA z dodatkiem materiału pierwotnego lub przy zoptymalizowanym procesie regranulacji uzyskiwano właściwości zbliżone do materiału bazowego. Dla rPET i rPETG raportowano osiąganie 80–90% pierwotnej wytrzymałości mechanicznej przy właściwym suszeniu filamentu i kontroli temperatury przetwórstwa [45].

4.4. Wpływ parametrów druku na właściwości wydruków z recyklatu

Parametry procesu druku mają kluczowe znaczenie dla kompensacji obniżonej jakości materiałów z recyklingu. Odpowiednia optymalizacja parametrów druku może częściowo zniwelować skutki degradacji recyklatu.

W literaturze zaleca się następujące strategie procesowe (w zależności od rodzaju polimeru i systemu druku): zwiększenie temperatury dyszy o około 5–10°C, co może poprawić adhezję międzywarstwową i ograniczyć porowatość [46]; ograniczenie prędkości druku (np. do 30–40 mm/s) w celu zapewnienia odpowiedniego czasu przebywania materiału w strefie grzewczej i lepszej homogenizacji stopu [47]; stosowanie mniejszej grubości warstwy (w zakresie 0,15–0,2 mm), co sprzyja zwiększeniu spójności strukturalnej i redukcji pustek międzywarstwowych [48]; oraz suszenie filamentu przed drukiem, ponieważ obecność wilgoci w recyklacie prowadzi do powstawania pęcherzy gazowych i wzrostu porowatości [48].

Ponadto jakość i wytrzymałość wydruków z recyklatu można poprawić poprzez odpowiednią kontrolę parametrów procesu wytwarzania filamentu, w szczególności optymalizację prędkości ślimaka oraz temperatur przetwórstwa [33]. Ograniczenie intensywności ścinania i przegrzewania materiału sprzyja zmniejszeniu stopnia degradacji łańcuchów polimerowych. Dodatkowo możliwe jest mechaniczne modyfikowanie materiału poprzez wprowadzenie włókien węglowych lub szklanych, które zwiększają sztywność kompozytu, jednak zwykle kosztem wzrostu kruchości wydruków [49]. W wybranych badaniach wykazano, że zastosowanie mieszaniny zawierającej 70% rPLA i 30% materiału pierwotnego pozwalało odzyskać do około 95% pierwotnej wytrzymałości mechanicznej, przy zachowaniu odpowiednich parametrów przetwórstwa [50].

5. Podsumowanie

Recyklaty polimerowe wykazują istotny potencjał aplikacyjny w kontekście zrównoważonego druku 3D, stanowiąc obiecującą alternatywę dla surowców pierwotnych. Ich efektywne wykorzystanie wymaga jednak świadomego podejścia technologicznego, ponieważ wielokrotne przetwarzanie termiczne prowadzi do degradacji łańcuchów polimerowych. Proces ten skutkuje pogorszeniem adhezji międzywarstwowej, spadkiem właściwości mechanicznych oraz wzrostem porowatości struktury wydruków.

Kluczowym warunkiem uzyskania materiału o stabilnych parametrach jest odpowiednie przygotowanie regranulatu oraz precyzyjna kontrola procesu wytwarzania filamentu. Analiza literatury wskazuje, że negatywne skutki recyklingu można w znacznym stopniu ograniczyć poprzez optymalizację parametrów druku oraz stosowanie modyfikacji materiałowych, takich jak dodatki stabilizujące, kompatybilizatory czy wzmocnienia włókniste.

W perspektywie dalszego rozwoju istotne znaczenie będzie mieć doskonalenie systemów monitorowania jakości ekstruzji oraz rozwój zaawansowanych metod kontroli właściwości recyklatów. Dalsze badania nad modyfikacją wtórnych surowców mogą przyczynić się do zwiększenia ich konkurencyjności względem materiałów pierwotnych i skuteczniejszego domykania obiegu tworzyw sztucznych w gospodarce.

Literatura

[1] G. Budzik, J. Woźniak, Ł. Przeszłowski „Druk 3D jako element przemysłu przyszłości. Analiza rynku I tendencje rozwoju", Oficyna Wydawnicza Politechniki Rzeszowskiej (2022).

[2] S. Kabir, Y. Li, Y. Lee "Critical Dialogs of 3D Printing Waste Recycling in Circular Fashion", Business, Management and Economics. IntechOpen (2025).

[3] M. Andrzejewska, M. Krawczyk, P. Postawa, „Tworzywa sztuczne. W gospodarce, nie w środowisku", Raport UNEP/GRID Warszawa (2024).

[4] Parlament Europejski „Recykling odpadów z tworzyw sztucznych w UE: fakty i liczby", artykuł informacyjny. Portal internetowy Parlamentu Europejskiego, URL: www.europarl.europa.eu/topics/pl/article/20181212STO21610/recykling-odpadow-z-tworzyw-sztucznych-w-ue-fakty-i-liczby (dostęp: 12.11.2025).

[5] F. M. Mwema, E. Akinlabi "Basics of Fused Deposition Modelling (FDM)", SpringerBriefs in Applied Sciences and Technology, Springer (2020).

[6] J. Blaut, W. Leszczyk, R. Rumin „Perspektywa zastosowania maszyn małoseryjnych do produkcji filamentu inżynierii produkcji i zarządzania odpadami", Systemy Wspomagania w Inżynierii Produkcji, vol. 13, nr 2, s. 174–182 (2024).

[7] F. J. Al Gawhari, A. A. Al-Zaidi "Types of Polymers Using in 3D Printing and Their Applications: A Brief Review", European Journal of Theoretical and Applied Sciences. AMO Publisher (2024).

[8] Spectrum Filaments „Smart ABS", Karta materiałowa/Specyfikacja. URL: spectrumfilaments.com/en/filament/smart-abs (dostęp: 12.11.2025).

[9] J. Ramian, J. Ramian, D. Dziob „Thermal Deformations of Thermoplast during 3D Printing: Warping in the Case of ABS", Materials, vol. 14, nr 22, nr art. 7070. MDPI (2021).

[10] P. Byrley, B. J. George, W. K. Boyes, K. Rogers „Particle emissions from fused deposition modeling 3D printers: Evaluation and meta-analysis", Science of The Total Environment, vol. 655, s. 395–407. Elsevier (2019).

[11] T. M. Joseph, A. Kallingal, A. M. Suresh, D. K. Mahapatra, M. S. Hasanin, J. Haponiuk, S. Thomas „3D printing of polylactic acid: recent advances and opportunities", International Journal of Advanced Manufacturing Technology, vol. 125, nr 3-4, s. 1015–1035. Springer Nature (2023).

[12] V. Cojocaru, D. Frunzaverde, C.-O. Miclosina, G. Marginean „The Influence of the Process Parameters on the Mechanical Properties of PLA Specimens Produced by Fused Filament Fabrication – A Review", Polymers, vol. 14, nr 5, nr art. 886. MDPI (2022).

[13] X. Zhao, J. Liu, J. Li, X. Liang, W. Zhou, S. Peng „Strategies and techniques for improving heat resistance and mechanical performances of poly(lactic acid) (PLA) biodegradable materials", International Journal of Biological Macromolecules, vol. 218, s. 423–443. Elsevier (2022).

[14] Spectrum Filaments „PLA Premium", Karta materiałowa/Specyfikacja. URL: spectrumfilaments.com/en/filament/pla-premium (dostęp: 12.11.2025).

[15] H.-H. Tsai, S.-J. Wu, Y.-D. Wu, W.-Z. Hong „Feasibility Study on the Fused Filaments of Injection-Molding-Grade Poly(Ethylene Terephthalate) for 3D Printing", Polymers, vol. 14, nr 11, nr art. 2276. MDPI (2022).

[16] K. Durgashyam, I. Reddy M, A. Balakrishna, K. Satyanarayana „Experimental investigation on mechanical properties of PETG material processed by fused deposition modeling method", Materials Today: Proceedings, vol. 18, nr 3, s. 1206–1214. Elsevier (2019).

[17] M. Garwacki, I. Cudnik, D. Dziadowiec, P. Szymczak, J. Andrzejewski „The Development of Sustainable Polyethylene Terephthalate Glycol-Based (PETG) Blends for Additive Manufacturing Processing – The Use of Multilayered Foil Waste as the Blend Component", Materials, vol. 17, nr 5, nr art. 1083. MDPI (2024).

[18] Spectrum Filaments „PCTG Premium", Karta materiałowa/Specyfikacja. URL: spectrumfilaments.com/en/filament/pctg-premium (dostęp: 12.11.2025).

[19] M. Shiferaw „Recycled Polymer for FDM 3D Printing Filament Material: Circular Economy for Sustainability of Additive Manufacturing", Advances of Science and Technology. Springer (2022).

[20] P. Wetzel, A. K. Sambale, K. Uhlig, M. Stommel, B. Schneider, J.-M. Kaiser „Hygromechanical Behavior of Polyamide 6.6: Experiments and Modeling", Polymers, vol. 15, nr 16, nr art. 3387. MDPI (2023).

[21] X. Zhang, W. Fan, T. Liu „Fused deposition modeling 3D printing of polyamide-based composites and its applications", Composites Communications, vol. 21, nr art. 100413. Elsevier (2020).

[22] X. Li, J. Mason „New Perspective on Heat Deflection Temperature of Glassy Polycarbonate", Macromolecular Materials and Engineering, vol. 294, nr 5, s. 306–314. Wiley-VCH (2009).

[23] L. Fang, Y. Yan, O. Agarwal, S. Yao i in. „Effects of Environmental Temperature and Humidity on the Geometry and Strength of Polycarbonate Specimens Prepared by Fused Filament Fabrication", Materials, vol. 13, nr 19, nr art. 4414. MDPI (2020).

[24] K. Bernat „Post-Consumer Plastic Waste Management: From Collection and Sortation to Mechanical Recycling", Energies, vol. 16, nr 8, nr art. 3504. MDPI (2023).

[25] M. Arese, I. Bolliri, G. Ciaccio, V. Brunella „Post-Industrial Recycled Polypropylene for Automotive Application: Mechanical Properties After Thermal Ageing", Processes, vol. 13, nr 2, nr art. 315. MDPI (2025).

[26] Dopak „Wyzwania recyklingu tworzyw termoplastycznych", Artykuł branżowy. URL: dopak.pl/wyzwania-recyklingu-tworzyw-termoplastycznych/ (dostęp: 12.11.2025).

[27] W. Frącz „Wpływ wielokrotnego przetwórstwa polimerów na parametry stanu tworzywa w formie wtryskowej", Zeszyty Naukowe Politechniki Rzeszowskiej nr 284. Oficyna Wydawnicza Politechniki Rzeszowskiej (2012).

[28] Spectrum Filaments „PA6 Neat Black", Karta materiałowa/Specyfikacja. URL: spectrum-filaments.com (dostęp: 25.02.2026).

[29] P. Gijsman „A review on the mechanism of action and applicability of Hindered Amine Stabilizers", Polymer Degradation and Stability, vol. 145, s. 2–10. Elsevier (2017).

[30] F. Awaja, D. Pavel „Recycling of PET", European Polymer Journal, vol. 41, nr 7, s. 1453––1477. Elsevier (2005).

[31] R. Francis "Recycling of Polymers. Methods, Characterization and Applications", John Wiley & Sons (2016).

[32] L. M. G. Araujo, A. Morales „Compatibilization of recycled polypropylene and recycled poly (ethylene terephthalate) blends with SEBS-g-MA", Polímeros, vol. 28, nr 1, s. 84––91. ABPol (2018).

[33] I. Ibrahim, A. G. Ashour, W. Zeiada i in. „ A Systematic Review on the Technical Performance and Sustainability of 3D Printing Filaments Using Recycled Plastic", Sustainability, vol. 16, nr 18, nr art. 8247. MDPI (2024).

[34] Fiberlogy „Easy PET-G", Karta materiałowa/Specyfikacja. URL: fiberlogy.com (dostęp: 23.02.2026).

[35] J. D. Seno Flores i in. "Sustainable polymer reclamation: recycling poly(ethylene terephthalate) glycol (PETG) for 3D printing applications", Journal of Materials Science: Materials in Engineering, vol. 19, nr art. 16. Springer (2024).

[36] N. Vidakis, M. Petousis, L. Tzounis, S. A. Grammatikos, E. Porfyrakis, A. Maniadi, N. Mountakis „Sustainable Additive Manufacturing: Mechanical Response of Polyethylene Terephthalate Glycol over Multiple Recycling Processes", Materials, vol. 14, nr 5, nr art. 1162. MDPI (2021).

[37] M. Kujawa, J. Głowacka, W. Pawlak, B. Sztorch, D. Pakuła, M. Frydrych, J. Sokolska, R. E. Przekop „Molybdenum Disulphide Modified Polylactide for 3D Printed (FDM/FFF) Filaments", Polymers, vol. 15, nr 10, nr art. 2236. MDPI (2023).

[38] J. Hege, G. Lee "Extrusion of Recycled PLA 3D Filament Compounded with Industrial Hemp", Raport z projektu badawczego (MC REU). Pennsylvania State University (2021).

[39] K. Çava, M. Aslan „Investigating Printability and Mechanical Performance of 3D Printed Recycled PET with PLA and TPU Hybrid Additives", International Journal of 3D Printing Technologies and Digital Industry, vol. 7, nr 2. DergiPark (2023).

[40] I. Anderson „Mechanical Properties of Specimens 3D Printed with Virgin and Recycled Polylactic Acid", 3D Printing and Additive Manufacturing, vol. 4, nr 2, Mary Ann Liebert, Inc. (2017).

[41] F. A. Cruz Sanchez, H. Boudaoud, M. Camargo, J. M. Pearce „Plastic recycling in additive manufacturing: A systematic literature review and opportunities for the circular economy", Journal of Cleaner Production, vol. 264, nr art. 121602. Elsevier (2020).

[42] A. D. Dobrzańska-Danikiewicz, B. Siwczyk, A. Bączyk, A. Romankiewicz „Mechanical properties of recycled PLA and PETG printed by FDM/FFM method", Journal of

Achievements in Materials and Manufacturing Engineering, vol. 119, nr 2, s. 49–59 (2023).

[43] M. Axinte, D. L. Chicet, R. Chelariu, R. I. Comăneci „The Influence of Processing Parameters on the Mechanical Properties of PLA 3D Printed Parts", Archives of Metallurgy and Materials, nr art. 151433 (2024).

[44] A. Al Rashid, M. Koç „Additive Manufacturing for Sustainability and Circular Economy: Needs, Challenges, and Opportunities for 3D Printing of Recycled Polymeric Waste", Materials Today Sustainability, vol. 24, nr art. 100529. Elsevier (2023).

[45] A. Rodríguez-Panes, J. Claver, A. M. Camacho „The Influence of Manufacturing Parameters on the Mechanical Behaviour of PLA and ABS Pieces Manufactured by FDM: A Comparative Analysis", Materials, vol. 11, nr 8, nr art. 1333. MDPI (2018).

[46] M. A. Caminero, J. M. Chacón, E. García-Plaza, P. J. Núñez, J. García-Moreno, G. Rodríguez „Additive manufacturing of 3D reinforced structures: effects of process parameters on mechanical properties", Revista de Materiales Compuestos, vol. 3, nr 3, Scipedia (2022).

[47] M. Behzadnasab, A. A. Yousefi, D. Ebrahimibagha, F. Nasiri „Effects of processing conditions on mechanical properties of PLA printed parts", Rapid Prototyping Journal, nr 2, s. 381–389. Emerald (2019).

[48] N. Vidakis i in. „Sustainable Additive Manufacturing: Mechanical Response of Polypropylene over Multiple Recycling Processes", Sustainability, vol. 13, nr 1, nr art. 159. MDPI (2021).

[49] N. van de Werken, H. Tekinalp, P. Khanbolouki, S. Ozcan, A. Williams, M. Tehrani „Additively manufactured carbon fiber-reinforced composites: State of the art and perspective", Additive Manufacturing, vol. 31, nr art. 100962. Elsevier (2020).

[50] N. Naveed, M. N. Anwar, M. Armstrong, F. Ahmad, M. I. U. Haq, G. Ridley „Enhancing Sustainability and Functionality with Recycled Materials in Multi-Material Additive Manufacturing", Sustainability, vol. 17, nr 13, nr art. 6105. MDPI (2025).

THE USE OF RECYCLED POLYMERS IN ADDITIVE MANUFACTURING – A REVIEW OF PROPERTIES AND TECHNOLOGICAL CHALLENGES

Summary

The role of additive manufacturing in the circular economy was analyzed, highlighting the potential of utilizing polymer waste as a raw material source for filament production. The suitability of popular thermoplastics for recycling was discussed, noting the significant differences between contaminated post-consumer waste and cleaner post-industrial waste. Filament manufacturing technologies were presented, emphasizing the necessity of proper material preparation—specifically cleaning and drying—to minimize structural defects. It was indicated that repeated processing of polymers leads to thermal-oxidative degradation and the shortening of macromolecular chains, resulting in deteriorated mechanical properties and reduced interlayer adhesion of the prints. Concurrently, methods to mitigate these effects were discussed, such as the use of stabilizers, blending recyclate with virgin material, and optimizing printing parameters. Literature analysis indicates that, provided a proper technological regime is maintained, recycled materials can serve as a valuable alternative to virgin plastics.

Keywords: recycling, additive manufacturing, filament, polymers, circular economy.

