



NAUKA – **TECHNIKA** – TECHNOLOGIA

S E R I A W Y D A W N I C Z A A G H

T O M 11



WYDAWNICTWA AGH

KRAKÓW 2025

NAUKA – TECHNIKA – TECHNOLOGIA

S E R I A W Y D A W N I C Z A A G H

T O M 11



WYDAWNICTWA AGH

KRAKÓW 2025

Wydawnictwa Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie

© Wydawnictwa AGH, Kraków 2025

ISBN 978-83-68219-80-7

ISSN 3071-7027

Monografia elektroniczna dostępna online.

DOI: <https://doi.org/10.7494/978-83-68219-80-7>

Publikacja może być rozpowszechniana zgodnie z licencją Creative Commons CC BY 4.0

Komitet Redakcyjny serii „Nauka – Technika – Technologia”:

redaktor: *prof. dr hab. inż. Andrzej Gonet* (Akademia Górniczo-Hutnicza, Wydział Wiertnictwa, Nafty i Gazu, Kraków)

sekretarz: *dr inż. Robert Rado* (Akademia Górniczo-Hutnicza, Wydział Wiertnictwa, Nafty i Gazu, Kraków)

Zamieszczone w tomie publikacje zostały pozytywnie zaopiniowane przez recenzentów wyznaczonych przez Komitet Redakcyjny serii.

Opracowanie redakcyjne i korekta: *Monika Filipek, Magdalena Grzech, Agnieszka Rusinek, Kamila Zimnicka*

Projekt okładki i strony tytułowej serii: *Paweł Sepielak*

Skład komputerowy: Andre

Wydawnictwa AGH

al. A. Mickiewicza 30, 30-059 Kraków

tel. 12 617 32 28, 12 636 40 38

e-mail: redakcja@wydawnictwoagh.pl

www.wydawnictwo.agh.edu.pl

Spis treści

Karol Adamiec

Projekt, optymalizacja i wykonanie prototypu sportowego zacisku hamulcowego do motocykli z wykorzystaniem technologii przyrostowych	5
---	---

Miłosz Gaszyna

Pozyskiwanie i analiza danych o cyberatakach z wykorzystaniem systemu honeypot	23
---	----

Filip Kowalski

Optymalizacja algorytmu <i>wave function collapse</i> pod kątem czasu wykonania	43
--	----

Jakub Kutek, Łukasz Kwinta, Mariusz Gibiec

Rozwój sterowników napędów elektrycznych i ich implementacja w łażiku planetarnym Kalman	61
---	----

Łukasz Łukańko, Joanna Wartak

Oil and gas industry challenges in light of the new EU methane law	73
--	----

Grzegorz Olszyna, Andrzej Tytko

Badania trwałości zmęczeniowej typowych lin stalowych stosowanych w górniczych wyciągach szybowych	83
---	----

Piotr Paniec, Maciej Stanuch, Andrzej Skalski

Segmentacja aorty z obrazów tomografii komputerowej przy użyciu głębokich sieci neuronowych	97
--	----

Michał Przepiórski, Konrad Staruch

Projekt i realizacja tunelu aerodynamicznego o obiegu zamkniętym	111
--	-----

Henryk Sechman, Sandra Płochocka, Jan Adam Sobon

Szczelność naftowych otworów wiertniczych w świetle powierzchniowych badań geochemicznych – studium przypadku z monokliny przedsudeckiej	121
--	-----

Anna Siwik

Polskie środowiska emigracyjne – różne strategie działania w obliczu klęski idei suwerennej Polski w 1945 roku	143
---	-----

Tomasz Styn, Michał Pitera, Mateusz Okoń, Patryk Harasik

Zastosowanie kryptografii postkwantowej w sieciach IoT	161
--	-----

Jan Ziaja, Robert Rado

Niestandardowe możliwości zastosowania technologii horyzontalnego przewiertu sterowanego – HDD	179
---	-----

Projekt, optymalizacja i wykonanie prototypu sportowego zacisku hamulcowego do motocykli z wykorzystaniem technologii przyrostowych

Karol Adamiec 

Akademia Górniczo-Hutnicza, Wydział Odlewnictwa, Kraków

Streszczenie: W niniejszym artykule przedstawiono prototyp motocyklowego zacisku hamulcowego, zaprojektowanego z myślą o produkcji przyrostowej technologią SLM ze stopu Ti6Al4V. Celem projektu była maksymalna redukcja masy w stosunku do pierwotnego korpusu zacisku przy jednoczesnym zachowaniu wymaganej wytrzymałości w warunkach eksploatacyjnych. W ramach pracy wykorzystano inżynierię odwrotną z zastosowaniem technologii skanowania 3D światłem strukturalnym, co pozwoliło na zachowanie kompatybilności z wybranym motocyklem. Geometrię prototypowego zacisku opracowano na podstawie analiz symulacji naprężeń oraz optymalizacji MES, uwzględniając wyznaczony układ obciążeń w warunkach maksymalnego hamowania. Artykuł prezentuje postępowanie projektowe, oparte na optymalizacji konstrukcji dla technik produkcji przyrostowych i nadające się do adaptacji w produkcji szerszej grupy podzespołów motoryzacyjnych.

Słowa kluczowe: optymalizacja geometrii, analiza wytrzymałości, metoda elementów skończonych, produkcja przyrostowa, selektywne stapianie laserowe, zacisk hamulcowy

DESIGN, OPTIMIZATION AND PRODUCTION OF A PROTOTYPE SPORTS BRAKE CALIPER FOR MOTORBIKES BASED ON ADDITIVE TECHNOLOGY

Abstract: This article presents design process of a prototype brake caliper intended for a motorcycle, developed with additive manufacturing in mind using the SLM technology and the Ti6Al4V alloy. The objective of the project was to achieve maximum weight reduction compared to the original caliper body while maintaining the required strength under operational conditions. As part of the work, reverse engineering was utilized with the application of 3D structured light scanning technology, ensuring compatibility with the previously selected motorcycle. The geometry of the prototype caliper was developed based on finite element method (FEM) stress simulation analyses and optimization, considering the calculated load system under maximum braking conditions. The paper outlines the design process based on structural optimization for additive manufacturing techniques, demonstrating an approach that can be adapted to the production of a broader range of automotive components.