Aneta Miazga
Koło Naukowo-Badawcze Turystyki „Tuptuś”

STRATEGIE OPTYMALIZACJI WYDATKÓW W TURYSTYCE STUDENCKIEJ – ANALIZA POSTAW POKOLENIA Z

Streszczenie

Rozdział analizuje ewolucję modelu podróżowania pokolenia Z, traktującego turystykę jako kluczowe narzędzie samorealizacji. Głównym celem publikacji było zidentyfikowanie strategii stosowanych przez studentów w celu niwelowania barier finansowych przy zachowaniu mobilności. Przeprowadzone badania empiryczne na grupie 115 osób jednoznacznie potwierdziły postawioną w pracy tezę, wykazując dominację modelu budżetowego nad standardowymi ofertami biur podróży. Współczesny młody podróżnik jawi się w badaniach jako świadomy „inżynier finansowy”, który odchodzi od archaicznego, pasywnego oszczędzania na rzecz aktywnej optymalizacji relacji jakości do nakładów. Fundamentem tej inżynierii jest transport, stanowiący główny obszar generowania oszczędności. Optymalizację w tym zakresie wymusza stała obserwacja fluktuacji rynkowych, wykorzystywanie algorytmów wyszukiwarek oraz radykalna elastyczność w doborze terminów. Analiza dowiodła także, iż pokolenie to masowo zaprzęga nowoczesne technologie cyfrowe do bezprowizyjnego zarządzania kapitałem. Wykorzystanie instrumentów fintechowych oraz aplikacji do precyzyjnego rozliczania kosztów grupowych trwale eliminuje błędy w sprawiedliwym podziale wydatków. Istotną składową strategii pozostaje eksploracja darmowych zasobów kultury, lokalizowanych głównie poprzez media społecznościowe, takie jak TikTok czy Instagram. Zamiast korzystać z usług komercyjnych, młodzi ludzie stawiają na cyfrowy spryt, który pozwala im wydłużyć czas trwania wyprawy poprzez konwersję oszczędności na dodatkowe dni w terenie. Całość publikacji prowadzi do konkluzji, iż wysoka biegłość technologiczna pozwala studentom na niezwykle skuteczne operowanie ograniczonym kapitałem. Dzięki temu wyjazd turystyczny przestaje być postrzegany jako koszt, a staje się precyzyjną, wysoce efektywną operacją logistyczną o charakterze inwestycyjnym.

Słowa kluczowe: turystyka niskobudżetowa, minimalizacja kosztów, inżynieria finansowa, mobilność cyfrowa.

1. Wprowadzenie

Współczesny model podróżowania uległ głębokiej przemianie, stając się – z luksusu dostępnego dla nielicznych – powszechnym elementem stylu życia. Ewolucja ta jest szczególnie widoczna w postawach pokolenia Z, dla którego mobilność turystyczna przestała być wyznacznikiem wysokiego statusu materialnego, a stała się standardowym narzędziem samorealizacji i zdobywania doświadczeń. W tym ujęciu bariery finansowe nie stanowią już głównej przeszkody w podróżowaniu, lecz stają się impulsem do poszukiwania nowych strategii optymalizacji wydatków. Istotnym przejawem tej zmiany stała się indywidualna turystyka niskobudżetowa, zwana także niskokosztową (ang. *low-cost tourism, low-budget tourism*). Obejmuje ona wyjazdy turystyczne, realizowane samodzielnie przez pojedyncze osoby lub grupy dwuosobowe, które odbywają się przy jak najniższym wykorzystaniu środków finansowych [4].

Dla pokolenia Z, czyli młodych ludzi wchodzących dopiero na rynek pracy, podróżowanie przestało być nagrodą za zgromadzony kapitał, a stało się stylem życia, do którego dostęp uzyskują dzięki sprytowi i biegłości technologicznej. Celem niniejszego rozdziału jest zbadanie, które metody minimalizacji kosztów podróży są najczęściej stosowane przez studentów – przedstawicieli pokolenia Z.

W opracowaniu postawiono tezę, według której studenci, stosując strategie optymalizacji wydatków, najczęściej wybierają budżetowe wyjazdy o charakterze turystycznym.

2. Podstawowe modele wyjazdów turystycznych

Dzisiejszy rynek turystyczny opiera się na trzech głównych modelach podróżowania. Różnią się one oszczędnością czasu potrzebnego na ich zaplanowanie, a także ceną i poziomem wygody.

Model budżetowy opiera się na paradygmacie samodzielności i optymalizacji wydatków. Turysta samodzielnie planuje, przygotowuje oraz realizuje wyjazd, rezygnując z gotowych ofert biur turystycznych. Jako formę transportu wybiera podróż autostopem, rowerem, pieszą wędrówkę, tanie loty, oferty last minute lub darmowe przewozy oferowane na grupach „facebookowych" [4]. Miejscem noclegu jest najczęściej hostel lub pokoje wieloosobowe, co sprzyja integracji z lokalną społecznością i innymi podróżnikami.

Alternatywę stanowi model standardowy, będący złotym środkiem między oszczędnością a wygodą. Najczęściej jest on realizowany poprzez korzystanie z gotowych pakietów oferowanych przez biura podróży, obejmujących nocleg, wyżywienie i transport [9]. Zakwaterowanie w ośrodkach o średnim standardzie (hotele trzygwiazdkowe) oraz wsparcie rezydenta gwarantują spokój ducha, eliminując konieczność samodzielnego planowania każdego etapu wyjazdu.

Dla najbardziej wymagających podróżników dedykowany jest model premium, stawiający na bezkompromisowy komfort i najwyższą jakość obsługi. Wybór luksusowych hoteli (cztero- i pięciogwiazdkowych) oraz prywatnych transferów eliminuje logistyczne niedogodności, pozwalając skupić się na estetyce otoczenia i relaksie. Dzięki indywidualnie dopasowanym programom zwiedzania i personelowi, który dyskretnie uprzedza potrzeby gości, wyjazd zmienia się w ekskluzywne doświadczenie regeneracyjne.

3. Teoretyczne podstawy zarządzania budżetem i nadrzędnymi potrzebami pokolenia Z

Zarządzanie budżetem w turystyce studenckiej wykroczyło poza ramy tradycyjnego oszczędzania, ewoluując w stronę świadomej i wielowymiarowej inżynierii finansowej. Proces ten nie opiera się już na pasywnej redukcji wydatków, lecz na aktywnym, strategicznym projektowaniu struktury nakładów w celu maksymalizacji ich efektywności. Dla pokolenia Z, ukształtowanego przez cyfryzację oraz globalny przepływ informacji, tradycyjny podział na konsumpcję i oszczędności traci na aktualności. W nowym paradygmacie podróż nie jest traktowana jako jednorazowy akt zużycia zasobów, lecz jako świadomy proces budowania kapitału osobistego. Stanowi ona inwestycję w rozwój umiejętności interpersonalnych, wzmocnienie poczucia niezależności, poszerzanie horyzontów, a także w refleksję nad własnym życiem [13].

Skuteczna logistyka wymaga precyzyjnego rozdzielenia kosztów stałych od zmiennych, co tworzy bazę świadomego planowania ekspedycji. Nakłady stałe – czyli wydatki nieuniknione, obejmujące transport do miejsca docelowego, nocleg, ubezpieczenie turystyczne, opłaty drogowe i urzędowe – są optymalizowane poprzez tzw. arbitraż czasowy. Wykorzystywanie algorytmów wyszukiwarek pozwala na rezerwację biletów z dużym wyprzedzeniem lub w momentach statystycznych korekt cenowych. Z kolei koszty zmienne, obejmujące wyżywienie, transport lokalny, atrakcje, bilety wstępu i pamiątki, podlegają dynamicznej optymalizacji. Decyzja o wyborze darmowych zasobów kultury zamiast usług komercyjnych bezpośrednio przekłada się na wydłużenie czasu trwania wyprawy, co konwertuje oszczędności na dodatkowe dni w terenie.

Istotnym elementem analizy teoretycznej jest zmiana postrzegania oszczędności na rzecz koncepcji *Value for Money* – optymalizacji relacji jakości do nakładów. Wybór konkretnej usługi nie jest determinowany wyłącznie najniższą kwotą na paragonie, lecz bilansem płynących z niej korzyści. W tej kalkulacji wyższa cena początkowa bywa akceptowalna, o ile gwarantuje wymierne oszczędności w innych obszarach, takich jak logistyka czy wyżywienie. Ostatecznym celem tak skonstruowanej strategii nie jest zatem najtańsza możliwa wyprawa, lecz osiągnięcie najwyższej efektywności posiadanego kapitału. Należy jednak zaznaczyć, że subiektywne postrzeganie stosunku jakości do ceny pozostaje zmienne i zależy od indywidualnych potrzeb podróżnika, miejsca oraz czasu realizacji wyjazdu [1].

4. Metody optymalizacji wydatków

Podróżowanie młodych ludzi stało się obszarem świadomego zarządzania kapitałem. Zamiast rezygnować z jakości, młodzi podróżnicy wykorzystują swoją elastyczność i algorytmy, aby zamienić każdą złotówkę w maksymalną wartość doświadczenia. Istotne znaczenie ma tu stwierdzenie, że podróżowanie w dobie technologii informacyjno-komunikacyjnych zmieniło swój dotychczasowy charakter, oferując bogaty zbiór udogodnień, w tym aplikacje mobilne [6], które zredefiniowały pojęcie budżetowania. Dla Pokolenia Z optymalizacja kosztów to proces wielopoziomowy, który ewoluuje od prostych nawyków po wyrafinowane techniki cyfrowe.

Fundamentem optymalizacji w przemieszczaniu się są radykalna elastyczność i adaptacja do mechanizmów rynkowych. Strategie te zakładają budowanie planów wyjazdowych w oparciu o okazje cenowe generowane przez globalne algorytmy wyszukiwarek takich jak Skyscanner, Kayak czy Momondo [10], co pozwala na monitorowanie wahań taryf w czasie rzeczywistym. Rezygnacja ze sztywnych dat wyjazdu na rzecz terminów dyktowanych przez niskie słupki wykresów cenowych umożliwia znaczną redukcję wydatków.

Istotną rolę odgrywa tu również transport multimodalny, w ramach którego aplikacje pokroju Koleo czy FlixBus pozwalają na łączenie tanich przelotów z nocnymi przejazdami autokarowymi lub kolejowymi [11]. Takie rozwiązanie pełni podwójną funkcję optymalizacyjną: realizuje cel przemieszczenia się przy jednoczesnej eliminacji kosztu noclegu. Wśród najbardziej zaawansowanych strategii wymienia się technikę hidden city ticketing, polegającą na rezerwowaniu lotów z przesiadką, w których faktycznym celem podróży jest miasto tranzytowe, co pozwala zmniejszyć płatność do nawet 70% [5]. W takim przypadku pasażerowie podróżujący do miasta pośredniego mają motywację, aby udawać, że podróżują do celu, wysiąść z samolotu w punkcie przesiadkowym i zrezygnować z niewykorzystanej części biletu [8].

Standardem optymalizacji kosztów w sektorze zakwaterowania pozostaje korzystanie z hosteli oraz pokoi wieloosobowych. Pozwala to na radykalną redukcję wydatków jednostkowych, jednocześnie sprzyjając nawiązywaniu relacji między podróżnymi. Strategie oszczędnościowe obejmują także decentralizację lokalizacji. Wybór obiektów położonych poza ścisłym centrum oraz w miejscach dysponujących sprawną infrastrukturą transportową może okazać się wysoce efektywny dzięki krótkim przejazdom metrem bądź autobusem.

Analiza zjawiska wskazuje również na rosnące znaczenie ekonomii współdzielenia, oznaczającej angażowanie się wspólnie z innymi osobami w proces konsumpcji dóbr lub usług [2] – platformy takie jak Couchsurfing czy rozbudowana sieć kontaktów z programu Erasmus pozwalają niemal całkowicie wyeliminować koszty zakwaterowania. Dodatkowo biegłość technologiczna młodych ludzi umożliwia wykorzystywanie sieci VPN do symulowania logowania z regionów

o niższym parytecie siły nabywczej, co może skutkować uzyskaniem korzystniejszych stawek na globalnych portalach rezerwacyjnych [12].

Z kolei optymalizacja wyżywienia opiera się na umiejętnym balansowaniu między samodzielnym przygotowywaniem posiłków a korzystaniem z lokalnych okazji. Podstawą jest wybór noclegów z dostępem do kuchni, co pozwala na zakupy w lokalnych dyskontach i radykalne obniżenie kosztów „życia na mieście".

W sferze konsumpcji zewnętrznej teoretyczne ramy wyznacza trend ratowania żywności, w ramach którego na popularności zyskują aplikacje takie jak Too Good To Go czy Foodsi, oferujące dostęp do wysokiej jakości posiłków restauracyjnych za ułamek ich ceny rynkowej. Praktyki te obejmują również poszukiwanie autentycznych smaków w strefach street foodu oraz korzystanie z tzw. „lunch menu", przy jednoczesnym unikaniu restauracji nastawionych typowo na turystów, zwłaszcza w godzinach wieczornych.

Ostatnim filarem budżetowania jest inteligentne zarządzanie dostępem do kultury oraz finansami. Strategia ta zakłada śledzenie kalendarza darmowych wejść do muzeów oraz udział w „free" walking tours – pieszych wycieczek opartych na systemie dobrowolnych napiwków [3]. Media społecznościowe, takie jak TikTok czy Instagram, ułatwiają znalezienie darmowych atrakcji.

Niezbędnym wsparciem są również instrumenty fintechowe, takie jak Revolut czy Wise, które eliminują spready bankowe i prowizje za przewalutowania, chroniąc kapitał przed ukrytymi kosztami [7]. Strategie te obejmują także mechanizmy cashback (np. LetyShops, Picodi), dzięki którym każda rezerwacja pozwala generować realny zwrot części gotówki. Dopełnieniem tych metod jest wykorzystywanie aplikacji typu Splitwise, Tricount oraz przelewów BLIK – narzędzia te eliminują błędy w rozliczeniach grupowych, zapewniając sprawiedliwy podział kosztów wspólnych w czasie rzeczywistym.

Współczesna technologia i spryt pozwalają na przekształcić zwykłe oszczędzanie w przemyślaną, cyfrową strategię. Dzięki instrumentom fintechowym, sieciom VPN oraz ekonomii współdzielenia podróż przestaje być jedynie kosztem, a staje się precyzyjną operacją logistyczną. To świat, w którym algorytmy pracują na korzyść podróżnika, a biegłość w obsłudze aplikacji realnie zwiększa siłę nabywczą jego budżetu. Ostatecznie bariery finansowe znikają tam, gdzie pojawiają się rzetelna informacja i umiejętność jej praktycznego wykorzystania.

5. Analiza wyników badań

W pracy dokonano analizy metod optymalizacji kosztów podróży, które są najczęściej stosowane przez studentów – przedstawicieli pokolenia Z.

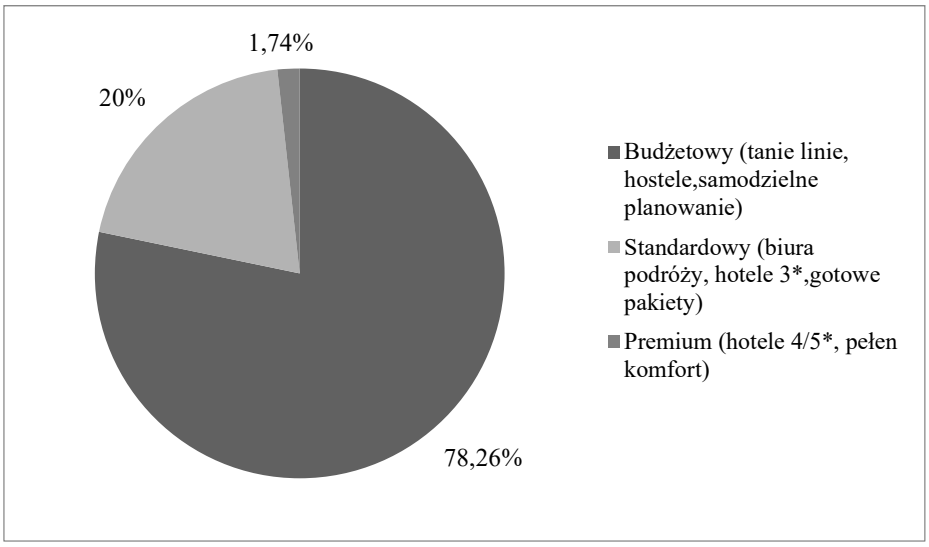
Analizie poddano 115 respondentów, którymi byli studenci uczelni wyższych. Odpowiedzi w ankiecie udzieliło 66 kobiet, co stanowiło 57,39% ogółu badanych oraz mężczyzn 49 (42,61%). Badani przedstawiciele pokolenia Z najczęściej odbywali podróż w celach turystycznych (min. 1 nocleg) raz w roku lub rzadziej (51,30%). 41,74% studentów podróżowało od 2 do 3 razy w ciągu roku, a niewielki

odsetek respondentów (6,96%) deklarowali 4 i więcej podróży w ciągu roku. Respondenci również wskazali mając możliwość wybrania kilku odpowiedzi jednocześnie, skąd najczęściej pochodzą środki na ich wyjazdy. Najczęściej wskazywano na środki pieniężne od rodziców lub rodziny (39,13%), następnie ze stypendium uczelnianego (35,33%) oraz z oszczędności z pracy stałej lub dorywczej (25%). Na pochodzenie środków pieniężnych na wyjazdy z kredytu, karty kredytowej lub debetu wskazało 0,54% ankietowanych. Istotnym elementem badania była deklaracja o wysokości wydatków jakie średnio przeznaczają studenci na jeden wyjazd. Najczęściej wskazywaną odpowiedzią był koszt poniżej 1000 zł, zaznaczyło go 67 respondentów, co stanowiło 58,26% ogółu. Drugim najczęściej wybieranym przedziałem był 1000-2000 zł i zaznaczyło go 37,39%. Niewielki odsetek badanych wskazał przedziały 2001-4000 zł (2,61%) i powyżej 4000 zł (1,74%).