Keywords: shape optimization, strength analysis, finite element method, additive manufacturing, selective laser melting, brake caliper

https://doi.org/10.7494/978-83-68219-80-7_1

1. Wprowadzenie

Motorsport stanowi główne źródło rozwoju branży motoryzacyjnej. Rozwiązania występujące we współczesnych pojazdach drogowych wielokrotnie poprzedzał debiut na torze wyścigowym. Wyścigi samochodowe i motocyklowe są nie tylko obszarem rywalizacji kierowców, ale również całych zespołów inżynierów, wyznaczających trendy w projektowaniu pojazdów sportowych i koncepcyjnych.

Współcześnie kluczowym kierunkiem rozwoju w branży motoryzacyjnej jest minimalizacja masy produkowanych podzespołów przy jednoczesnym zachowaniu wysokich parametrów wytrzymałościowych. Taka kombinacja właściwości zapewnia poprawę niemal całej charakterystyki pojazdu: przyspieszenia, prowadzenia, drogi hamowania czy zużycia paliwa – czynników szczególnie istotnych podczas zawodów sportowych.

Odpowiedzią na wygórowane wymagania techniczne wobec samochodów i motocykli sportowych jest stosowanie materiałów ultralekkich zarówno do produkcji elementów o walorach estetycznych, jak i konstrukcyjnych. Wymienione cechy spełniają między innymi stopy tytanu i magnezu oraz kompozyty węglowe.

Rozwój nowoczesnych technik projektowania i produkcji umożliwia wykorzystanie programów opartych na metodzie elementów skończonych (MES) do optymalizacji geometrii części pod kątem wykonania ich w technologii przyrostowej. Tak zaprojektowane elementy, wykonywane metodą produkcji addytywnej ze stopów tytanu, stanowią szczytowe osiągnięcie współczesnej branży samochodowej.

Celem niniejszego projektu było opracowanie prototypowego korpusu zacisku hamulcowego przeznaczonego do motocykla. Przy wytycznych projektowych uwzględniono optymalizację geometrii produkowanych tradycyjnie rozwiązań, zakładając zmniejszenie masy zacisku, przy jednoczesnym zachowaniu wymaganej wytrzymałości w warunkach eksploatacyjnych oraz utrzymaniu kompatybilności z innymi elementami motocykla. Projekt został przystosowany do późniejszej produkcji korpusu zacisku w technologii przyrostowej SLM (ang. *selective laser melting* – selektywnego stapiania laserowego) ze stopu Ti6Al4V.

Projekt zakładał zaprojektowanie korpusu zacisku hamulcowego do motocykla o przeznaczeniu wyścigowym, co pozwoliło skupić się na maksymalnej poprawie osiągow, pomijając jednocześnie aspekty ekonomiczne wykonania podzespołów. Daje to możliwość doboru najwyższej klasy materiałów i technologii produkcji zaprojektowanego elementu.

W celu uzyskania możliwie rzeczywistych podstaw do projektu zdecydowano się na zastosowanie inżynierii odwrotnej z wykorzystaniem istniejącego motocyklowego zacisku hamulcowego. Ze względu na dostępność informacji oraz cenę zakupu potrzebnych podzespołów projekt opiera się na modelu motocykla drogowego, sportowo-turystycznego. Dostępne dane techniczne na temat motocykla, układu hamulcowego i zacisków zostały zaczerpnięte z instrukcji serwisowej pojazdu oraz z własnych pomia-

rów, zaś w przypadku brakujących informacji przyjęto uogólnione założenia dla tej grupy motocykli. Takie podejście znacznie ułatwiło realizację projektu, a jednocześnie pozwoliło na przedstawienie całego procesu, który może być finalnie wykorzystany przy produkcji motocykli wyścigowych.

Geometria prototypowego zacisku została opracowana na podstawie symulacji wytrzymałościowej i optymalizacji metodą elementów skończonych (MES), przeprowadzonych dla warunków odpowiadających maksymalnej sile hamowania motocykla. Proces optymalizacji geometrii korpusu ograniczono w taki sposób, aby zapewnić pełną kompatybilność elementu z motocyklem, w którym wykorzystano pierwotne rozwiązanie.

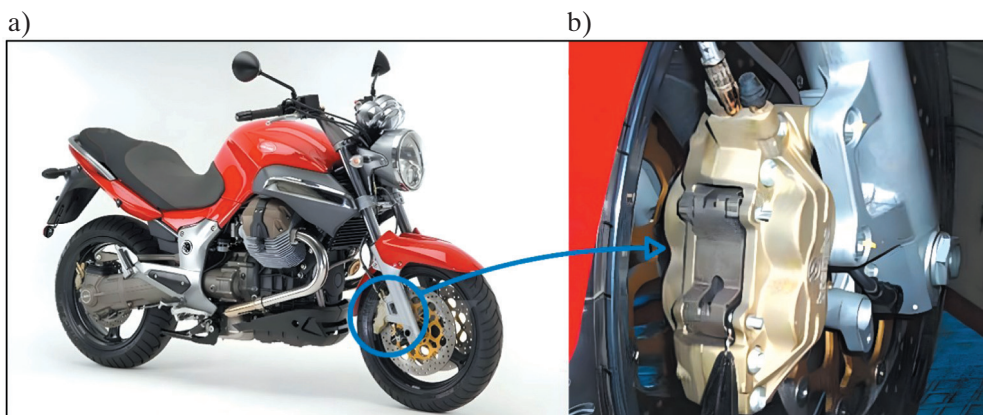
Wraz ze zmianą geometrii zacisku dopuszczono ingerencję w materiał i technologię produkcji zacisku. Aby umożliwić najwyższy stopień optymalizacji bryły pod kątem naprężeń, wybrano produkcję techniką addytywną SLM (selektywnego spiekania laserowego) z wykorzystaniem stopu tytanu Ti6Al4V.