Uwzględniając odpowiedzi ankietowanych można scharakteryzować grupę badawczą jako studentów odbywających najczęściej od 1 do 3 podróży w trakcie roku i przeznaczających maksymalnie 2000 zł na podróże, z czego uzyskiwali te środki od rodziców, bądź z stypendium uczelnianego.

5.1. Wybierane modele wyjazdów

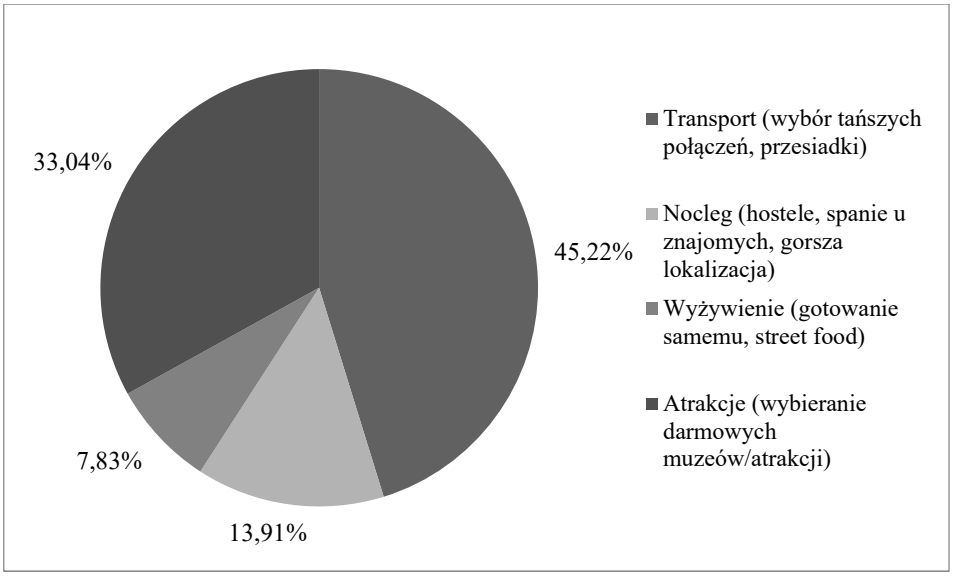
Respondenci pytani o model wyjazdów jaki najczęściej wybierają, wskazali na budżetowy (78,26%); tanie linie, hostele, samodzielne planowanie, 20% wybrało model standardowy (biura podróży, hotele 3*,gotowe pakiety), natomiast – 1,74% respondentów wersję premium; hotele 4/5*, pełen komfort (rys. 1.).



Rysunek 1. Najczęściej wybierany model wyjazdu
Figure 1. The most frequently chosen trip model
Źródło: Opracowanie własne.

5.2. Generowanie największych oszczędności w podróży turystycznej

Na rysunku 2. przedstawiono najchętniej wybierane filary stanowiące źródło największych oszczędności w podróży turystycznej. Studenci najwięcej oszczędzają na transporcie (wybór tańszych połączeń, przesiadki) – 45,22%, a także na atrakcjach (wybieranie darmowych muzeów/atrakcji) – 33,04%. Noclegi (hostele, spanie u znajomych, gorsza lokalizacja) wskazało 13,91% respondentów, natomiast wyżywienie (samodzielne przygotowanie posiłków, street food) – 7,83%.



Rysunek 2. Filar podróży turystycznej generujący największe oszczędności
Figure 2. The pillar of tourist travel that generates the greatest savings
Źródło: Opracowanie własne.

Tabela 1. Filar podróży turystycznej generujący największe oszczędności według płci

Table 1. The pillar of tourism travel that generates the greatest savings by gender

Filar podróży turystycznej generujący największe oszczędności	Kobieta	Mężczyzna
Atrakcje (wybieranie darmowych muzeów/atrakcji)	36,36%	28,57%
Nocleg (hostele, spanie u znajomych, gorsza lokalizacja)	10,61%	18,37%
Transport (wybór tańszych połączeń, przesiadki)	43,94%	46,94%
Wyżywienie (gotowanie samemu, street food)	9,09%	6,12%
Łącznie	100,00%	100,00%

Źródło: Opracowanie własne.

W tabeli 1. ukazano preferencje dotyczące filaru podróży turystycznej generującej największe oszczędności w zależności od płci respondentów. Zarówno

kobiety jak i mężczyźni najczęściej wybierali transport (wybór tańszych połączeń, przesiadki) – 43,94% i 46,94% oraz atrakcje (wybieranie darmowych muzeów/atrakcji) – 36,36% i 28,57%.

W tabeli 2. ukazano preferencje dotyczące filaru podróży turystycznej generującej największe oszczędności w zależności od częstości podróżowania. Najwięcej na transporcie (wybór tańszych połączeń, przesiadki) oszczędzało podróżujących 4 razy w roku i częściej oraz od 2 do 3 razy w roku – 75% i 58,33%, natomiast podróżujący raz w roku i rzadziej najbardziej optymalizowali koszty na atrakcjach (wybieranie darmowych muzeów/atrakcji) – 45,76%. Respondenci podróżujący od 2 do 3 razy w roku oraz raz w roku lub rzadziej optymalizowali koszty na wyżywieniu – 6,25% i 10,17%, natomiast podróżujący 4 razy w roku i częściej wcale na tym nie oszczędzali.

Tabela 2. Filar podróży turystycznej generujący największe oszczędności według częstości podróżowania

Table 2. Pillar of tourist travel generating the greatest savings by travel frequency

Filar podróży turystycznej generujący największe oszczędności	**2-3 razy w roku**	**4 razy w roku i częściej**	**Raz w roku lub rzadziej**
Atrakcje (wybieranie darmowych muzeów/atrakcji)	20,83%	12,50%	45,76%
Nocleg (hostele, spanie u znajomych, gorsza lokalizacja)	14,58%	12,50%	13,56%
Transport (wybór tańszych połączeń, przesiadki)	58,33%	75,00%	30,51%
Wyżywienie (gotowanie samemu, street food)	6,25%	0,00%	10,17%
Łącznie	100,00%	100,00%	100,00%

Źródło: Opracowanie własne.

W tabeli 3. ukazano preferencje dotyczące filaru podróży turystycznej generującej największe oszczędności w zależności od pochodzenia środków pieniężnych na wyjazdy. Respondenci, których środki pieniężne na podróże pochodziły z kredytu, karty kredytowej lub debetu oszczędzali tylko i wyłącznie na atrakcjach (wybieranie darmowych muzeów lub atrakcji) – 100%, natomiast w pozostałych grupach odpowiedzi były zróżnicowane. Zarówno ankietowani, których środki pochodziły oszczędności z pracy stałej lub dorywczej, stypendium uczelnianego oraz były od rodziców lub rodziny, najbardziej minimalizowali koszty na transporcie (wybór tańszych połączeń, przesiadki) – 52,94%, 45,33% i 38,89%.

Tabela 3. Filar podróży turystycznej generujący największe oszczędności według pochodzenia środków pieniężnych na wyjazdy

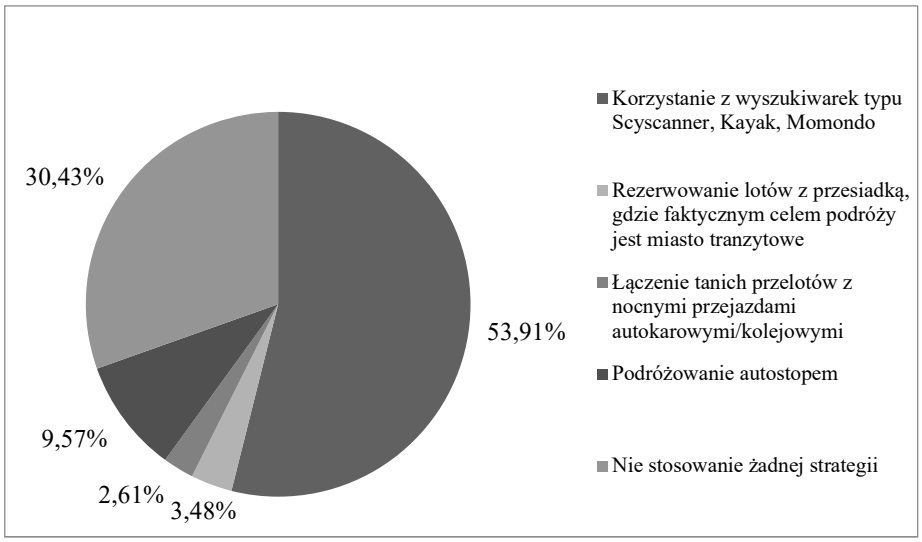
Table 3. Pillar of tourist travel generating the greatest savings according to the origin of funds for trips

Filar podróży turystycznej generujący największe oszczędności	Oszczędności z pracy (stałej lub dorywczej)	Stypendium uczelniane	Środki pieniężne z kredytu / karty kredytowej / debetu	Środki pieniężne od rodziców / rodziny
Atrakcje (wybieranie darmowych muzeów/atrakcji)	29,41%	40,00%	100,00%	36,11%
Nocleg (hostele, spanie u znajomych, gorsza lokalizacja)	8,82%	10,67%	0,00%	15,28%
Transport (wybór tańszych połączeń, przesiadki)	52,94%	45,33%	0,00%	38,89%
Wyżywienie (gotowanie samemu, street food)	8,82%	4,00%	0,00%	9,72%
Łącznie	100%	100%	100%	100%

Źródło: Opracowanie własne.

5.3. Strategie minimalizacji kosztów

Zbadano, jakie strategie minimalizacji kosztów są najczęściej wybierane w konkretnych obszarach. Na rysunku 3. przedstawiono najczęstsze metody optymalizacji kosztów transportu.



Rysunek 3. Stosowane strategie minimalizacji kosztów transportu

Figure 3. Strategies used to minimize transportation costs

Źródło: Opracowanie własne.

Wśród badanych studentów 53,91% zadeklarowało, że korzysta z wyszukiwarek typu Scyscanner, Kayak, Momondo. Żadnej ze strategii nie stosowało 30,43% respondentów, a podróżowało autostopem 9,57%. Badanych rezerwujących loty z przesiadką, gdzie faktycznym celem podróży jest miasto tranzytowe było 3,48%, natomiast łączących tanie przeloty z nocnymi przejazdami autokarowymi/kolejowymi 2,61%.

W tabeli 4. przedstawiono rozkład wskazań respondentów według płci w aspekcie stosowanej strategii minimalizacji kosztów transportu. Zarówno kobiety jak i mężczyźni stawiali na korzystanie z wyszukiwarek typu Scyscanner, Kayak, Momondo (56,06% i 51,02%). Żadnej strategii nie stosowało 34,85% kobiet, a mężczyzn 24,49%. W przypadku mężczyzn aż 14,29% szukała oszczędności, podróżując autostopem, natomiast w przypadku kobiet odsetek ten wynosił raptem 6,06% czyli o ponad 8 punktów procentowych mniej.

Tabela 4. Stosowane strategie minimalizacji kosztów transportu według płci

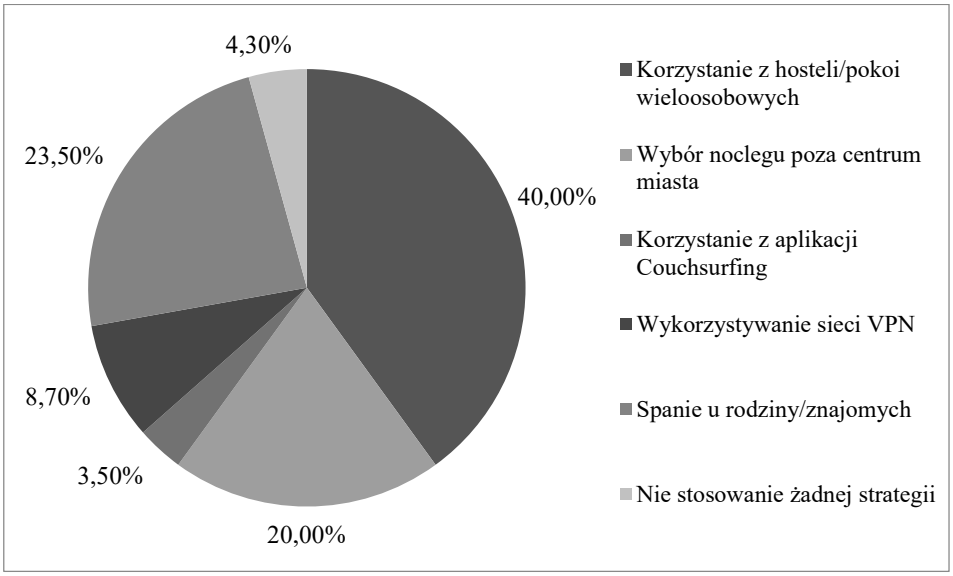
Table 4. Strategies used to minimize transportation costs by gender

Strategia minimalizacji kosztów transportu	**Kobieta**	**Mężczyzna**
Korzystanie z wyszukiwarek typu Scyscanner, Kayak, Momondo	56,06%	51,02%
Łączenie tanich lotów z nocnymi przejazdami autokarowymi/kolejowymi	1,52%	4,08%
Nie stosowanie żadnej strategii	34,85%	24,49%
Podróżowanie autostopem	6,06%	14,29%
Rezerwowanie lotów z przesiadką, gdzie faktycznym celem podróży jest miasto tranzytowe	1,52%	6,12%
Łącznie	100,00%	100,00%

Źródło: Opracowanie własne.

Z kolei na rysunku 4. przedstawiono wybierane metody optymalizacji kosztów zakwaterowania. Najczęściej wskazywano na wybór noclegów w hostelach lub pokojach wieloosobowych (40%), następnie na noclegów u rodziny lub znajomych (23,50%) oraz poza centrum miasta (20%). Inne formy optymalizacyjne – sieci VPN wybrało 8,70% respondentów, zaś korzystanie z aplikacji Couchsurfing – 3,50%. Żadnej ze strategii nie wskazało 3,50% badanych.

W tabeli 5. zaprezentowano rozkład wskazań respondentów według płci w aspekcie stosowanej strategii minimalizacji kosztów zakwaterowania. W przypadku kobiet aż 46,97% korzystało z hosteli lub pokoi wieloosobowych, natomiast w przypadku mężczyzn odsetek ten wynosił 30,61%, czyli o ponad 16 punktów procentowych mniej. Znaczącą dysproporcję w odpowiedziach stanowiło również zakwaterowanie u rodziny lub znajomych (25,76% kobiet i 20,41% mężczyzn), jak i poza centrum miasta (16,67% i 24,49%). Popularność do minimalizacji kosztów wśród mężczyzn zdobyło wykorzystywanie sieci VPN (20,41%), natomiast żadna z kobiet nie preferowała takiej strategii (0%). Żadnej ze strategii nie wskazało 7,58% kobiet, a mężczyzn 0%.



Rysunek 4. Stosowane strategie minimalizacji kosztów zakwaterowania
Figure 4. Strategies used to minimize accommodation costs
Źródło: Opracowanie własne.

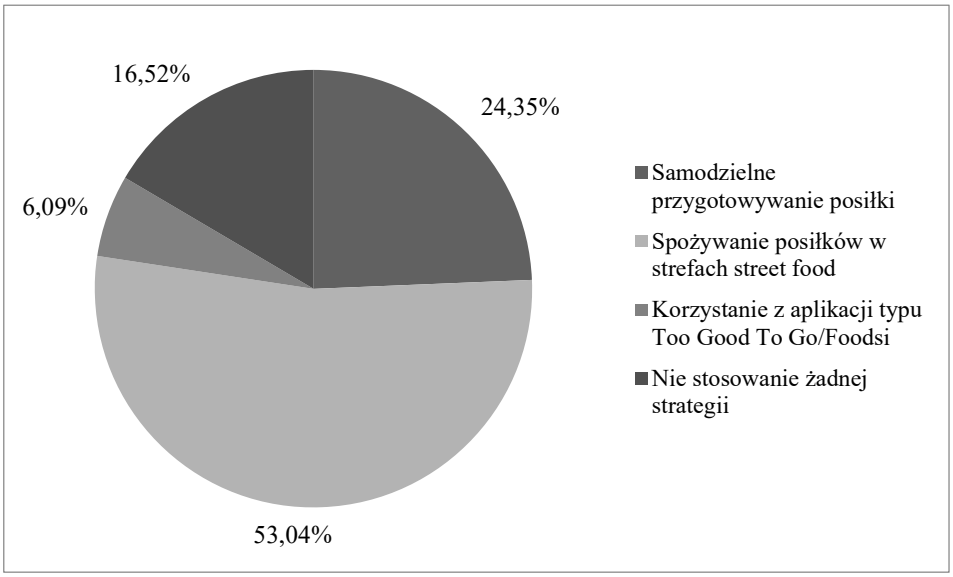
Tabela 5. Stosowane strategie minimalizacji kosztów zakwaterowania według płci

Table 5. Strategies used to minimize accommodation costs by gender

Strategia minimalizacji kosztów zakwaterowania	**Kobieta**	**Mężczyzna**
Korzystanie z aplikacji Couchsurfing	3,03%	4,08%
Korzystanie z hosteli/pokoi wieloosobowych	46,97%	30,61%
Nie stosowanie żadnej strategii	7,58%	0,00%
Spanie u rodziny/znajomych	25,76%	20,41%
Wybieranie noclegu poza centrum miasta	16,67%	24,49%
Wykorzystywanie sieci VPN	0,00%	20,41%
Łącznie	100,00%	100,00%

Źródło: Opracowanie własne.

Na rysunku 5. zaprezentowano wybierane strategie optymalizacji kosztów związanych z wyżywieniem w podróży przez studentów. Ponad połowa (53,04%) respondentów spożywało posiłki w strefach street food, samodzielnie przygotowywało posiłki – 24,35%. 16,52% wskazało, że nie stosowało żadnej z dostępnych w odpowiedziach strategii, natomiast 6,09% respondentów używało aplikacji typu Too Good To Go lub Foodsi.



Rysunek 5. Strategie minimalizacji kosztów wyżywienia
Figure 5. Strategies for minimizing food costs
Źródło: Opracowanie własne.

W tabeli 6. przedstawiono rozkład wskazań respondentów według płci w aspekcie stosowanej strategii minimalizacji kosztów wyżywienia. Zarówno kobiety jak i mężczyźni stawiali na spożywanie posiłków w strefach street food – 53,03% i 53,06%. Na samodzielnie przygotowywanie posiłków wskazało 27,27% kobiet, a mężczyzn 20,41%, czyli prawie o 7 punktów procentowych mniej. W przypadku kobiet tylko 12,12% nie stosowało żadnej strategii, natomiast w przypadku mężczyzn odsetek ten wynosił 22,45%.

Tabela 6. Stosowane strategie minimalizacji kosztów według płci

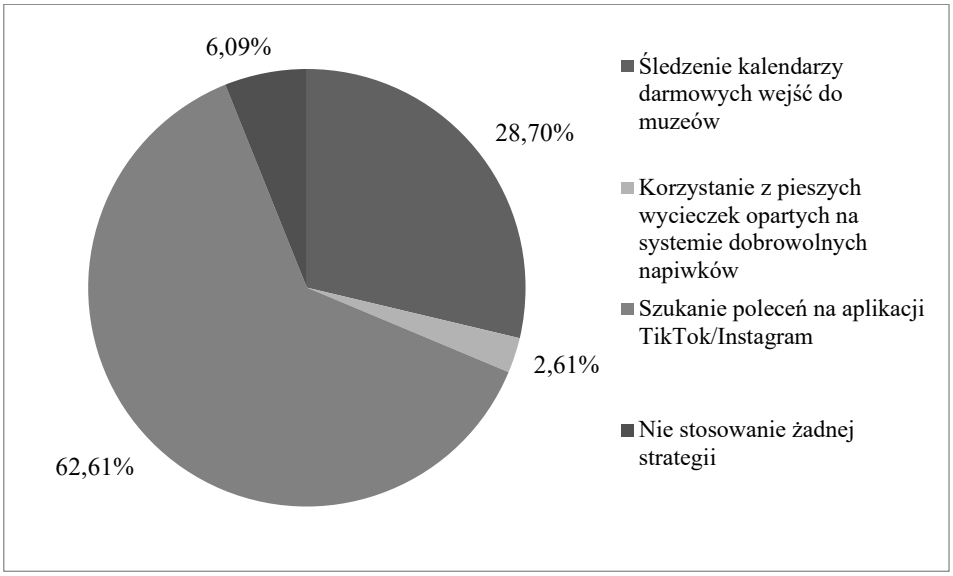
Table 6. Cost minimization strategies used by gender

Strategia minimalizacji kosztów wyżywienia	**Kobieta**	**Mężczyzna**
Spożywanie posiłków w strefach street food	53,03%	53,06%
Korzystanie z aplikacji typu Too Good To Go/Foodsi	7,58%	4,08%
Nie stosowanie żadnej strategii	12,12%	22,45%
Samodzielnie przygotowywanie posiłków	27,27%	20,41%
Łącznie	100,00%	100,00%

Źródło: Opracowanie własne.

Na rysunku 6. przedstawiono wybór metod optymalizacji związanych z obniżaniem kosztów atrakcji turystycznych. Najchętniej studenci szukali poleceń miejsc na aplikacji TikTok, jak również Instagram – 62,61% respondentów,

28,70% śledziła kalendarze darmowych wejść do muzeów, żadnych koncepcji nie posiadało 6,09% badanych, natomiast 2,61% korzystało z pieszych wycieczek opartych na systemie dobrowolnych napiwków.



Rysunek 6. Strategie minimalizacji kosztów atrakcji turystycznych
Figure 6. Strategies for minimizing the costs of tourist attractions
Źródło: Opracowanie własne.

W tabeli 7. przedstawiono rozkład wskazań respondentów według płci w aspekcie stosowanej strategii minimalizacji kosztów atrakcji turystycznych. Kobiety i mężczyźni preferowali szukanie poleceń na aplikacji TikTok/Instagram – 66,67% i 57,14% oraz śledzenie kalendarzy darmowych wejść do muzeów – 31,82% i 24,49%. W przypadku mężczyzn aż 14,29% nie stosowało żadnej strategii, natomiast w przypadku kobiet odsetek ten wynosił 0%.

Tabela 7. Stosowane strategie minimalizacji kosztów atrakcji turystycznych według płci
Table 7. Strategies used to minimize the costs of tourist attractions by gender

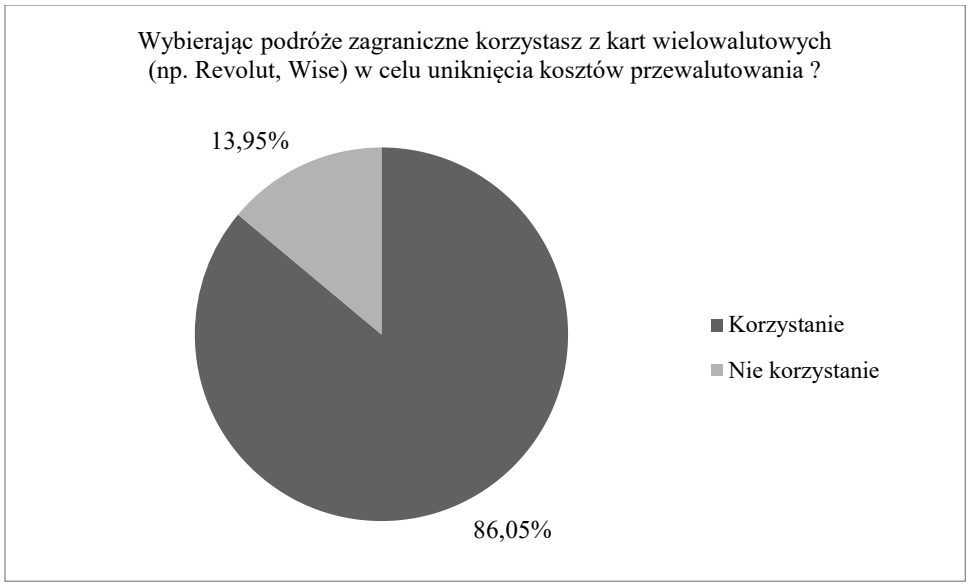
Strategia minimalizacji kosztów atrakcji turystycznych	**Kobieta**	**Mężczyzna**
Korzystanie z pieszych wycieczek opartych na systemie dobrowolnych napiwków	1,52%	4,08%
Nie stosowanie żadnej strategii	0,00%	14,29%
Szukanie poleceń na aplikacji TikTok/Instagram	66,67%	57,14%
Śledzenie kalendarzy darmowych wejść do muzeów	31,82%	24,49%
Łącznie	100,00%	100,00%

Źródło: Opracowanie własne.

5.4. Dodatkowe strategie minimalizacji kosztów

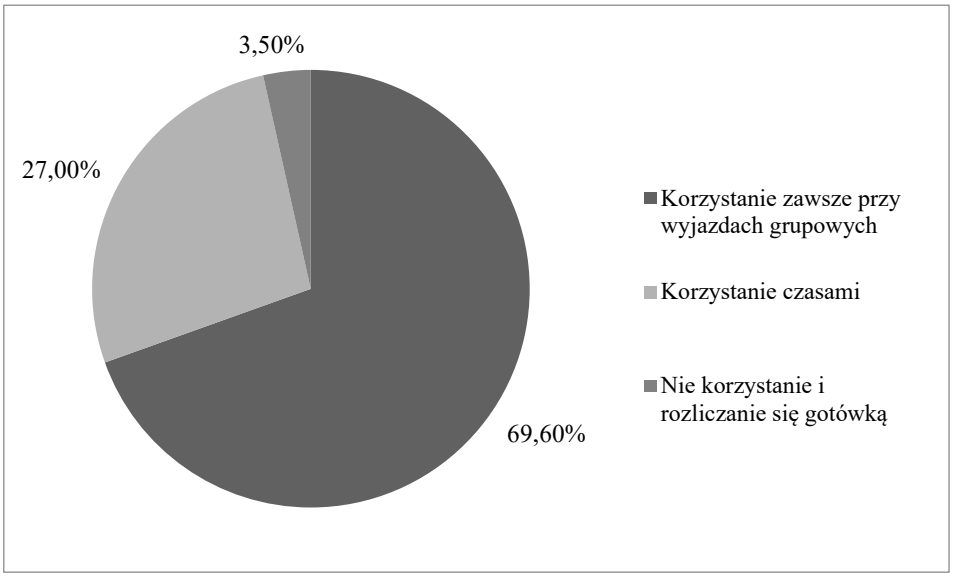
W badaniu 67,83% respondentów korzystało z podróży zagranicznych, natomiast 32,17% nie korzystało.

Na rysunku 7. przedstawiono skalę korzystania z kart wielowalutowych (np. Revolut, Wise). 86,05% ankietowanych korzystało z kart wielowalutowych, a nie korzystało zaledwie 13,95%.



Rysunek 7. Skala korzystania kart wielowalutowych (np. Revolut, Wise)
Figure 7. The scale of use of multi-currency cards (e.g. Revolut, Wise)
Źródło: Opracowanie własne.

Z kolei na rysunku 8. przedstawiono skalę korzystania z aplikacji do wspólnego rozliczania kosztów ze znajomymi (np. Splitwise, Tricount, przelewy na telefon BLIK). Wśród badanych studentów 62,61% zadeklarowało, że korzystało zawsze przy wyjazdach grupowych, a czasami korzystało 27%. Natomiast nie korzystało z aplikacji do wspólnego rozliczania kosztów tylko 3,50% respondentów.



Rysunek 8. Skala korzystania z aplikacji do wspólnego rozliczania kosztów ze znajomymi (np. Splitwise, Tricount, przelewy na telefon BLIK)
Figure 8. The scale of use of apps for joint settlement of costs with friends (e.g. Splitwise, Tricount, BLIK phone transfers)
Źródło: Opracowanie własne.

6. Podsumowanie

Analiza przeprowadzona na grupie 115 respondentów potwierdziła postawione założenia:

- Zdecydowana większość studentów (78,26% – 90 osób) wybiera budżetowy model podróżowania, oparty na tanich liniach lotniczych, hostelach i samodzielnym planowaniu.
- Studenci najchętniej optymalizowali wydatki związane z transportem (45,22%) oraz atrakcjami (33,04%).
- W strategiach optymalizacji kosztów transportu dominowały wyszukiwarki lotów typu Skyscanner lub Kayak (75% wskazań w tej kategorii).
- W zakwaterowaniu najczęstszą strategią oszczędnościową dla osób stawiających na ten filar był nocleg u rodziny lub znajomych (62,5%).
- W sferze kultury aż 68,42% osób oszczędzających na atrakcjach szukało inspiracji na TikToku i Instagramie.
- Pokolenie Z masowo korzystało z nowoczesnych technologii finansowych – 92,31% podróżujących za granicę używa kart wielowalutowych (np. Revolut), a blisko 70% zawsze stosuje aplikacje do rozliczeń grupowych, takich jak Splitwise czy BLIK.

- Typowy wyjazd studenta kosztował poniżej 1000 zł (58,26%), a środki na ten cel pochodziły głównie od rodziców lub ze stypendiów.
- Kobiety częściej wybierały hostele i pokoje wieloosobowe (46,97%) niż mężczyźni (30,61%), co sugerowało wyższą skłonność kobiet do kompromisu w zakresie prywatności na rzecz optymalizacji kosztów noclegu.
- Teoretyczne strategie sugerowały samodzielne przygotowywanie posiłków jako bazę oszczędności, natomiast badania wykazały dominację modelu street food (53,04%). Sugeruje to, że dla pokolenia Z autentyczność doświadczeń kulinarnych była ważniejsza niż maksymalna redukcja kosztów poprzez gotowanie.
- Narzędzia takie jak VPN pozwalały studentom (głównie mężczyznom – 20,41%) na obchodzenie barier geograficznych w cennikach, co było formą cyfrowego sprytu w walce o niższe stawki rezerwacyjne.
- 78,26% badanych studentów rezygnowało z usług biur podróży na rzecz samodzielnego planowania (tanie linie, hostele). Świadczyło to o wysokim poczuciu sprawstwa i kompetencjach logistycznych młodych podróżników.

Na podstawie analizy można wyznaczyć konkretne rekomendacje:

- Firmy powinny tworzyć pakiety modułowe, które pozwalają na samodzielną rezerwację (np. tylko noclegu), zamiast oferować sztywne pakiety, których pokolenie Z unika.
- Kluczowe atrakcje i punkty gastronomiczne powinny koncentrować działania promocyjne na TikToku i Instagramie, prezentując się jako miejsca budżetowe i autentyczne.
- Muzea i instytucje powinny tworzyć łatwo dostępne, mobilne kalendarze darmowych wejść, trafiając w potrzeby 33% studentów optymalizujących ten obszar.
- Miasta powinny promować bazę noclegową w dzielnicach ościennych, podkreślając sprawność transportu publicznego, co wpisuje się w strategię "Value for Money" badanych osób.
- Skoro 35,33% środków na wyjazdy pochodzi ze stypendiów , programy te są kluczowym czynnikiem umożliwiającym studentom rozwój kompetencji interpersonalnych poprzez podróże.
- Warto promować szkolenia z zakresu bezpiecznego korzystania z technologii optymalizacyjnych (np. VPN), aby wyrównać szanse i biegłość technologiczną bez względu na płeć.

Literatura

[1] Barton R., Aibinu A.A., Oliveros R. J.: The Value for Money Concept in investment evaluation: Deconstructing its meaning for better decision-making; Project Management Journal, 50(2), 2019, s. 210-225.

[2] Kauf S.: Ekonomia współdzielenia (sharing economy) jako narzędzie kreowania smart city, Zeszyty Naukowe Politechniki Śląskiej, Seria: Organizacja i Zarządzanie z. 120, 2018, s. 141-151.

[3] Koerts D.: Amsterdam: Interpreting the city in “free” walking tours, Conference Proceedings 5th International Research Forum on Guided Tours, Roskilde University, Denmark, 2017, s. 112-123.

[4] Kozłowski L., Brzykca A., Rudnicki R.: Niskobudżetowa turystyka indywidualna jako forma aktywności turystycznej, Ekonomiczne Problemy Turystyki, 1 (33), 2016, s. 259-271.

[5] Liu Q.: Paying More for a Shorter Flight? - Hidden City Ticketing, Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, USA, s. 1-36.

[6] Niemczyk A.: Aplikacje mobilne jako determinanta zachowań turystycznych (na przykładzie Krakowa), Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 473, 2017, s. 370-380.

[7] Olszowy J., Fetela K., Balmas K.: Finansowe labirynty. Odkrywanie nowych ścieżek w świecie kapitału i innowacji, symbioza teorii i praktyki: innowacje, perspektywy w ekonomii i finansach, Łódź 2024, s.43-66.

[8] Wang Z., Ye Y.: Hidden-City Ticketing: the Cause and Impact, September 14, 2013, s. 1-31.

[9] Widz M.: Typologia pakietów turystycznych na podstawie kryteriów systemów rezerwacyjnych, Annales Universitatis Mariae Curie-Skłodowska Lublin – Polonia, Vol. LXXIX, Sectio B, 2024, s. 115-132.

Źródła internetowe

[10] B. Dziwak: Ranking wyszukiwarek tanich lotów: TOP 13 portali, które warto poznać, planując wyjazd, https://bartekwpodrozy.pl/ranking-wyszukiwarek-tanich-lotow-top-portali-ktore-warto-poznac-planujac-wyjazd/ (dostęp: 22.02.2026 r.).

[11] https://turystyka.rp.pl/autobusy/art43228261-flixbus-i-koleo-w-jednej-aplikacji-najwieksze-polaczenie-oferty-autobusowej-i-kolejowej-w-polsce (dostęp: 22.02.2026 r.).

[12] https://worldbusinessoutlook.com/how-to-save-money-when-shopping-online-with-a-vpn/ (dostęp: 23.02.2026 r.).

[13] K. Pietrzak, Czy podróże zmieniają człowieka? Fakty i Mity, https://studentwpodrozy.pl/czy-podroze-zmieniaja-czlowieka-fakty-i-mity/#:~:text=Psycho%2Dodporno%C5%9B%C4%87%20i%20rozw%C3%B3j%20osobisty%20Podr%C3%B3%C5%BCe%20stawiaj%C4%85%20przed,nowe%20umiej%C4%99tno%C5%9Bci%20rozwi%C4%85zywania%20problem%C3%B3w%2C%20rozwijamy%20pewno%C5%9B%C4%87%20siebie (dostęp: 20.02.2026 r.).