Finalny projekt korpusu zacisku został zaprezentowany w formie komputerowego modelu CAD oraz jako prototyp koncepcyjny wydrukowany w 3D z tworzywa sztucznego, mający na celu prezentację rzeczywistej geometrii.

2. Inżynieria odwrotna

2.1. Opis i charakterystyka pierwotnego zacisku hamulcowego

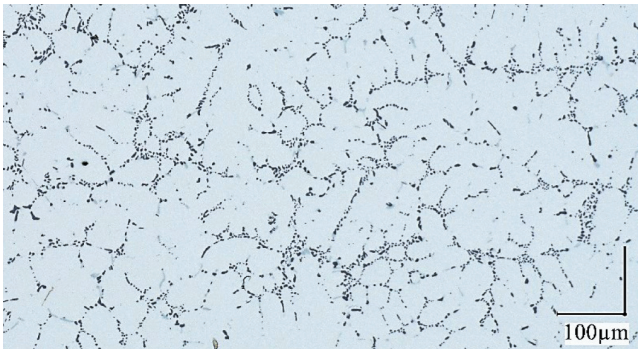
Do projektu wykorzystano motocykl włoskiej marki Moto Guzzi. Wybrany model, Brevia V 1100 (rys. 1a), wyposażony jest w układ o trzech hamulcach tarczowych – dwóch przednich (rys. 1b) i pojedynczego tylnego (Moto Guzzi). Przednie zaciski hamulcowe oraz zacisk tylny wyprodukowała firma Brembo.



Rys. 1. Wybrany motocykl (a) (fot. F. Paolillo);
przedni zacisk hamulcowy (b) (fot. Moto Guzzi)

Tabela 1
Najważniejsze dane opisujące pierwotny zacisk hamulcowy

Parametr	Wartość	Źródło danych
Średnica tłoczków/cylinderków 2	34 mm	Pomiar własny wykonany z wykorzystaniem suwmiarki lub wagi elektronicznej
Średnica tłoczków/cylinderków 3	30 mm	
Wysokość tłoczków 2 i 3	20 mm	
Głębokość cylinderków 2 i 3	21 mm	
Wymiary klocków hamulcowych	75 × 55 mm	
Rozstaw otworów pod klocki hamulcowe	51 mm	
Masa połówki A	323 g	
Masa połówki B	244 g	
Masa śrub łączących połówki A i B	4 × 17 g	
Masa całkowita korpusu zacisku	635 g	Instrukcja serwisowa (Moto Guzzi)
Średnica tarczy hamulcowej	320 mm	



Rys. 3. Mikrostruktura stopu, zgląd metalograficzny nietrawiony, mikroskop optyczny

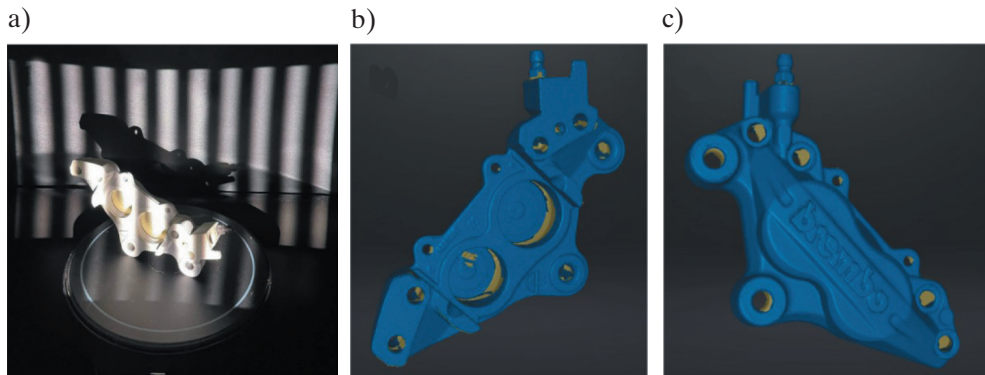
2.2. Skan 3D zacisku hamulcowego i cyfrowy bliźniak

Dążąc do stworzenia parametrycznego modelu CAD pierwotnego zacisku, posłu-
żono się inżynierią odwrotną. W celu uzupełnienia informacji o geometrii zacisku
i ułatwienia późniejszego procesu modelowania CAD zastosowano skaner 3D.

Skan 3D pozwolił na zebranie szczegółowych danych wymiarowych zacisku ha-
mulcowego w formie geometrycznej siatki. Dokładność wymiarowa wzrosła znacząco
względem możliwości oferowanych przez pomiary z wykorzystaniem ręcznych przy-
rządów. Zastosowane w projekcie urządzenie skanujące firmy Shining 3D, model
EinScan-SE V2, działa z wykorzystaniem technologii światła strukturalnego (rys. 4a).

Uzyskanie pełnego modelu zacisku hamulcowego wymagało wykonania skanów
obu połówek korpusu osobno. Dla każdej z nich proces przeprowadzono trzykrotnie,
zmieniając przy tym ułożenie połówki korpusu względem ruchomego stołu. Efektem

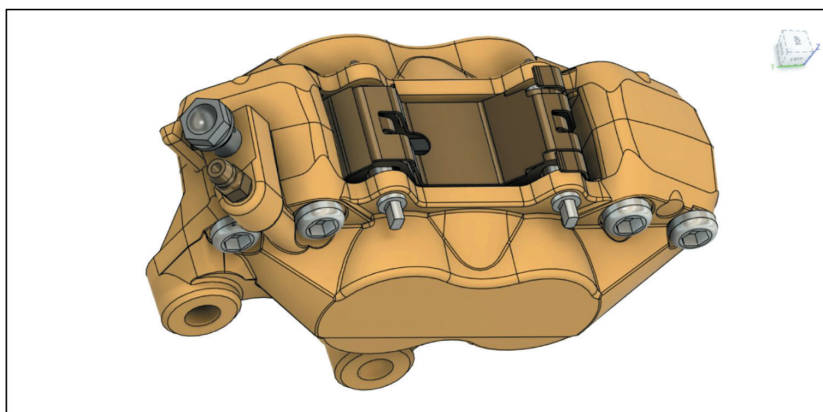
procesu skanowania była chmura punktów, z której w programie EinScan utworzono model powierzchni skanowanego elementu (rys. 4b i 4c). Naprawa i łatanie siatki skanu umożliwiły wygenerowanie zamkniętej bryły powierzchniowej i eksport w formacie STL (.stl) do programu CAD.