STRATEGIES FOR OPTIMIZING EXPENSES IN STUDENT TOURISM – ANALYSIS OF GENERATION Z ATTITUDES

Summary

This article analyzes the evolution of Generation Z's travel model, which views tourism as a key tool for self-fulfillment. The primary goal of the publication was to identify strategies used by students to overcome financial barriers while maintaining mobility. Empirical research

conducted on a sample of 115 individuals clearly confirmed the thesis, demonstrating the dominance of the budget model over standard travel agency offerings. The contemporary young traveler is presented in the study as a conscious "financial engineer" who is abandoning archaic, passive saving in favor of actively optimizing the value-for-money ratio. The foundation of this engineering is transportation, which is the primary area for generating savings. Optimization in this area is driven by constant monitoring of market fluctuations, the use of search engine algorithms, and radical flexibility in choosing travel dates. The analysis also demonstrated that this generation is massively harnessing modern digital technologies for commission-free capital management. The use of fintech instruments and applications for precise group cost accounting permanently eliminates errors in the fair distribution of expenses. A key component of the strategy remains the exploration of free cultural resources, located primarily through social media such as TikTok and Instagram. Instead of using commercial services, young people are relying on digital savvy, which allows them to extend the duration of their trips by converting savings into additional days in the field. The publication concludes that high technological proficiency allows students to manage limited capital extremely effectively. As a result, a tourist trip is no longer perceived as a cost, but instead becomes a precise, highly efficient logistical operation with an investment character.

Keywords: low-budget tourism, cost minimization, financial engineering, digital mobility.

Mikołaj Lech, Mirosław Sobaszek
Koło Naukowe Elektroników

Mirosław Sobaszek
opiekun Koła Naukowego Elektroników

MODEL IMPEDANCYJNY GŁOŚNIKA DYNAMICZNEGO W OBUDOWIE ZAMKNIĘTEJ

Streszczenie

Celem pracy było opracowanie, implementacja oraz eksperymentalna weryfikacja modelu impedancji głośnika dynamicznego pracującego w obudowie zamkniętej. Analizowany model oparto na parametrach Thiele'a-Smalla oraz elektromechanicznym schemacie zastępczym typu RLC, umożliwiającym opis zespolonej impedancji w funkcji częstotliwości. W modelu uwzględniono zarówno właściwości elektryczne cewki drgającej, jak i rezonansowy charakter części mechanicznej oraz wpływ objętości i tłumienia obudowy na zachowanie całego układu. W części teoretycznej przedstawiono zależności opisujące wpływ rezystancji i indukcyjności cewki, parametrów układu drgającego oraz sprężystości powietrza zamkniętego w obudowie na przebieg charakterystyki impedancyjnej. Model rozszerzono o dodatkowy element reprezentujący straty energii występujące w objętości powietrza i materiałach tłumiących, co umożliwiło dokładniejsze odwzorowanie obniżenia i poszerzenia piku rezonansowego obserwowanego w rzeczywistych układach. Na podstawie opracowanego modelu wykonano symulacje numeryczne w środowisku MATLAB, wyznaczając charakterystyki modułu i fazy impedancji w szerokim zakresie częstotliwości. Następnie przeprowadzono pomiary eksperymentalne z wykorzystaniem oscyloskopu Keysight MSOX4024A wyposażonego w funkcję analizy odpowiedzi częstotliwościowej (FRA). Pomiar realizowano metodą jednoczesnej rejestracji napięcia i prądu przy użyciu sondy napięciowej oraz sondy prądowej Keysight 1147B. Uzyskane wyniki potwierdziły charakterystyczne zachowanie głośnika w obudowie zamkniętej – występowanie pojedynczego maksimum impedancji w pobliżu częstotliwości rezonansowej, minimalnej wartości impedancji zbliżonej do rezystancji cewki przy niskich częstotliwościach oraz wzrostu impedancji w zakresie wyższych częstotliwości wynikającego z indukcyjności cewki. Porównanie charakterystyk symulacyjnych i pomiarowych wykazało dobrą zgodność jakościową pod względem kształtu przebiegów oraz położenia maksimum rezonansowego. Na podstawie przeprowadzonej analizy stwierdzono, że opracowany model poprawnie opisuje rzeczywiste właściwości elektromechaniczne głośnika i może stanowić użyteczne narzędzie wspoma-

gające projektowanie obudów głośnikowych oraz analizę współpracy głośnika ze zwrotnicami i wzmacniaczem. Zaobserwowane różnice pomiędzy symulacją i pomiarem wynikały głównie z zastosowania parametrów katalogowych zamiast parametrów wyznaczonych eksperymentalnie oraz z niepewności pomiarowych i wpływu warunków pracy. Wykazano również, że uwzględnienie tłumienia obudowy istotnie zwiększa zgodność modelu z wynikami pomiarów.

Słowa kluczowe: parametry Thiele'a-Smalla, MATLAB, analiza częstotliwościowa, układ rezonansowy, symulacja.

1. Wprowadzenie

Głośnik dynamiczny jest jednym z podstawowych elementów toru elektroakustycznego, a jego właściwości elektryczne, mechaniczne i akustyczne mają bezpośredni wpływ na jakość reprodukowanego dźwięku. Jednym z kluczowych parametrów charakteryzujących głośnik jest jego impedancja, która w przeciwieństwie do idealnego obciążenia rezystancyjnego ma charakter zespolony i silnie zależny od częstotliwości. Zależność ta wynika zarówno z własności elektrycznych cewki drgającej, jak i z rezonansowego charakteru układu mechanicznego membrany, zawieszenia oraz współpracującego z nimi powietrza.

W praktyce projektowej, szczególnie podczas projektowania pasywnych zwrotnic głośnikowych, często stosuje się uproszczone założenie stałej impedancji znamionowej głośnika. Prowadzi to jednak do rozbieżności pomiędzy charakterystykami teoretycznymi a rzeczywistym zachowaniem układu, zwłaszcza w pobliżu częstotliwości rezonansowej oraz w zakresie wyższych częstotliwości, gdzie istotny wpływ ma indukcyjność cewki drgającej. Dokładniejszy opis zachowania głośnika umożliwiają parametry Thiele'a-Smalla oraz elektromechaniczne modele zastępcze, które pozwalają na uwzględnienie wpływu obudowy oraz strat mechanicznych i elektrycznych. W pracy wykorzystano symulacje numeryczne wykonane w środowisku MATLAB oraz pomiary przeprowadzone z użyciem oscyloskopu wyposażonego w funkcję analizy odpowiedzi częstotliwościowej FRA.

2. Model impedancji głośnika w obudowie zamkniętej

Impedancja głośnika dynamicznego pracującego w obudowie zamkniętej ma charakter wyraźnie rezonansowy i może być opisana za pomocą elektromechanicznego schematu zastępczego. Model ten łączy w sobie elementy elektryczne, mechaniczne i akustyczne, umożliwiając analizę zachowania przetwornika w funkcji częstotliwości. Dzięki temu możliwe jest przewidywanie zarówno przebiegu impedancji, jak i wpływu objętości obudowy oraz strat energetycznych na właściwości układu.

Od strony elektrycznej głośnik reprezentowany jest przez rezystancję cewki drgającej R_e oraz jej indukcyjność L_e. Rezystancja R_e odpowiada rzeczywistej rezystancji uzwojenia i wyznacza minimalną wartość impedancji w zakresie bardzo niskich częstotliwości, gdzie wpływ elementów mechanicznych jest niewielki. Indukcyjność L_e opisuje zjawiska związane z polem magnetycznym cewki i powoduje wzrost impedancji wraz ze wzrostem częstotliwości. W zakresie wysokotonowym impedancja głośnika ma zatem charakter rosnący, zbliżony do liniowej funkcji częstotliwości.

Część mechaniczna modelu odwzorowuje ruch układu drgającego, obejmującego membranę, cewkę, resor dolny, zawieszenie górne oraz powietrze poruszane przez membranę. Masa drgająca reprezentowana jest przez indukcyjność L, której wartość odpowiada bezwładności układu ruchomego. Podatność mechaniczna zawieszenia modelowana jest przez pojemność C – im większa podatność (miększe zawieszenie), tym większa wartość pojemności w modelu. Straty energii występujące w układzie mechanicznym, wynikające z tarcia w zawieszeniu, strat materiałowych oraz oporów powietrza, reprezentowane są przez rezystancję R. Elementy L, C i R tworzą równoległy obwód rezonansowy, odpowiedzialny za powstawanie maksimum impedancji przy częstotliwości rezonansowej.

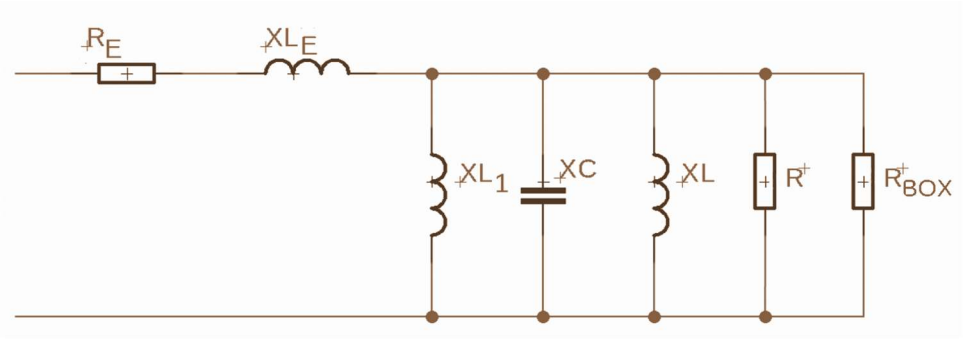
Umieszczenie głośnika w obudowie zamkniętej wprowadza dodatkowy element sprężysty – powietrze uwięzione wewnątrz obudowy. Powietrze to ulega sprężaniu i rozprężaniu podczas ruchu membrany, zachowując się jak dodatkowa sprężyna mechaniczna. W modelu zastępczym zjawisko to reprezentowane jest przez dodatkową indukcyjność L_1 dołączoną równolegle do indukcyjności L. Skutkiem tego jest zmniejszenie całkowitej podatności układu drgającego, co prowadzi do zwiększenia częstotliwości rezonansowej z wartości F_s (rezonans w wolnym powietrzu) do wartości F_c (rezonans w obudowie). Zmiana ta zależy bezpośrednio od stosunku objętości równoważnej Vas do objętości obudowy V_b. Mniejsza objętość obudowy oznacza większe usztywnienie, wyższą częstotliwość rezonansową i silniejszą ingerencję obudowy w pracę przetwornika.

W praktyce powietrze w obudowie nie jest elementem idealnie sprężystym. Występują w nim straty energii wynikające z lepkości powietrza, obecności materiałów tłumiących (wełna, pianka, włóknina), a także z ewentualnych nieszczelności konstrukcji. Aby uwzględnić te zjawiska w modelu matematycznym, wprowadza się dodatkową rezystancję R_{box}, dołączoną równolegle do elementów mechanicznych. Rezystancja ta modeluje tłumienie akustyczne obudowy i odpowiada za rozpraszanie energii w objętości powietrza. Jej obecność powoduje zmniejszenie dobroci rezonansu oraz obniżenie wysokości maksimum impedancji.

Po uwzględnieniu tłumienia obudowy część mechaniczna modelu ma postać równoległego połączenia czterech elementów: indukcyjności Lw (wypadkowej indukcyjności reprezentującej masę drgającą oraz sprężystość powietrza w obudowie), pojemności C, rezystancji mechanicznej R oraz rezystancji strat obudowy R_{box}. Schemat zastępczy analizowanego układu przedstawiono na rysunku 1. Całkowita dobroć mechaniczna układu maleje zgodnie z zależnością:

$$Q_{total} = \frac{Q_{ms} * Q_{box}}{Q_{ms} + Q_{box}} \quad (1)$$

gdzie: Q_{total} – łączna dobroć,
Q_{ms} – dobroć mechaniczna,
Q_{box} – tłumienie obudowy.

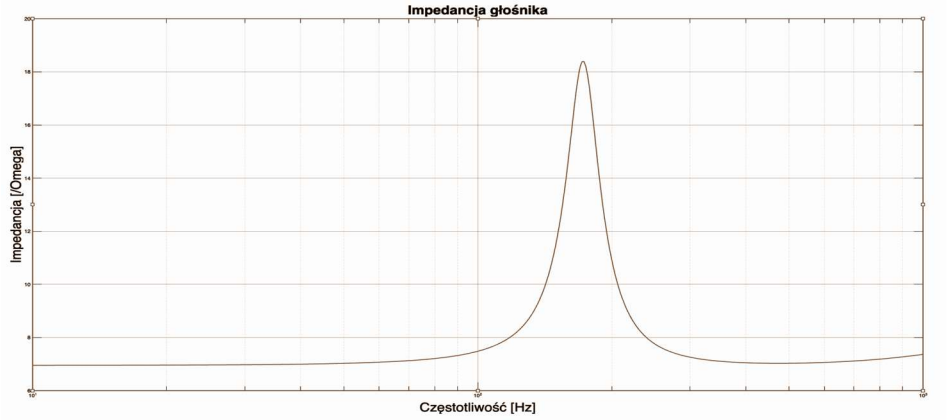


Rys. 1. Schemat zastępczy modelu impedancji głośnika w obudowie zamkniętej
Fig. 1. Equivalent circuit model of the impedance of a loudspeaker in a sealed enclosure

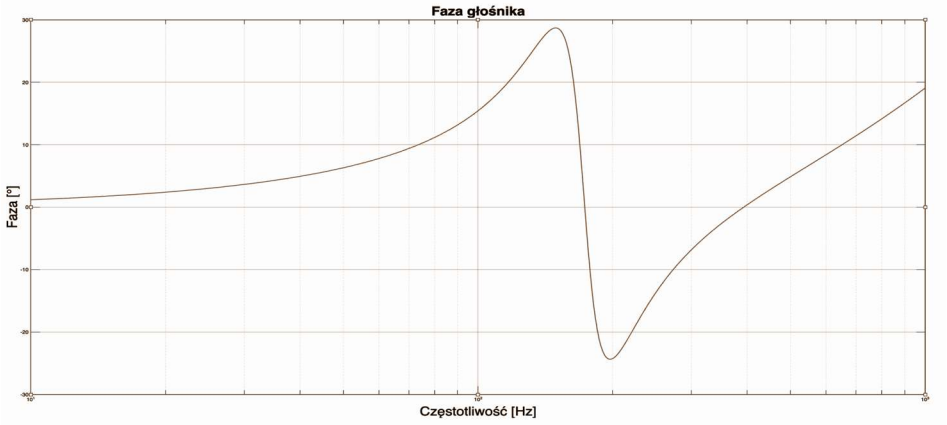
Oznacza to, że wzrost tłumienia w obudowie prowadzi do zmniejszenia amplitudy rezonansu oraz jego poszerzenia. W charakterystyce impedancji obserwuje się wówczas niższy i szerszy pik, co przekłada się na bardziej kontrolowaną odpowiedź impulsową i mniejszą skłonność do podbarwień w zakresie niskich częstotliwości.

Charakterystyka impedancji głośnika w obudowie zamkniętej ma zatem trzy wyraźne obszary. W zakresie bardzo niskich częstotliwości impedancja zbliża się do wartości Re. W pobliżu częstotliwości rezonansowej pojawia się pojedyncze maksimum, którego wysokość i szerokość zależą od parametrów Q_{ms}, Q_{es} oraz Q_{box}. W zakresie wysokich częstotliwości dominuje wpływ indukcyjności L_e, powodując systematyczny wzrost impedancji. Przebieg modułu impedancji i fazy przedstawiono na rysunku 2. i rysunku 3.

Zachowanie układu opisują parametry Thiele'a-Smalla, które stanowią podstawę projektowania obudów. Parametr Vas pozwala określić zależność między podatnością zawieszenia a objętością obudowy, natomiast dobroci Q_{ms}, Q_{es} i Q_{ts} określają poziom tłumienia. Po umieszczeniu głośnika w obudowie zamkniętej częstotliwość rezonansowa wzrasta do F_c, a dobroć przyjmuje wartość Q_{tc}, zależną od V_b oraz od tłumienia obudowy. Właściwy dobór objętości obudowy i poziomu tłumienia umożliwia kształtowanie charakteru basu – od odpowiedzi bardziej miękkiej i rozciągniętej (duża obudowa, mniejsze usztywnienie) do bardziej zwartej i kontrolowanej (mniejsza obudowa, większe tłumienie).



Rys. 2. Wykres przebiegu modułu impedancji głośnika
Fig. 2. Plot of the magnitude of the loudspeaker impedance



Rys. 3. Wykres przebiegu fazy impedancji
Fig. 3. Plot of the impedance phase response

Uwzględnienie rezystancji R_{box} w modelu zwiększa jego zgodność z rzeczywistymi pomiarami impedancji, ponieważ pozwala odtworzyć obserwowane obniżenie i poszerzenie piku rezonansowego. Dzięki temu model matematyczny wierniej odwzorowuje rzeczywiste zachowanie głośnika w obudowie zamkniętej i umożliwia dokładniejszą analizę oraz projektowanie układu elektroakustycznego.

Parametry elementów modelu:

$$X_L(s) = s * A * C_1$$

$$X_{L1}(s) = s * A * C_2$$

$$R = \frac{A}{R_1}$$

$$X_C(s) = \frac{s * M_1}{A}$$

$$R_{box}(s) = A * \omega * C_2 * Q_{box}$$

$$X_{Lw}(s) = \frac{X_{L1} * X_L}{X_{L1} + X_L}$$

gdzie:

$$A = \frac{R_E * 2{,}27 * 10^7}{f_s * V_{AS} * Q_{ES}}$$

$$M_1 = \frac{3{,}55 * 10^6}{{F_s}^2 * V_{AS}}$$

$$C_1 = V_{AS} * 0{,}7 * 10^{-8}$$

$$R_1 = \frac{F_S * M_1 * 6{,}28}{Q_{MS}}$$

$$C_2 = V_B * 0{,}7 * 10^{-8}$$

Impedancję zespoloną głośnika można zapisać w postaci:

$$Z(s) = R_e + sL_e + \frac{1}{(\frac{1}{sL\omega} + sC + \frac{1}{R} + \frac{1}{Rbox})} \quad (2)$$

Tabela 1. Parametry głośnika niskotonowego GDN 11/30 użytego do badań ze strony producenta

Table 1. Parameters of the GDN 11/30 woofer used for testing from the manufacturer's website

Impedancja	8 ohm
Rezystancja cewki	6,75 ohm
Pasmo przenoszenia	100 - 10000 Hz
Częstotliwość rezonansowa FS	55 Hz
Moc nominalna	30 W
Moc maksymalna	60 W
Efektywność	85 dB
Objętość ekwiwalentna VAS	0,006dm^3
Dobroć mechaniczna QMS	3,89
Dobroć elektryczna QES	0,58
Dobroć całkowita QTS	0,5
Tłumienie obudowy	0,3

3. Pomiar rzeczywistej impedancji głośnika i weryfikacja modelu

Pomiar impedancji głośnika z wykorzystaniem oscyloskopu Keysight wyposażonego w funkcję FRA (Frequency Response Analysis) polega na wyznaczeniu charakterystyki zespolonej impedancji w funkcji częstotliwości na podstawie jednoczesnego pomiaru napięcia na zaciskach głośnika oraz prądu płynącego przez jego cewkę. Do pomiaru wykorzystywany jest wbudowany w oscyloskop generator sygnału sinusoidalnego, który poprzez wzmacniacz audio pobudza głośnik sygnałem o niewielkiej amplitudzie, zapewniającej pracę wzmacniacza i głośnika w zakresie liniowym. Napięcie na głośniku mierzone jest za pomocą sondy napięciowej podłączonej równolegle do jego zacisków, natomiast prąd mierzony jest sondą prądową umieszczoną szeregowo z głośnikiem lub obejmującą jeden z przewodów zasilających. Sygnały z obu sond doprowadzane są do dwóch kanałów oscyloskopu, z których jeden pełni rolę sygnału odniesienia, a drugi sygnału odpowiedzi w analizie FRA. Stanowisko pomiarowe przedstawiono na rysunku 4. Składa się ono z oscyloskopu Keysight MSOX4024A, sondy prądowej Keysight 1147B, sondy napięciowej Keysight N2841A, wzmacniacza audio Onkyo TX-7820.

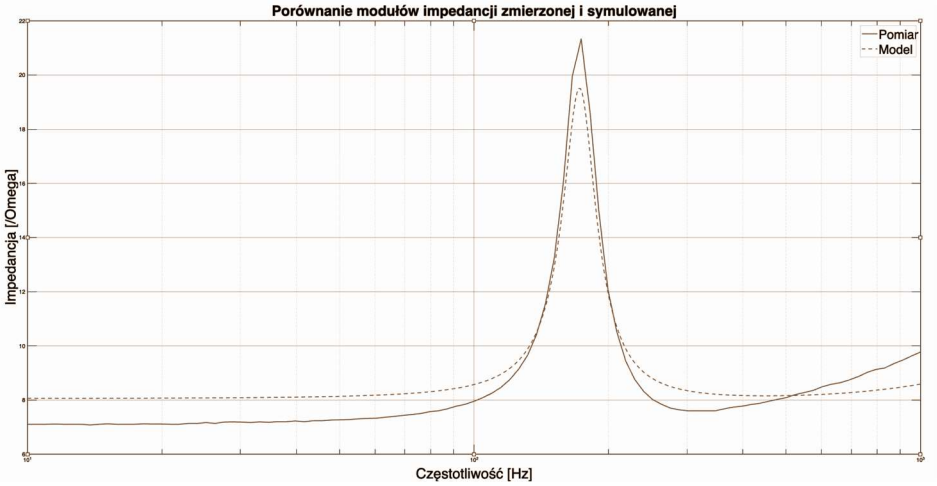


Rys. 4. Stanowisko pomiarowe

Fig. 4. The measurement station

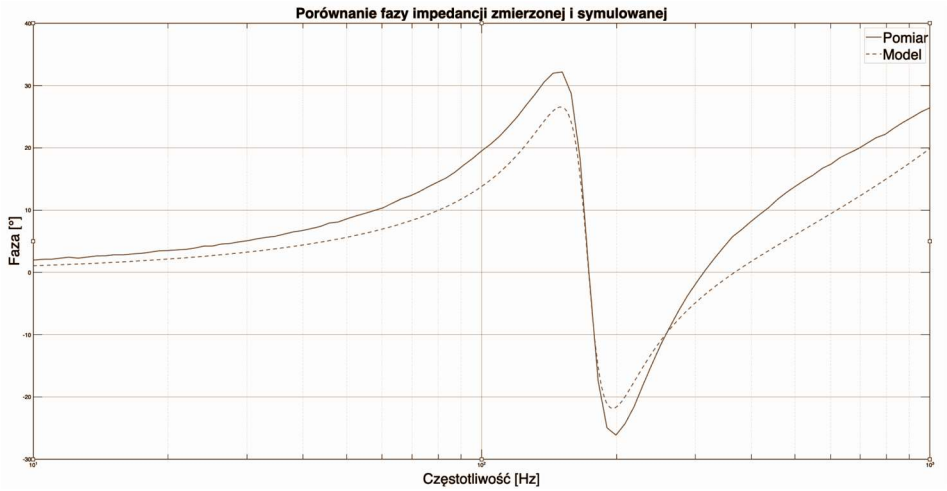
Funkcja FRA realizuje przemiatanie częstotliwości w zakresie od 10 Hz do 10 kHz, rejestrując dla każdej częstotliwości amplitudę oraz przesunięcie fazowe pomiędzy napięciem i prądem. Na tej podstawie oscyloskop oblicza moduł impedancji jako iloraz amplitudy napięcia i amplitudy prądu oraz kąt fazowy impedancji jako różnicę faz pomiędzy tymi wielkościami. Przed wykonaniem właściwego pomiaru konieczne jest przeprowadzenie kompensacji sondy napięciowej oraz wyzerowanie i skalibrowanie sondy prądowej, a także uwzględnienie jej współczynnika przeliczeniowego. Dla zwiększenia dokładności pomiaru często wykonuje się również pomiar referencyjny na znanym rezystorze, co pozwala zweryfikować poprawność toru pomiarowego i ustawień FRA.

Wynikiem pomiaru jest charakterystyka impedancji głośnika w funkcji częstotliwości, obejmująca zarówno zmianę wartości modułu, jak i kąta fazowego. Charakterystyka ta umożliwia identyfikację częstotliwości rezonansowej głośnika, obserwację minimum impedancji w zakresie niskich częstotliwości oraz wzrostu impedancji przy wyższych częstotliwościach, wynikającego z indukcyjnego charakteru cewki drgającej. Metoda oparta na analizie FRA i pomiarze napięcia oraz prądu jest szczególnie użyteczna, ponieważ pozwala na szybkie i dokładne wyznaczenie impedancji zespolonej bez konieczności stosowania dodatkowych mostków pomiarowych, a jednocześnie odzwierciedla rzeczywiste warunki pracy głośnika zasilanego ze wzmacniacza. Porównanie przebiegu modułu impedancji i fazy symulowanej oraz zmierzonej przedstawiono na rysunku 5. i rysunku 6.



Rys. 5. Porównanie modułu impedancji zmierzonego z modułem z modelu w MATLAB

Fig. 5. The comparison of the measured impedance magnitude with the magnitude obtained from the MATLAB model



Rys. 6. Porównanie fazy impedancji zmierzonego z modułem z modelu w MATLAB

Fig. 6. The comparison of the measured impedance phase with the phase obtained from the MATLAB model

Różnice w przebiegach wynikają z przyjęcia parametrów TS z noty katalogowej, a nie zmierzonych dla konkretnego głośnika, które zmieniają się pod wpływem temperatury, mocy dostarczonej do głośnika, mogą się one także różnić dla poszczególnych głośników z danej serii produkcyjnej.

4. Podsumowanie

Celem niniejszej pracy było opracowanie, implementacja oraz weryfikacja modelu impedancji głośnika dynamicznego pracującego w obudowie zamkniętej, z uwzględnieniem rzeczywistych zjawisk stratnych występujących w układzie mechanicznym i akustycznym. W ramach pracy przedstawiono elektromechaniczny model zastępczy oparty na parametrach Thiele'a-Smalla, który umożliwia opis zespolonej impedancji przetwornika w funkcji częstotliwości. Model rozszerzono o dodatkowy element reprezentujący tłumienie powietrza w obudowie, co pozwoliło na bardziej realistyczne odwzorowanie przebiegu charakterystyki impedancyjnej.

W części teoretycznej omówiono znaczenie poszczególnych parametrów modelu oraz ich interpretację fizyczną. Wykazano, że impedancja głośnika jest wypadkową oddziaływania części elektrycznej (rezystancja i indukcyjność cewki drgającej) oraz części mechaniczno-akustycznej, która ma charakter rezonansowy. Umieszczenie głośnika w obudowie zamkniętej powoduje zmianę warunków pracy przetwornika poprzez wprowadzenie dodatkowej sprężystości powietrza, co skutkuje wzrostem częstotliwości rezonansowej oraz zmianą dobroci układu.

Uwzględnienie rezystancji R_{box} modelującej straty w obudowie pozwoliło opisać zjawisko tłumienia rezonansu oraz obniżenia amplitudy maksimum impedancji.

Symulacje numeryczne wykonane w środowisku MATLAB umożliwiły wyznaczenie charakterystyk modułu i fazy impedancji w szerokim zakresie częstotliwości. Otrzymane przebiegi potwierdziły typowe właściwości głośnika w obudowie zamkniętej:

- obecność pojedynczego maksimum impedancji w pobliżu częstotliwości rezonansowej F_c,
- minimalną wartość impedancji zbliżoną do rezystancji cewki R_e w zakresie niskich częstotliwości,
- systematyczny wzrost impedancji w zakresie wyższych częstotliwości wynikający z indukcyjności L_e,
- zmianę kąta fazowego przechodzącą z charakteru pojemnościowego przez rezonans do charakteru indukcyjnego.

Wprowadzenie tłumienia obudowy wpłynęło na zmniejszenie wysokości piku rezonansowego oraz jego poszerzenie, co odpowiada zmniejszeniu dobroci całkowitej układu. Zjawisko to ma istotne znaczenie praktyczne, ponieważ wpływa na odpowiedź impulsową głośnika oraz subiektywnie odbierany charakter niskich częstotliwości. Większe tłumienie prowadzi do bardziej kontrolowanej i mniej podbarwionej reprodukcji basu, natomiast mniejsze tłumienie skutkuje wyraźniejszym wzmocnieniem w pobliżu częstotliwości rezonansowej.

W części eksperymentalnej przeprowadzono pomiar rzeczywistej impedancji głośnika z wykorzystaniem oscyloskopu wyposażonego w funkcję analizy odpowiedzi częstotliwościowej (FRA). Metoda ta umożliwiła bezpośrednie wyznaczenie modułu oraz kąta fazowego impedancji poprzez pomiar napięcia i prądu w funkcji częstotliwości. Uzyskane charakterystyki pomiarowe zostały następnie porównane z wynikami symulacji.

Porównanie wyników wykazało dobrą zgodność jakościową pomiędzy modelem a rzeczywistym zachowaniem głośnika. Kształt charakterystyki, położenie maksimum rezonansowego oraz ogólny trend zmian impedancji w funkcji częstotliwości zostały poprawnie odwzorowane przez model. Zaobserwowane różnice ilościowe dotyczące dokładnej wartości częstotliwości rezonansowej oraz wysokości piku impedancji wynikają przede wszystkim z wykorzystania parametrów katalogowych zamiast parametrów wyznaczonych eksperymentalnie dla konkretnego egzemplarza głośnika. Dodatkowo na wyniki pomiarów wpływają czynniki takie jak temperatura cewki drgającej, poziom sygnału pobudzającego, tolerancje produkcyjne oraz rzeczywiste właściwości materiałów tłumiących zastosowanych w obudowie.

Przeprowadzona analiza potwierdziła, że elektromechaniczny model RLC oparty na parametrach Thiele'a-Smalla stanowi skuteczne narzędzie do modelowania impedancji głośnika w obudowie zamkniętej. Uwzględnienie tłumienia obudowy zwiększa dokładność modelu oraz jego zgodność z pomiarami eksperymentalnymi. Model taki może być wykorzystywany w procesie projektowania

zestawów głośnikowych, doboru objętości obudowy, analizy współpracy z układami zwrotnic pasywnych oraz oceny obciążenia wzmacniacza audio.

Dokładność pomiaru napięcia za pomocą oscyloskopu Keysight MSOX4024A: ±0,1div±2mV±1% offsetu. Dokładność pomiaru prądu za pomocą sondy Keysight 1147B: ±(1%·I+10mA).

Literatura

1. Mrozek B., Mrozek Z.: *MATLAB i Simulink. Poradnik użytkownika*, Wydawnictwo Helion, Gliwice 2017.
2. Herres D.: *Oscilloscopes: A Manual for Students, Engineers, and Scientists*, Springer International Publishing, Cham 2020.
3. Witort A.: *Zestawy głośnikowe*, Wydawnictwo Czasopism i Książek Technicznych NOT-SIGMA, Warszawa 1986.
4. *Kolumny głośnikowe* [online]. Dostępne w Internecie: https://kolumny.prv.pl/ (dostęp: 20 stycznia 2026 r.).
5. *GDN 11/30/8 Ω* [online]. Dostępne w Internecie: https://www.skleptonsil.pl/gdn-11-30-8-ohm,id563.html (dostęp: 20 stycznia 2026 r.).
6. Krajewski J.: *Głośniki i zestawy głośnikowe. Budowa, działanie, zastosowania*, Wydawnictwa Komunikacji i Łączności WKŁ, Warszawa 2008.
7. Kulka Z.: *Elektroakustyka dla akustyków*, Wydawnictwo Naukowe PWN, Warszawa 2005.
8. Górecki K.: *Podstawy elektroakustyki*, Wydawnictwo BTC, Legionowo 2004.
9. Śliwiński A.: *Akustyka architektoniczna*, Wydawnictwo Naukowe PWN, Warszawa 2015.

MODELLING THE IMPEDANCE OF A DYNAMIC SPEAKER IN A SEALED ENCLOSURE

Summary

The aim of this study was to develop, implement, and experimentally verify an impedance model of a dynamic loudspeaker operating in a sealed enclosure. The analyzed model was based on Thiele–Small parameters and an electromechanical RLC equivalent circuit, enabling the description of the complex impedance as a function of frequency. The model takes into account both the electrical properties of the voice coil and the resonant behavior of the mechanical system, as well as the influence of enclosure volume and damping on the overall system performance.

The theoretical part presents the relationships describing the influence of voice coil resistance and inductance, moving system parameters, and the stiffness of the air enclosed in the cabinet on the impedance response. The model was extended by introducing an additional element representing energy losses occurring in the enclosed air volume and damping materials, allowing the reduction and broadening of the resonance peak observed in real systems to be reproduced more accurately.

Based on the developed model, numerical simulations were performed in MATLAB to determine the magnitude and phase characteristics of impedance over a wide frequency range. Experimental measurements were then carried out using a Keysight MSOX4024A oscilloscope

equipped with Frequency Response Analysis (FRA) functionality. The measurements were performed by simultaneously recording voltage and current using a voltage probe and a Keysight 1147B current probe.

The comparison between simulation and measurement results demonstrated good qualitative agreement of the obtained characteristics, confirming the validity of the proposed model. The observed quantitative discrepancies were mainly caused by the use of catalog parameters instead of parameters measured for the tested loudspeaker unit and by the influence of measurement conditions. The developed model can be applied in loudspeaker enclosure design, crossover analysis, and evaluation of audio amplifier loading conditions.

Keywords: Thiele–Small parameters, MATLAB, frequency analysis, resonant system, simulation.

Małgorzata Żukowicz
Koło Naukowo-Badawcze Turystyki „Tuptuś”

Jan Polaszczyk
opiekun Koła Naukowego

PODATKI TURYSTYCZNE W POLSCE I EUROPIE – ANALIZA PORÓWNAWCZA SYSTEMÓW

Streszczenie

Wraz z dynamicznym rozwojem turystyki rośnie rola mechanizmów finansowych, umożliwiających skuteczne zarządzanie ruchem turystycznym i wspieranie lokalnych społeczności. Ważnym narzędziem w tym zakresie są podatki turystyczne, które nie tylko przynoszą dochody publiczne, ale także pozwalają finansować infrastrukturę, chronić środowisko i utrzymywać atrakcyjność regionów dla turystów. Celem rozdziału jest przedstawienie i porównanie systemów podatków turystycznych w Polsce oraz wybranych krajach europejskich, w tym Francji, Hiszpanii, Włoszech i Grecji. Analizie poddano strukturę systemów, mechanizmy poboru, wysokość stawek oraz efektywność fiskalną, opierając się na metodzie porównawczej. Skuteczność systemów podatków turystycznych zależy od właściwego wyważenia wysokości stawek i sprawności mechanizmu poboru. Dobrze zaprojektowane systemy nie tylko zwiększają dochody publiczne, ale także wspierają rozwój turystyki i lokalnej gospodarki, podnosząc atrakcyjność regionów dla odwiedzających.

Słowa kluczowe: turystyka, podatek, opłata turystyczna, system podatkowy, Europa.

1. Wprowadzenie

Turystyka stanowi ważny sektor gospodarki zarówno w Polsce jak i innych krajach europejskich, generując znaczące dochody oraz wspierając rozwój lokalnych społeczności i infrastruktury turystycznej. Podatki i opłaty turystyczne odgrywają tu istotną rolę, umożliwiając finansowanie infrastruktury, ochronę środowiska oraz kontrolę natężenia ruchu turystycznego.