Rys. 4. Skanowana połówka A: a) w trakcie procesu skanowania;
b) i c) w interfejsie programu EinScan

Dalsze etapy pracy oparte na modelowaniu geometrii zacisku pierwotnego oraz własnego rozwiązania zostały przeprowadzone z wykorzystaniem programu Fusion 360 firmy Autodesk. Program pozwala na import zeskanowanej geometrii w formacie STL jako bryły typu *mesh*.

Zaimportowane skany po wzajemnym dopasowaniu układów współrzędnych posłużyły jako podstawa do utworzenia parametrycznych modeli połówek korpusu zacisku. Pozostałe elementy złożenia (rys. 5) narysowano na podstawie pomiarów rzeczywistych elementów. Powstałe złożenie jest kompletnym źródłem danych geometrycznych służących do stworzenia własnego prototypu rozwiązania korpusu zacisku hamulcowego.

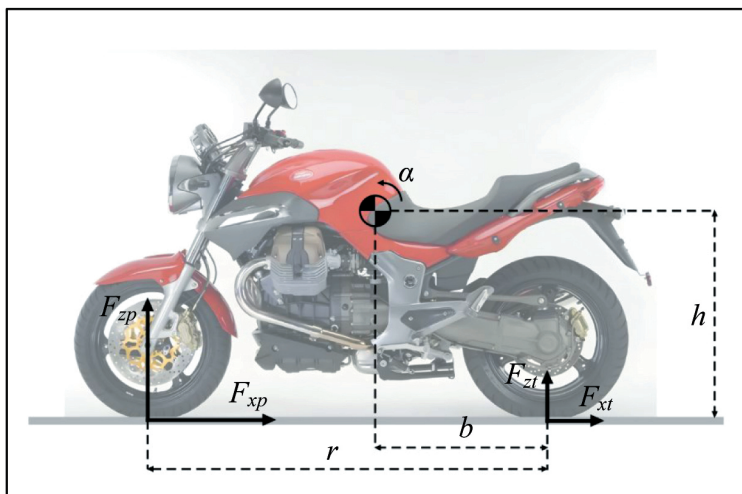


Rys. 5. Model CAD pierwotnego zacisku hamulcowego – złożenie

3. Wyznaczenie obciążeń i dobór materiału do produkcji zacisku

3.1. Dynamika procesu hamowania

Zachowanie jednośladów podczas hamowania jest wyjątkowe na tle innych pojazdów. Niska masa, krótki rozstaw osi i jednocześnie stosunkowo wysoki środek ciężkości sprawiają, że technika hamowania motocyklem znacząco różni się od techniki stosowanej w samochodach (Cossalter i in. 2004). Analiza tego procesu wymaga uwzględnienia wielu czynników, takich jak rozkład sił pomiędzy przednim a tylnym kołem czy efekt przeniesienia masy i kompresji zawieszenia (Croccolo i in. 2013). Aby wyznaczyć maksymalną siłę hamowania, można posłużyć się układem sił działających na motocykl i poszczególne koła (rys. 6) (Corno i in. 2008).



Rys. 6. Schemat rozkładu sił podczas hamowania
naniesiony na wybrany motocykl (fot. instrukcja serwisowa)

Źródło: Corno i in. (2008)

W tym przypadku układ sił został uproszczony przez pominięcie zjawiska kompresji zawieszenia oraz aerodynamiki, a zaznaczony środek ciężkości odpowiada warunkom rzeczywistego hamowania, tj. wraz z kierującym. Oznaczenie p odnosi się do przedniego, a t do tylnego koła. Układ równań równowagi prezentuje się następująco (Corno i in. 2008):

$$\begin{cases} ma_x = -F_x \\ ma_z = F_{zt} + F_{zp} - mg \\ I\alpha = F_{zt}b - F_{zp}(r-b) + F_xh \end{cases} \quad (1)$$

gdzie:

- r – rozstaw kół,
- b – odległość między tylną osią a środkiem ciężkości w osi x ,
- h – wysokość środka ciężkości względem podłoża,
- m – masa motocykla razem z kierowcą,
- I – moment bezwładności,
- a – przyspieszenie w danej osi,
- α – przyspieszenie kątowe,
- g – przyspieszenie grawitacyjne,
- F – siła działająca w danej osi i punkcie.

Silne hamowanie w przypadku motocykla odbywa się przy użyciu wyłącznie przedniego hamulca. Zasadniczo szacowanie sił polega na uwzględnieniu najtrudniejszego warunku hamowania, mianowicie sytuacji, przy której obciążenie na przednim kole osiąga najwyższą wartość, obciążenie na tylnym kole maleje do zera, a tylna oś znajduje się na granicy oderwania od ziemi. Dla powyższego modelu prezentują to równania:

$$a_z = 0; \quad F_{zt} = 0; \quad \alpha = 0 \quad (2)$$

Układ równań (1) upraszcza się do postaci:

$$F_x = mg \frac{r-b}{h} \quad (3)$$

Oznacza to, że maksymalna możliwa siła hamowania – w sytuacji, gdy przyczepność przedniej opony jest na tyle duża, że może dojść do oderwania tylnego koła od podłoża – zależy od masy kierowcy, motocykla, położenia ich wspólnego środka ciężkości oraz rozstawu osi pojazdu.