Celem rozdziału jest przedstawienie i porównanie systemów podatków turystycznych w Polsce oraz wybranych krajach europejskich, ze szczególnym uwzględnieniem ich struktury, mechanizmów poboru i skuteczności fiskalnej. Analiza została przeprowadzona metodą porównawczą, co pozwala zestawić różne modele opłat turystycznych i wskazać zarówno ich podobieństwa, jak i różnice.

Zakres badań obejmuje Polskę oraz wybrane kraje europejskie, w tym Francję, Hiszpanię, Włochy i Grecję, a okres analizowany obejmuje ostatnie dziesięć lat.

Analiza wskazuje, że systemy opłat turystycznych różnią się zarówno wysokością stawek, jak i sposobem organizacji oraz podziałem kompetencji między administracją centralną, a lokalną. Najwyższą skuteczność fiskalną wykazuje system francuski, oparty na jednolitych zasadach poboru i sprawnym nadzorze. W Hiszpanii i we Włoszech efektywność jest umiarkowana z uwagi na zdecentralizowany charakter systemów, natomiast w Polsce wpływy z opłat turystycznych pozostają ograniczone z powodu niskich stawek i wąskiego zakresu stosowania.

Rozdział podzielono na cztery zasadnicze części. Pierwsza z nich omawia podatek turystyczny w ujęciu teoretycznym, jego istotę, funkcjonowanie w Polsce oraz analizę funkcji, korzyści i ograniczeń tego rodzaju opłat. Druga część opisuje systemy opłat turystycznych w wybranych krajach europejskich. Następnie przeprowadzono analizę porównawczą, ukazując podobieństwa i różnice między systemami. Zakończenie przedstawia wnioski wynikające z badań oraz rekomendacje dotyczące projektowania skutecznych i efektywnych systemów podatków turystycznych.

Podsumowując efektywność systemów podatków turystycznych w dużej mierze zależy od odpowiedniego połączenia wysokości stawek z prawidłową organizacją poboru. Poprawnie zaprojektowane mechanizmy pozwalają zarówno zwiększać dochody publiczne, jak i wspierać rozwój turystyki oraz zachować atrakcyjność regionów dla odwiedzających.

2. Podatek turystyczny w ujęciu teoretycznym

2.1. Istota podatku turystycznego

Podatki są jednym z instrumentów finansów publicznych wykorzystywanych w turystyce. Ich znaczenie wzrasta wraz z dynamicznym rozwojem turystyki oraz rosnącymi potrzebami społeczności lokalnych.

Podatek turystyczny jest opłatą pobieraną od turystów za nocleg w hotelach, apartamentach i innych rodzajach tymczasowego zakwaterowania. W niektórych przypadkach ten podatek może być również pobierany za odwiedzanie określonych stref turystycznych i atrakcji[1]. Podatkiem turystycznym jest również forma opłaty, wiążącą się z dodatkowymi kosztami ponoszonymi przez samorząd lokalny[2]. Inne podejście przedstawia podatek turystyczny jako opłatę nakładaną na dobra i usługi nabywane głównie przez turystów. Zalicza się do niej najczęściej

[1] Trendway Travel, *Podatek turystyczny: co to jest i dlaczego jest potrzebny*, https://www.trendway-travel.pl/blog/podatek-turystyczny-dlaczego (dostęp: 26.02.2026).

[2] T. Wołowiec, N. Podolchak, *SPA and Local Tourist Fees as a Tributes of a Consumption Nature*, Teka Komisji Prawniczej PAN Oddział w Lublinie, vol. XV, 2022, no. 1, s. 378.

opłatę klimatyczną, która jest wnoszona przez hotelarzy za każdy sprzedany nocleg, opłaty wstępu do niektórych miejsc i opłaty za transport[3].

Przedstawione ujęcia wskazują, że podatek turystyczny może występować w różnych formach, jednak łączy je jego bezpośredni związek z ruchem turystycznym oraz korzystaniem z usług i infrastruktury turystycznej. Bez względu na sposób jego definiowania, obciążenie to dotyczy przede wszystkim turystów, a uzyskane wpływy służą finansowaniu wydatków ponoszonych przez jednostki samorządu terytorialnego. Podatek ten spełnia nie tylko funkcję fiskalną, ale również wpływa na kształtowanie rozwoju turystyki oraz ograniczanie jej niekorzystnych skutków.

2.2. System podatku turystycznego w Polsce

Podatek turystyczny stanowi szczególną kategorię świadczeń publicznych, którego obowiązek zapłaty wiąże się z czasowym pobytem w danym regionie turystycznym. W przeciwieństwie do klasycznych podatków, nie jest on związany z dochodem czy majątkiem podatnika, lecz z faktem korzystania z walorów środowiskowych, infrastruktury oraz usług dostępnych na obszarze odwiedzanym przez turystów. Jego zasadniczym celem jest częściowe przeniesienie kosztów utrzymania tych zasobów na osoby, które z nich korzystają, ale nie są stałymi mieszkańcami danego regionu[4].

W obowiązującym w Polsce systemie prawnym nie występuje bezpośrednio kategoria podatku turystycznego. Funkcje zbliżone do takiego instrumentu pełnią jednak opłaty lokalne, przede wszystkim opłata miejscowa oraz opłata uzdrowiskowa[5]. Podstawy ich funkcjonowania określa ustawa o podatkach i opłatach lokalnych, która umożliwia jednostkom samorządu terytorialnego wprowadzanie opłat na osoby czasowo przebywające na ich terenie[6]. W praktyce oznacza to, że każda gmina samodzielnie podejmuje decyzję o wprowadzeniu tych opłat oraz o ich wysokości, mieszcząc się w granicach określonych przez przepisy krajowe.

Opłata miejscowa jest stosowana w miejscowościach wyróżniających się określonymi walorami środowiskowymi i turystycznymi. Obejmuje osoby fizyczne przebywające na ich terenie dłużej niż jedną dobę, głównie w celach wypoczynkowych, turystycznych lub szkoleniowych[7]. Jej wprowadzenie nie jest obowiązkowe i zależy od spełnienia ustawowych kryteriów, takich jak korzystne

---

[3] A. Pawlowicz, Podatek Pigou na rynku turystycznym, Ekonomiczne Problemy Turystyki nr 13, Uniwersytet Szczeciński Zeszyty Naukowe nr 568, Szczecin 2009, s. 117-118.

[4] A. Hybel, *Opłata miejscowa – charakterystyka prawna i funkcjonalna*, Biuletyn Uzdrowiskowy, Urząd Miasta Nowego Sącza.

[5] K. Wałachowski, *Opłata turystyczna – propozycja zmian w Polsce*, Kraków 2025, s. 3.

[6] Ustawa z dnia 12 stycznia 1991 r. o podatkach i opłatach lokalnych (Dz.U. z 2025 r. poz. 707).

[7] Ibidem.

warunki klimatyczne, atrakcyjność krajobrazowa oraz odpowiednie zaplecze umożliwiające pobyt turystów[8].

Z kolei opłata uzdrowiskowa pobierana jest w miejscowościach posiadających status uzdrowiska i jest związana z wykorzystywaniem ich walorów leczniczych. Uzyskanie takiego statusu wymaga spełnienia szeregu wymogów określonych w przepisach, w tym posiadania naturalnych surowców leczniczych, odpowiednich warunków klimatycznych oraz rozwiniętej infrastruktury lecznictwa uzdrowiskowego[9]. W porównaniu z opłatą miejscową jej stawki są zwykle wyższe, co wynika ze specyfiki funkcjonowania tego typu obszarów i potrzeby utrzymania ich walorów. Choć formalnie za pobór opłat odpowiadają gminy, w praktyce zadanie to realizują podmioty świadczące usługi noclegowe. Hotele, pensjonaty oraz kwatery prywatne pełnią funkcję inkasentów, pobierając opłatę od turystów i przekazując ją do budżetu jednostki samorządu terytorialnego.

Przepisy przewidują również liczne zwolnienia z obowiązku ponoszenia opłat. Dotyczą one m.in. przedstawicieli personelu dyplomatycznego, osób przebywających w szpitalach, osób niewidomych wraz z przewodnikami, właścicieli domów letniskowych oraz zorganizowanych grup dzieci i młodzieży[10]. Dodatkowo jednostki samorządu terytorialnego mogą wprowadzić własne ulgi i zwolnienia, co zwiększa zróżnicowanie stosowanych rozwiązań[11]. Warto zaznaczyć, że na jednym obszarze nie jest dopuszczalne jednoczesne pobieranie opłaty miejscowej i uzdrowiskowej.

W wielu krajach europejskich stosuje się podobne opłaty, które różnią się formy i nazwą, np. opłaty klimatyczne, miejskie czy podatki od noclegu. Niezależnie od formy opłaty te są pobierane za czasowy pobyt i mają wspierać rozwój turystyki oraz utrzymanie atrakcyjności danego obszaru[12].

2.3. Funkcje podatków turystycznych

Podatki turystyczne pełnią ważną rolę w funkcjonowaniu gospodarki turystycznej oraz systemu finansów publicznych. Ich znaczenie nie ogranicza się wyłącznie do pozyskiwania dochodów, lecz obejmuje również wpływ na sposób zarządzania ruchem turystycznym oraz relacje pomiędzy odwiedzającymi

---

[8] Rozporządzenie Rady Ministrów z dnia 18 grudnia 2007 r. w sprawie warunków, jakie powinna spełniać miejscowość, w której można pobierać opłatę miejscową, (Dz.U. z 2007 r. nr 249 poz. 1851).

[9] Ustawa z dnia 28 lipca 2005 r. o lecznictwie uzdrowiskowym, uzdrowiskach i obszarach ochrony uzdrowiskowej oraz o gminach uzdrowiskowych, (Dz.U. z 2005 r. nr 167 poz. 1399).

[10] Art. 17 ust.2 ustawy dnia 12 stycznia 1991 r. o podatkach i opłatach lokalnych.

[11] D. Rynio, A. Zakrzewska-Półtorak, *Opłata miejscowa i uzdrowiskowa jako źródło dochodów lokalnych - analiza wysokości stawek na wybranych przykładach*, „Finanse Komunalne" 2023 nr 1, s. 33.

[12] OECD,*TourismTrends and Policies 2024*, OECD Publishing, Paris 2024, https://www.oecd.org/en/publications/oecd-tourism-trends-and-policies-2024_80885d8b-en.html (dostęp: 26.02.2026).

a społecznościami lokalnymi. Wśród najważniejszych funkcji podatków turystycznych wyróżnia się funkcję fiskalną, redystrybucyjną oraz regulacyjną.

Funkcja fiskalna stanowi podstawowe zadanie podatków turystycznych i polega na dostarczaniu środków finansowych do budżetów publicznych[13], zwłaszcza na poziomie lokalnym. Dochody te są istotne dla gmin i regionów o dużym ruchu turystycznym, gdzie pozwalają utrzymać infrastrukturę i wysoki standard usług, a także wspierają inwestycje w infrastrukturę turystyczną, poprawę estetyki przestrzeni publicznej, ochronę środowiska i promocję regionu, co przyczynia się do rozwoju lokalnej gospodarki i wzrostu atrakcyjności destynacji.

Z kolei funkcja redystrybucyjna wiąże się z rozkładem kosztów wynikających z obsługi ruchu turystycznego[14]. Obecność turystów generuje dodatkowe wydatki, takie jak utrzymanie dróg, zwiększone koszty sprzątania, gospodarki odpadami czy eksploatacji infrastruktury rekreacyjnej. Wprowadzenie podatków turystycznych pozwala na częściowe pokrycie tych kosztów przez osoby korzystające z lokalnych zasobów, zmniejszając tym samym obciążenie finansowe mieszkańców. Mechanizm ten wpisuje się w zasadę, że użytkownicy powinni partycypować w kosztach korzystania z dóbr publicznych.

Funkcja regulacyjna odnosi się do możliwości oddziaływania na wielkość i strukturę ruchu turystycznego[15]. Zróżnicowanie stawek w sezonie i poza nim można równomiernie rozłożyć ruch turystyczny oraz ograniczyć negatywne skutki nadmiernej koncentracji turystów, wspierając w ten sposób zrównoważony rozwój turystyki.

2.4. Podatki turystyczne – korzyści i ograniczenia

Podatki turystyczne stanowią jedno z narzędzi stosowanych przez władze publiczne w celu zarządzania rozwojem turystyki oraz finansowania zadań związanych z obsługą ruchu turystycznego. Ich wprowadzenie wiąże się jednak z koniecznością rozważenia zarówno potencjalnych korzyści, jak i możliwych negatywnych konsekwencji, szczególnie w kontekście konkurencyjności ośrodka turystycznego[16].

Korzyści wynikające z opłat turystycznych dotyczą przede wszystkim aspektów finansowych i organizacyjnych. Środki te mogą być wykorzystywane na rozwój infrastruktury turystycznej, utrzymanie przestrzeni publicznej, ochronę środowiska oraz działania promocyjne. Dzięki temu oferta turystyczna staje się bardziej atrakcyjna, a część kosztów obsługi turystów zostaje przeniesiona na osoby odwiedzające, zmniejszając obciążenie finansowe mieszkańców[17].

---

[13] G. Szczodrowski, *Polski system podatkowy*, Wydawnictwo Naukowe PWN, Warszawa 2007, s. 30.
[14] Ibidem.
[15] A. Majchrzycka-Guzowska, *Finanse i prawo finansowe*, Wolters Kluwer, Warszawa 2019, s. 115.
[16] K. Wałachowski, *Opłata turystyczna…,* op. cit.
[17] Ibidem.

Wskazuje się także na ograniczenia związane z wprowadzeniem takich obciążeń. Jednym z głównych argumentów jest ryzyko spadku atrakcyjności cenowej regionu, co może wpływać na decyzje turystów, szczególnie w warunkach silnej konkurencji, kierując ich do innych miejsc o podobnych walorach bez dodatkowych kosztów[18]. Dodatkowo wprowadzenie opłat wiąże się z obowiązkami administracyjnymi, takimi jak pobór i ewidencja środków[19].

Podatki turystyczne oddziałują na konkurencyjność regionów w sposób złożony, a ich efekt zależy od wielu czynników, takich jak wysokość stawek, metoda poboru i sposób wykorzystania uzyskanych środków. Odpowiednio skonstruowany system opłat może poprawiać jakość infrastruktury i usług, zwiększając atrakcyjność obszaru i wzmacniając jego pozycję na rynku turystycznym. Zbyt wysokie opłaty mogą zmniejszać zainteresowanie turystów i osłabiać konkurencyjność, szczególnie w regionach o podobnej ofercie.

Podsumowując, oddziaływanie podatków turystycznych na konkurencyjność regionów jest złożone i uzależnione od takich czynników, jak wysokość opłat, sposób ich wykorzystania oraz atrakcyjność danego miejsca. Ważne jest zachowanie równowagi między pozyskiwaniem dochodów a utrzymaniem wysokiego standardu oferty turystycznej.

3. Podatki turystyczne w wybranych krajach europejskich

Podatki turystyczne odgrywają istotną rolę w polityce finansowej wielu państw europejskich. Stanowią one nie tylko źródło dochodów dla budżetów lokalnych, ale również narzędzie wspomagające zarządzanie ruchem turystycznym. Ich forma, wysokość stawek oraz sposób poboru różnią się w zależności od kraju, co wynika zarówno z odmiennych uwarunkowań prawnych, jak i specyfiki poszczególnych rynków turystycznych.

W Grecji funkcjonuje opłata klimatyczna powiązana z korzystaniem z usług noclegowych. Jest naliczana kwotowo za każdą dobę pobytu, a jej wysokość zależy od standardu obiektu zakwaterowania, najczęściej określanego liczbą gwiazdek. W praktyce oznacza to, że osoby korzystające z obiektów o wyższym standardzie ponoszą większe obciążenia finansowe. Pobór opłaty realizowany jest przez obiekty noclegowe, które następnie przekazują środki do budżetu państwa. Uzyskane dochody wykorzystywane są m.in. na rozwój infrastruktury turystycznej, działania promocyjne oraz ograniczanie negatywnych skutków intensywnego ruchu turystycznego, szczególnie w najpopularniejszych miejscach kraju[20].

---

[18] M. Burzec, *Opłata miejscowa – relikt przeszłości czy efektywne źródło dochodów gmin*, Kwartalnik Prawno-Finansowy 2019, nr 3, Katolicki Uniwersytet Lubelski Jana Pawła II, s. 15-18.

[19] A. Hybel, *Opłata miejscowa…,* op. cit.

[20] Acropolis Greece, *Greece Implements New ClimateResilienceLevy for Tourists*, https://acropolisgreece.com/2024/01/08/greece-implements-new-climate-resilience-levy-for-tourists/ (dostęp: 26.02.2026).

Odmienny model opłat turystycznych przyjęto w Hiszpanii, który ma charakter zdecentralizowany, a władze regionalne decydują o ich wprowadzeniu i wysokości, co prowadzi do zróżnicowania w poszczególnych regionach. Opłaty pobierane są od osób korzystających z różnych form zakwaterowania, a ich wysokość zależy od standardu obiektu, jego lokalizacji czy długości pobytupodmioty świadczącychsystemy są najbardziej rozbudowane, a niekiedy stosuje się ograniczenia liczby dni objętych opłatą. Środki te przeznaczane są głównie na rozwój infrastruktury, ochronę środowiska i zarządzanie ruchem turystycznym w najbardziej obciążonych miejscach[21].