Maksymalna siła hamowania odpowiada sile tarcia generowanej przez oba przednie hamulce. Opisuje ją zależność siły nacisku klocków hamulcowych na tarczę i ich wzajemnego współczynnika tarcia, wyrażona wzorem:

$$f = 2\mu N \quad (4)$$

gdzie:

- f – siła tarcia,
- μ – współczynnik tarcia,
- N – siła nacisku klocków hamulcowych.

Należy zauważyć, że w przypadku zastosowania dwóch tarcz i zacisków na przednim kole każdy z hamulców generuje połowę siły hamowania. Ostateczną siłę nacisku, jaką wywierają klocki hamulcowe, określa równanie:

$$N = \frac{mg(r-b)}{2\mu h} \quad (5)$$

Taka sama siła co do wartości, ale o przeciwnym zwrocie, działa na zacisk hamulcowy. Powstające w jej wyniku naprężenia zginające są podstawą do wyznaczania wytrzymałości zacisku i projektowania jego geometrii. Należy pamiętać, że mają one charakter cykliczny, jednostronny (odzerowo-tętniący) o zmiennej amplitudzie i częstotliwości w zależności od siły i momentu hamowania. Wzór (5) stanowi podstawę do określenia reakcji siły rozwierającej zacisk w trakcie hamowania, której poznanie jest konieczne do przeprowadzania symulacji wytrzymałościowej. Siła rozwierająca jest wynikiem parcia płynu hamulcowego na powierzchnię tłoczków. Ciśnienie to wyraża się wzorem:

$$p = \frac{2N}{\pi(d_2^2 + d_3^2)} \quad (6)$$

gdzie d_2 i d_3 to średnice par tłoczków 2 i 3.

Oprócz działania siły rozwierającej należy uwzględnić również siłę reakcji wynikającą z tarcia klocków hamulcowych o tarczę. Klocki, zamocowane sztywno w zacisku, przekazują tę siłę na mocowanie zacisku do konstrukcji motocykla. Odpowiada ona połowie (ze względu na zastosowanie dwóch hamulców przedniego koła) maksymalnej siły hamowania. Zależność tę przedstawia równanie:

$$f_{\frac{1}{2}} = \mu N \quad (7)$$

Dane do obliczeń, przedstawione w tabeli 2, bazują na rzeczywistych wartościach dotyczących motocykla Moto Guzzi Breda V 1100, a w przypadku ich braku przyjęto założenia na podstawie ogólnych informacji charakterystycznych dla tego typu motocykli.

Symbol m w równaniu (5) oznacza sumę masy (8) kierującego (m_k), przyjętej jako 80 kg, oraz motocykla (m_m):

$$m = m_m + m_k \quad (8)$$

Dane dotyczące położenia środka ciężkości dla wybranego modelu motocykla nie są ogólnie dostępne, dlatego w obliczeniach posłużono się wartościami uogólnionymi, charakterystycznymi dla motocykli tej klasy.

Tabela 2

Dane do obliczeń sił działających podczas manewru gwałtownego hamowania

Parametr	Wartość	Źródło danych
Masa motocykla m_m	248 kg	instrukcja serwisowa
Rozstaw kół r	1,495 m	
Masa kierującego m_k	80 kg	założenie
Przyspieszenie ziemskie g	9,81	
Rozkład masy z kierującym P/T	49,5/50,5	założenie (Cameron 2021)
Wysokość środka ciężkości motocykla h_m	0,6 m	
Wysokość środka ciężkości kierującego h_k	0,9 m	założenie
Współczynnik tarcia μ	0,4	
Średnica pary tłoczków 2 d_2	34 mm	pomiar własny
Średnica pary tłoczków 3 d_3	30 mm	

Położenie środka ciężkości motocykla jest bezpośrednio powiązane z rozłożeniem masy na przednie i tylne koło. W przypadku motocykli sportowych rozłożenie masy zwykle przyjmuje stosunek 52 do 48 z większą częścią masy na przednim kole (Cameron 2021). W motocyklach turystycznych środek ciężkości przesunięty jest bardziej do tyłu. Dla większości obliczeń wystarczająco dokładnym przybliżeniem jest rozłożenie masy w proporcji 50 do 50. Po uwzględnieniu masy kierującego przyjęto rozłożenie masy 49,5 do 50,5 (tab. 2) z większą częścią masy na tylnym kole. Odległość b środka ciężkości od tylnej osi opisuje równanie:

$$b = r \cdot \frac{P}{100} \quad (9)$$

gdzie P oznacza część masy na przednim kole wyrażoną w procentach ($P = 49,5\%$).

Różne źródła w niejednoznaczny sposób określają wysokość środka ciężkości motocykla. Jego położenie może mieścić się w zakresie od około 50 cm (mierzone od podłoża), do około 75 cm, przy czym wartość tę często przyjmuje się jako połowę rozstawu osi. Ze względu na brak dokładnych informacji przyjęto wysokość środka ciężkości wynoszącą 60 cm od podłoża. Położenie środka ciężkości kierującego określono w przybliżeniu na wysokości bioder, co po naniesieniu geometrii na profil boczny motocykla daje około 90 cm nad podłożem. Wysokość wspólnego środka ciężkości opisuje równanie:

$$h = \frac{m_m \cdot h_m + m_k \cdot h_k}{m_m + m_k} \quad (10)$$

gdzie h_m i h_k to odpowiednio wysokość środka ciężkości motocykla i kierującego.

Współczynnik tarcia dla hamulców tarczowych różni się w zależności od producenta klocków hamulcowych oraz materiałów zastosowanych do ich produkcji. W klockach wykorzystywanych w pojazdach drogowych wynosi on zwykle 0,3–0,4, natomiast w przypadku klocków o przeznaczeniu wyścigowym – 0,4–0,5. Przy realizacji prezentowanego projektu przyjęto wartość 0,4 (tab. 2). Podstawiając dane do równań (8)–(10), otrzymujemy:

$$m = 248 + 80 = 324 \text{ kg} \quad (11)$$

$$b = 1,495 \cdot \frac{49,5}{100} \approx 0,74 \text{ m} \quad (12)$$

$$h = \frac{248 \cdot 0,6 + 80 \cdot 0,9}{248 + 80} \approx 0,68 \text{ m} \quad (13)$$

Po podstawieniu obliczonych wartości do równania (5) uzyskujemy:

$$N = \frac{324 \cdot 9,81(1,495 - 0,74)}{2 \cdot 0,4 \cdot 0,68} \approx 4411 \text{ N} \quad (14)$$

Wartość ciśnienia rozwierającego zacisk i siły tarcia klocków obliczono na podstawie równań (6) i (7):

$$p = \frac{2 \cdot 4411}{3,14(34^2 + 30^2)} = 1,37 \text{ MPa} = 13,7 \text{ bar} = 198,7 \text{ psi} \quad (15)$$

$$f_{1/2} = 0,4 \cdot 4411 = 1764 \text{ N} \quad (16)$$

3.2. Materiał do produkcji zacisku

Aby umożliwić wykonanie zacisku o wysokim stopniu złożoności geometrycznej, zdecydowano się na jego wytworzenie z zastosowaniem technologii przyrostowej. Spośród technik druku z wykorzystaniem materiałów metalowych wybrano selektywne stapianie laserowe, które zapewnia pełne przetopienie warstwy materiału, a tym samym zachowanie ciągłości struktury metalicznej (Liu i in. 2019, Nguyen i in. 2022). Poszukując materiału odpowiedniego do tej technologii, charakteryzującego się wysoką wytrzymałością, odpornością na podwyższone temperatury, zmęczenie oraz korozję, a jednocześnie cechującego się możliwie małą gęstością, wybrano stop Ti6Al4V. Dane materiałowe, takie jak gęstość, moduł Younga, wytrzymałość na rozciąganie, granica plastyczności oraz przewodność cieplna (tab. 3), zaczerpnięto z biblioteki materiałowej będącej częścią Autodesk Fusion 360 dla materiału Ti6Al4V przeznaczonego do produkcji przyrostowej.

Tabela 3
Charakterystyka stopu Ti6Al4V wykorzystanego w dalszych symulacjach
(Hrabe i in. 2017)

Cecha materiałowa	Wartość
Gęstość	4,43 g/cm ²
Przewodność cieplna	6,70 W/mm·°C
Moduł Younga	113,76 GPa
Wytrzymałość na rozciąganie	1034,21 MPa
Granica plastyczności przy 20°C	882,52 MPa
Granica plastyczności przy 200°C	około 750 MPa
Wytrzymałość zmęczeniowa	550–650 MPa przy 10 ⁷ cykli

Należy zaznaczyć, że naprężenia powstające w zacisku hamulcowym mają charakter zmienny, z tego względu istotne jest uwzględnienie w rozważaniach wytrzymałościowych właściwości zmęczeniowych (tab. 3). Wytrzymałość zmęczeniową stopu Ti6Al4V uzyskanego metodami przyrostowymi można znacząco poprawić z wykorzystaniem obróbki cieplnej. Na przykład proces HIP (ang. *hot isostatic pressing*) polegający na przeprowadzaniu obróbki cieplnej pod wpływem ciśnienia obojętnego gazu pozwala poprawić wytrzymałość zmęczeniową z 200–250 MPa dla 10⁷ cykli do nawet 550–600 MPa (Hrabe i in. 2017).

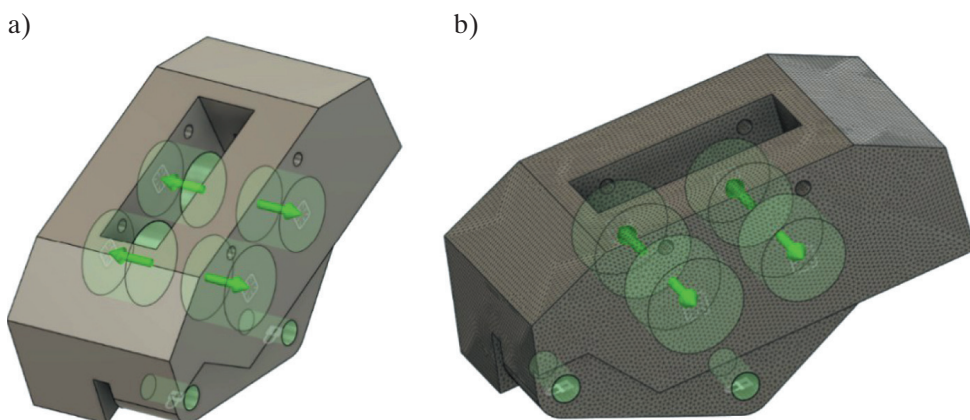
Istotnym aspektem pod względem wytrzymałości badanego elementu jest temperatura pracy. Temperatura wrzenia współczesnych płynów hamulcowych, takich jak np. DOT 5.1, wynosi 260°C. Osiągnięcie temperatury wrzenia płynu całkowicie uniemożliwia działanie układu hamulcowego. W dalszych rozważaniach przyjęto temperaturę pracy płynu i zacisku na poziomie 200°C. Przy przyjętej temperaturze 200°C granica plastyczności stopu Ti6Al4V spada o mniej więcej 15% względem wartości w temperaturze pokojowej (tj. 882,52 MPa – tab. 3).

4. Optimalizacja geometrii zacisku z wykorzystaniem symulacji MES

Proces optymalizacji geometrii zacisku przeprowadzono z wykorzystaniem programu Fusion 360 firmy Autodesk. Pozwoliło to na utworzenie geometrii w module Design i implementację ich do modułu Simulation w celu przygotowania optymalizacji

opartej na metodzie SIMP (ang. *solid isotropic material with penalization* – lity materiał izotropowy z penalizacją). Mimo stosunkowo uproszczonego względem konkurencyjnego oprogramowania interfejsu symulacyjnego program umożliwił osiągnięcie wysokiej szczegółowości zadanych siatek MES symulowanych geometrii.