W Portugalii opłaty turystyczne mają głównie charakter lokalny i są wprowadzane przez władze samorządowe, głównie w największych ośrodkach turystycznych. Najczęściej naliczane są jako stała kwota za każdą dobę pobytu, często z ograniczeniem liczby dni objętych opłatą. Wysokość stawek ustalana jest na poziomie lokalnym, co pozwala na dostosowanie ich do specyfiki danego miasta oraz skali ruchu turystycznego. Stawki ustala się lokalnie, dostosowując je do specyfiki miasta i natężenia ruchu turystycznego. Pobór realizują obiekty noclegowe, które przekazują środki do budżetu samorządu, wykorzystywane na utrzymanie infrastruktury, poprawę przestrzeni publicznej oraz promocję turystyki. Opłaty te odpowiadają na rosnącą liczbę turystów i związane z tym wyzwania organizacyjne oraz środowiskowe[22].

We Włoszech system opłat turystycznych charakteryzuje się dużą elastycznością wynikającą z szerokich kompetencji gmin. To one decydują o wprowadzeniu opłaty, jej wysokości oraz zasadach stosowania. Opłata naliczana jest zwykle za każdą dobę pobytu, a jej poziom zależy od standardu obiektu noclegowego i atrakcyjności danej miejscowości. W popularnych destynacjach stawki są wyższe, odzwierciedlając koszty utrzymania infrastruktury oraz ochrony dziedzictwa kulturowego. Często stosuje się także limity dni objętych opłatą oraz różnego rodzaju zwolnienia, np. dla dzieci, osób z niepełnosprawnościami, turystów przebywających w celach zdrowotnych lub służbowych oraz mieszkańców gminy[23].

Francja wyróżnia się stosunkowo uporządkowanym i przejrzystym systemem opłat turystycznych. Podatek znany jako taxe de séjour, posiada umocowanie w przepisach krajowych, jednak jego szczegóły mogą być ustalane lokalnie. Opłata pobierana jest od osób korzystających z noclegów i może przyjmować zarówno formę stałej kwoty za dobę pobytu, jak i stawki procentowej powiązanej z ceną usługi. Jej wysokość zależy zazwyczaj od standardu obiektu zakwaterowania. Pobór realizują obiekty noclegowe, w tym platformy pośredniczące

---

[21] Sekrety Hiszpanii, *Podatek turystyczny Hiszpania 2026: Gdzie płacisz, gdzie nie*, https://www.sekretyhiszpanii.pl/praktyczne/podatek-turystyczny-hiszpania-warto-wiedziec/ (dostęp: 26.02.2026).

[22] Idealista/news, *TouristTax in Portugal: 2025 rates for Lisbon, the Algarve& Madeira*, https://www.idealista.pt/en/news/legal-advice-in-portugal/2025/08/05/70982-tourist-tax-in-portugal-2025-rates-for-lisbon-the-algarve-madeira (dostęp: 26.02.2026).

[23] Portale del Contribuente, *Tourist Tax, What it is, who it applies to, how to proceed, and when*, https://www.portaledelcontribuente.it/en/thematic-area/tourist-tax/ (dostęp: 26.02.2026).

w najmie krótkoterminowym, co w ostatnich latach zwiększyło efektywność systemu[24].

Pomimo różnic w systemach opłat turystycznych, w większości państw europejskich opierają się one na podobnych zasadach. Najczęściej mają formę opłat związanych z noclegiem, pobieranych przez podmioty świadczące usługi zakwaterowania, a ich wysokość uzależniona jest od standardu obiektu lub długości pobytu. Uzyskane środki wykorzystywane są przede wszystkim na rozwój infrastruktury, utrzymanie przestrzeni publicznej, ochronę środowiska oraz promocję turystyki.

4. Analiza porównawcza systemów opłat turystycznych

Systemy opłat turystycznych w Europie różnią się pod względem struktury, wysokości stawek, metod poboru oraz stosowanych zwolnień, co wynika zarówno z odmiennych regulacji prawnych, jak i specyfiki lokalnych rynków turystycznych. Ich głównym celem jest pozyskiwanie dodatkowych środków dla władz lokalnych oraz częściowe pokrycie kosztów związanych z obsługą ruchu turystycznego, ochroną środowiska i dziedzictwa kulturowego. Analiza systemów w Grecji, Hiszpanii, Portugalii, Włoszech, Francji i Polsce pokazuje wspólne cechy, takie jak powiązanie opłat z noclegami i przeznaczenie środków na rozwój infrastruktury, oraz różnice w decentralizacji, sposobie naliczania i autonomii władz lokalnych. Wyboru tych krajów dokonano ze względu na ich zróżnicowane podejście do opłat turystycznych i istotną rolę w europejskim ruchu turystycznym, co umożliwia przedstawienie szerokiego spektrum rozwiązań. Tabela 1. przedstawia porównanie wybranych systemów opłat turystycznych w tych państwach w 2026 roku, uwzględniając rodzaj opłat, wysokość stawek oraz zasady poboru, co pozwala na syntetyczne zestawienie podobieństw i różnic.

[24] Travel Insider*, France TouristTax 2026 Paris & French RivieraVisitorFeesExplained*, https://travelinsider.blog/france-tourist-tax-2025/ (dostęp: 26.02.2026).

Tabela 1. Opłata turystyczna w wybranych państwach Europy w 2026 roku

Table 1. Tourist tax in selected European countries in 2026

Kraj i sposób naliczania podatku	**Stawki 2026**	**Podmiot pobierający opłatę**	**Zwolnienia/ preferencje**	**Efektywność fiskalna**
Polska (osoba/ doba)	Opłata miejscowa: 3,46 zł Opłata uzdrowiskowa: 6,67 zł[25]	Pobierane przez hotele, pensjonaty, kwatery prywatne; środki przekazywane do budżetu gminy	Dzieci, osoby niepełnosprawne, zorganizowane grupy szkolne, mieszkańcy gminy	Niska–średnia; ograniczone wpływy i duża różnorodność stawek między gminami
Grecja (pokój/noc)	Hotele 1–2★: 2 € 3★: 5 € 4★: 10 € 5★: 15 €	Pobierane bezpośrednio w obiektach noclegowych, środki przekazywane do budżetu państwa	Dzieci, osoby niepełnosprawne	Średnia; efektywnie zbierane w popularnych regionach, ograniczone w mniejszych miejscowościach
Hiszpania (Barcelona) **(osoba/noc)**	Hostele/ apartamenty:4–7 € Hotele 4★:6–10 € Hotele 5★:8–12 €	Pobierane przez obiekty noclegowe; władze regionalne ustalają stawki	Ograniczenie liczby dni pobytu, dzieci	Średnia-wysoka; system zdecentralizowany umożliwia dostosowanie do lokalnych potrzeb, ale utrudnia kontrolę
Portugalia (osoba/noc)	Lizbona: 4 € (maks. 7 nocy) Porto: 3 € (maks. 7 nocy) Algarve i mniejsze miasta: 1–2 €	Pobierane przez obiekty noclegowe, przekazywane do samorządów	Dzieci, osoby niepełnosprawne	Średnia; niski poziom stawek ogranicza dochody, ale pobór jest prosty
Włochy (Mediolan) **(osoba/noc)**	Hotele 2★: 5 € 3★: 7,40 € 4–5★: 10 € B&B/apartamenty: 9,50 €	Pobierane przez obiekty noclegowe; gminy decydują o wprowadzeniu i wysokości opłaty	Dzieci, osoby niepełnosprawne, przebywający w celach zdrowotnych lub służbowych, mieszkańcy gminy	Średnia; wysoka elastyczność, zróżnicowanie stawek utrudnia pełną kontrolę
Francja (osoba/noc)	Większość miast: 0,25–5 € Paryż:8,45–15,93€	Pobierane przez obiekty noclegowe i platformy najmu krótkoterminowego	Dzieci, osoby niepełnosprawne	Wysoka; jednolity system i wsparcie platform zwiększają efektywność fiskalną

Źródło: opracowanie własne na podstawie danych: Infor, *Podatki i opłaty lokalne w 2026 r. – stawki maksymalne*,InteriaBiznes, *Zmiany dla podróżujących do Grecji. Wakacje 2025 będą droższe*, Sekrety Hiszpanii, *Podatek turystyczny Hiszpania 2026: Gdzie płacisz, gdzie nie*, Idealista/News, *TouristTax in Portugal: 2025 rates for Lisbon, the Algarve & Madeira*, Comune de Milano HostCity, *Imposta di soggiorno: informazioni*, Service Public Entreprendre, *Touristtax: whatare the rates?*

[25] Podane kwoty są stawkami maksymalnymi.

W krajach europejskich funkcjonują różne rozwiązania dotyczące opłat związanych z turystyką, które różnią się zarówno wysokością stawek, jak i sposobem ich organizacji. W Polsce obciążenia są stosunkowo niewielkie i mają charakter lokalny. W innych państwach opłaty są znacznie wyższe, a ich wysokość najczęściej zależy od standardu obiektów noclegowych lub atrakcyjności turystycznej danego miejsca. Największe koszty ponoszą turyści odwiedzający popularne metropolie, takie jak Barcelona czy Paryż, gdzie ruch turystyczny jest szczególnie intensywny.

Mimo różnic w wysokości stawek, mechanizm poboru opłat w większości analizowanych krajów jest podobny. Za naliczanie i odprowadzanie opłat odpowiadają podmioty świadczące usługi noclegowe, co upraszcza funkcjonowanie systemu oraz zwiększa jego skuteczność. Zróżnicowanie pojawia się przede wszystkim w organizacji systemu i podziale kompetencji między administracją centralną a lokalną. W Hiszpanii i we Włoszech władze lokalne odgrywają kluczową rolę, samodzielnie ustalając stawki i zasady poboru, co pozwala dopasować system do warunków regionalnych, ale jednocześnie prowadzi do dużego rozdrobnienia. We Francji natomiast wprowadzono bardziej jednolity i przejrzysty model, który ułatwia kontrolę i podnosi skuteczność poboru. W Grecji natomiast wpływy z opłat trafiają do budżetu centralnego, ograniczając udział samorządów w dystrybucji środków.

Wspólne cechy systemów opłat turystycznych w badanych krajach można dostrzec przede wszystkim w zakresie stosowanych zwolnień i ulg. Najczęściej dotyczą one dzieci oraz osób z niepełnosprawnościami, choć w niektórych państwach katalog uprzywilejowanych grup jest szerszy. Zakres tych zwolnień bezpośrednio wpływa na poziom generowanych dochodów. Efektywność fiskalna stanowi z kolei jeden z kluczowych czynników różnicujących poszczególne systemy. Najlepsze rezultaty osiągają rozwiązania łączące odpowiedni poziom stawek z przejrzystymi zasadami poboru oraz sprawnym nadzorem, co ma miejsce we Francji. W Hiszpanii i we Włoszech efektywność jest umiarkowana, pomimo bardziej zdecentralizowanej organizacji systemów. W Polsce natomiast wpływy z opłat turystycznych są ograniczone ze względu na niskie stawki oraz węższy zakres stosowania.

Podsumowując efektywność systemu opłat turystycznych zależy od zachowania równowagi między wysokością stawek a sprawnością organizacyjną całego mechanizmu. Odpowiednio skonstruowane rozwiązania pozwalają nie tylko na generowanie dochodów publicznych, ale także na finansowanie działań wspierających rozwój turystyki, przy jednoczesnym utrzymaniu atrakcyjności regionu dla odwiedzających.

Literatura

[1] Burzec M.: Opłata miejscowa – relikt przeszłości czy efektywne źródło dochodów gmin, Kwartalnik Prawno-Finansowy 2019, nr 3, Katolicki Uniwersytet Lubelski Jana Pawła II, s. 15-18.

[2] Hybel A.: Opłata miejscowa – charakterystyka prawna i funkcjonalna, Biuletyn Uzdrowiskowy, Urząd Miasta Nowego Sącza.

[3] Majchrzycka-Guzowska A.: Finanse i prawo finansowe, Wolters Kluwer, Warszawa 2019, s. 115.

[4] Pawlowicz A., Podatek Pigou na rynku turystycznym, Ekonomiczne Problemy Turystyki nr 13, Uniwersytet Szczeciński Zeszyty Naukowe nr 568, Szczecin 2009, s. 117-118.

[5] Rynio D., Zakrzewska-Półtorak A.: Opłata miejscowa i uzdrowiskowa jako źródło dochodów lokalnych - analiza wysokości stawek na wybranych przykładach, „Finanse Komunalne" 2023 nr 1, s. 33.

[6] Szczodrowski G.: Polski system podatkowy, Wydawnictwo Naukowe PWN, Warszawa 2007, s. 30.

[7] Wałachowski K.: Opłata turystyczna – propozycja zmian w Polsce, Kraków 2025, s. 3.

[8] Wołowiec T., Podolchak N., *SPA and Local Tourist Fees as a Tributes of a Consumption Nature*, Teka Komisji Prawniczej PAN Oddział w Lublinie, vol. XV, 2022, no. 1, s. 378.

Akty prawne

[1] Ustawa z dnia 12 stycznia 1991 r. o podatkach i opłatach lokalnych (Dz.U. z 2025 r., poz. 707).

[2] Ustawa z dnia 28 lipca 2005 r. o lecznictwie uzdrowiskowym, uzdrowiskach i obszarach ochrony uzdrowiskowej oraz o gminach uzdrowiskowych, (Dz.U. z 2005 r., nr 167, poz. 1399).

[3] Rozporządzenie Rady Ministrów z dnia 18 grudnia 2007 r. w sprawie warunków, jakie powinna spełniać miejscowość, w której można pobierać opłatę miejscową, (Dz.U. z 2007 r., nr 249, poz. 1851).

Źródła internetowe

[1] Acropolis Greece: Greece Implements New ClimateResilienceLevy for Tourists, https://acropolis-greece.com/2024/01/08/greece-implements-new-climate-resilience-levy-for-tourists/ (dostęp: 26.02.2026).

[2] Comune de Milano Host City: Imposta di soggiorno: informazioni,https://www.comune.milano.it/argomenti/tributi/imposta-di-soggiorno-informazioni (dostęp: 26.02.2026).

[3] Idealista/news: TouristTaxin Portugal: 2025 rates for Lisbon, the Algarve& Madeira, https://www.idealista.pt/en/news/legal-advice-in-portugal/2025/08/05/70982-tourist-tax-in-portugal-2025-rates-for-lisbon-the-algarve-madeira (dostęp: 26.02.2026).

[4] Infor: Podatki i opłaty lokalne w 2026 r. - stawki maksymalne, https://ksiegowosc.infor.pl/podatki/podatki-osobiste/podatki-i-oplaty-lokalne/7213935,podatki-i-oplaty-lokalne-w-2026-r-stawki-maksymalne.html (dostęp: 26.02.2026).

[5] InteriaBiznes: Zmiany dla podróżujących do Grecji. Wakacje 2025 będą droższe, https://biznes.interia.pl/finanse/news-zmiany-dla-podrozujacych-do-grecji-wakacje-2025-beda-drozsze,nId,7888422 (dostęp: 26.02.2026).

[6] OECD: TourismTrends and Policies 2024, OECD Publishing, Paris 2024, https://www.oecd.org/en/publications/oecd-tourism-trends-and-policies-2024_8 0885d8b-en.html (dostęp: 26.02.2026).

[7] Portale del Contribuente: Tourist Tax, https://www.portaledelcontribuente.it/en/thematic-area/tourist-tax/ (dostęp: 26.02.2026).

[8] Sekrety Hiszpanii: Podatek turystyczny Hiszpania 2026: Gdzie płacisz, gdzie nie, https://www.sekretyhiszpanii.pl/praktyczne/podatek-turystyczny-hiszpania-warto-wiedziec/ (dostęp: 26.02.2026).

[9] Service Public Entreprendre: Touristtax: whatare the rates?, https://entreprendre.service-public.gouv.fr/vosdroits/F31635?lang=en (dostęp: 18.2.2026).

[10] Travel Insider: France TouristTax 2026 Paris & French RivieraVisitor Fees Explained, https://travelinsider.blog/france-tourist-tax-2025/ (dostęp: 26.02.2026).

[11] Trendway Travel: Podatek turystyczny: co to jest i dlaczego jest potrzebny, https://www.trendwaytravel.pl/blog/podatek-turystyczny-dlaczego (dostęp: 26.02.2026).

TOURIST TAXES IN POLAND AND EUROPE – A COMPARATIVE ANALYSIS OF SYSTEMS

Summary

With the dynamic growth of tourism, the role of financial mechanisms enabling effective management of tourist flows and support for local communities is increasing. An important tool in this context is tourism taxes, which not only generate public revenue but also allow for financing infrastructure, protecting the environment, and maintaining the attractiveness of regions for visitors. The aim of this article is to present and compare tourism tax systems in Poland and selected European countries, including France, Spain, Italy, and Greece. The analysis focuses on the structure of the systems, collection mechanisms, the level of rates, and fiscal efficiency, based on a comparative method. The effectiveness of tourism tax systems depends on the proper balance between the level of rates and the efficiency of the collection mechanism. Well-designed systems not only increase public revenues but also support the development of tourism and the local economy, enhancing the attractiveness of regions for visitors.

Keywords: tourism, tax, tourism fee, tax system, Europe.

DOCUMENT CREATED WITH

PDF COMBINER



PDF Combiner is a free application that you can use to combine multiple PDF documents into one.

Three simple steps are needed to merge several PDF documents. First, we must add files to the program. This can be done using the Add files button or by dragging files to the list via the Drag and Drop mechanism. Then you need to adjust the order of files if list order is not suitable. The last step is joining files. To do this, click button Combine PDFs.

Main features:

secure PDF merging - everything is done on your computer and documents are not sent anywhere

simplicity - you need to follow three steps to merge documents

possibility to rearrange document - change the order of merged documents and page selection

reliability - application is not modifying a content of merged documents.

Visit the homepage to download the application:

www.jankowskimichal.pl/pdf-combiner

To remove this page from your document, please donate a project.