Pierwszym krokiem było przeprowadzenie ogólnej optymalizacji topologicznej zacisku typu monoblock pod kątem wytrzymałości i zachowania sztywności pod wpływem działania maksymalnej siły rozwierającej. W tym celu utworzono uproszczoną bryłę (rys. 7a), zachowując wymiary konieczne do zapewnienia kompatybilności zacisku, tj. wielkość i rozmieszczenie cylinderek oraz otworów mocujących zacisk do motocykla, przestrzeń pod tarczę i klocki hamulcowe oraz otwory mocujące klocki. W module Simulation zadano materiał, tj. stop Ti6Al4V z biblioteki materiałowej Autodesk (tab. 3), unieruchomiono punkty mocowania zacisku i zadano ciśnienie wywierane na powierzchnię płaską cylinderek o wartości 1,37 MPa (rys. 7a). Wygenerowano siatkę (rys. 7b).

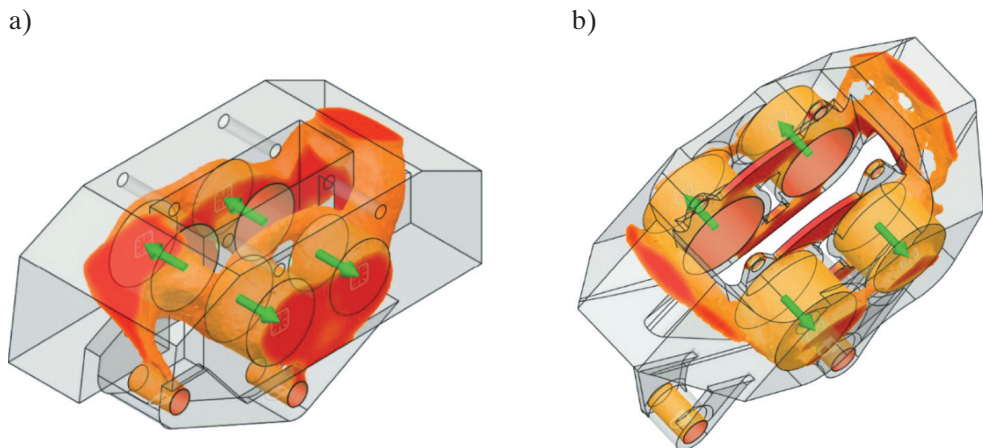


Rys. 7. Uproszczona bryła: a) po zadaniu obciążeń i warunków brzegowych;
b) po wygenerowaniu siatki

Wynikiem procesu optymalizacji topologicznej było uzyskanie interaktywnej geometrii, dla której za pomocą suwaka wraz ze skalą kolorystyczną można było regulować ilość pozostawionego nadmiaru materiału. Rysunek 8a przedstawia bryłę przy zachowaniu 18,82% materiału względem modelu uproszczonego (rys. 7).

Wyniki tej symulacji pozwoliły na dalsze definiowanie geometrii projektowanego zacisku polegające na wyznaczaniu kluczowych z punktu widzenia sztywności i wytrzymałości obszarów. Symulacja zwróciła również uwagę na aspekty niebędące częścią założeń symulacyjnych, takie jak estetyka, zmniejszona ochrona zacisku przed czynnikami

zewnątrznymi oraz odprowadzanie ciepła. Kolejnym krokiem było zaprojektowanie bryły zacisku o geometrii zbliżonej do wyniku pierwszej optymalizacji topologicznej i powtórzenie wyżej opisanego procesu. Wyniki przedstawia rysunek 8b.



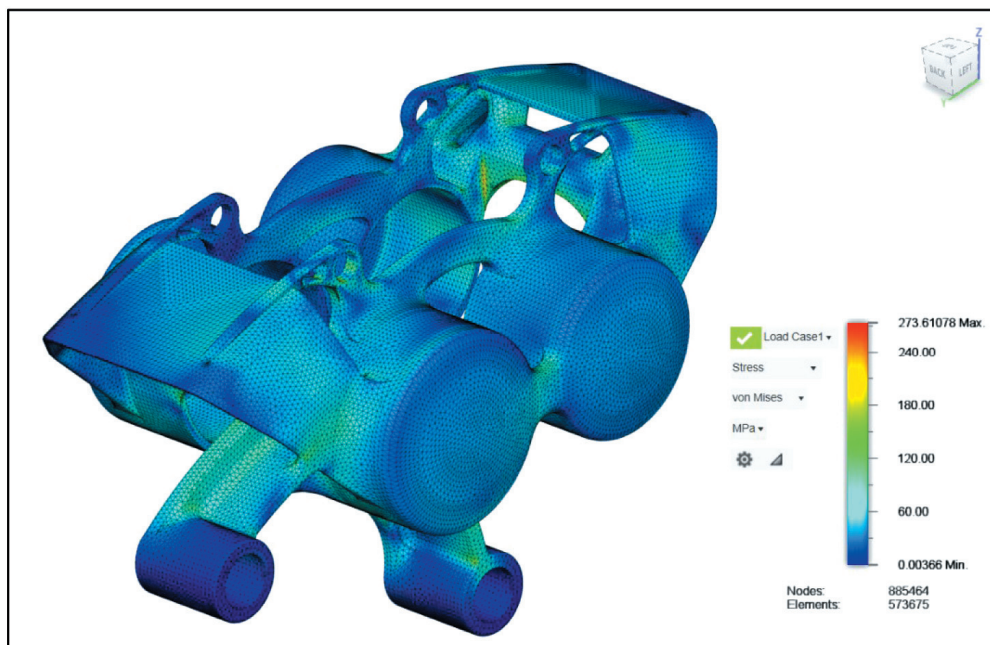
Rys. 8. Wyniki procesu optymalizacji topologicznej: a) pierwsza iteracja z zachowaniem 400 g materiału; b) druga iteracja z zachowaniem 105 g materiału

Drugi proces optymalizacji pozwolił na skrajne ograniczenie masy elementu do około 106 g. Należy zauważyć, że poszczególne części modelu uległy rozłączeniu w wyniku tak znacznego usunięcia materiału. Z tego względu wygenerowany model posłużył jako wyznacznik do stworzenia nowego prototypu (rys. 9). Został on narysowany od początku według założeń projektowych, dążących do finalnej redukcji masy. W ramach dalszych symulacji uwzględniono reakcję siły tarcia działającą na mocowanie zacisku, zadaną na powierzchni oparcia się klocków hamulcowych o korpus zacisku.

Wyniki najnowszej symulacji wytrzymałościowej wskazują, że maksymalne naprężenie równoważne wynosi 273,6 MPa (rys. 9). Odpowiada to współczynnikowi bezpieczeństwa wynoszącemu 3,23 przy temperaturze 20°C oraz 2,74 – przy 200°C. Dodatkowo (także przy uwzględnieniu 15-procentowego osłabienia konstrukcji w temperaturze 200°C) wartości naprężeń pozostają znacząco poniżej granicy wytrzymałości zmęczenia, wynoszącej 550–600 MPa. Masa korpusu o takiej geometrii, wykonanego ze stopu Ti6Al4V, wynosi 248 g. Oznacza to, że udało się uzyskać jej redukcję do 39% masy pierwotnego rozwiązania, a więc zmniejszyła się o 387 g. Jeśli uwzględni się oba przednie zaciski, to całkowita redukcja masy wynosi 774 g.

Finalny prototyp pomija zagadnienie projektowania kanałów pod płyn hamulcowy, ich wymiarów oraz rozmieszczenia. Konieczne może okazać się wzmocnienie konstrukcji w miejscach doprowadzania kanałów oraz podłączenia przewodów hamul-

cowych. Dodatkowym krokiem wynikającym z charakterystyki technologii przyrostowych (nakładania materiału warstwowo) jest określenie naddatków, metody obróbki powierzchniowej i sposobu wykończenia finalnego produktu. Aby zagwarantować zakładaną wytrzymałość zmęczeniową, konieczne jest wygładzenie powierzchni w celu wyeliminowania ryzyka wystąpienia naturalnych karbów spowodowanych warstwową technologią produkcji.



Rys. 9. Wynik symulacji naprężeń dla finalnego prototypu z uwzględnieniem złożonego układu sił; skala kolorystyczna odpowiada naprężeniu w materiale

Model CAD prototypu spełnia teoretyczne wymagania wytrzymałościowe. Aby sprawdzić, czy konstrukcja jest możliwa do złożenia i zamontowania na motocyklu, należy wykonać model fizyczny, który nie musi posiadać pełnych właściwości mechanicznych. Z tego względu dopuszczalne jest wykonanie prototypu z tworzywa sztucznego w technologii druku FDM (ang. *fused deposition modelling*). Ze względu na inną charakterystykę druku w technologii FDM (większe ograniczenia w ułożeniu wydruku na stole roboczym drukarki niż w SLM) oraz brak walorów funkcjonalnych prototypu z tworzywa zdecydowano się na podzielenie modelu CAD na cztery części, aby zapewnić wyższą jakość wydruku. Wydruk wykonano z tworzywa PLA (poliaktyd) za pomocą drukarki Bambu Lab A1 (rys. 10a). Wydrukowane elementy zostały oczyszczone z materiału podporowego i sklejone w jedną całość (rys. 10b).

a)



b)



Rys. 10. Rzeczywisty prototyp geometrii korpusu zacisku hamulcowego wykonany z tworzywa sztucznego (PLA) w technologii druku FDM:
a) w trakcie druku; b) gotowy element

5. Podsumowanie

W ramach prezentowanego projektu udało się stworzyć prototypowe rozwiązanie zacisku hamulcowego przeznaczonego do motocykla sportowego. Spełnione zostało główne założenie zaprojektowania lekkiej konstrukcji o wysokiej wytrzymałości. Przedstawiony prototyp waży zaledwie 39% masy tradycyjnego rozwiązania, jednocześnie cechują go wysokie współczynniki bezpieczeństwa (3,2–2,7) dla maksymalnych założonych obciążeń.

Na podstawie istniejącego rozwiązania zachowano jednakowy wymiar kluczowych elementów, takich jak tłoczki, klocki hamulcowe czy mocowanie zacisku, co zapewnia kompatybilność z motocyklem. Zmiana sposobu produkcji na technologie przyrostowe pozwoliła poprawić geometrię zacisku – dwie połówki zastąpiono jednym elementem (zacisk typu monoblock).

Sposób mocowania zacisku do motocykla to element, który w największym stopniu wymaga poprawek. Dalsza redukcja masy przy zachowaniu jednakowej wytrzymałości jest możliwa przy zastosowaniu radialnych zacisków hamulcowych. Taki eksper-

ment wykracza jednak poza założenia tej pracy ze względu na ingerencję w konstrukcję motocykla poza korpusem zacisku hamulcowego.

Zastosowanie stopu Ti6Al4V do wykonania elementów konstrukcyjnych pozwala na uzyskanie wysokich właściwości wytrzymałościowych i/lub niskiej wagi. Jednocześnie produkcja metodami przyrostowymi oraz wysokie koszty ograniczają możliwości wprowadzenia tej technologii jedynie do luksusowych/sportowych pojazdów wytwarzanych w małych seriach lub nawet jednostkowo.

Narzędzia optymalizacyjne i symulacyjne, których działanie opiera się na metodzie elementów skończonych, rewolucjonizują podejście do projektowania, a ich rozwój umożliwia wprowadzanie zaawansowanych rozwiązań technicznych, takich jak prototyp przedstawiony w prezentowanej pracy.

Możliwość wykonania i praktycznego zastosowania zaprezentowanego w pracy prototypu zależy od dokładności poczynionych założeń. W przypadku projektowania faktycznego produktu konieczne jest przeprowadzenie szerszych badań w celu pozyskania rzeczywistych danych do wszystkich aspektów projektu, które nie mogły być zrealizowane w ramach tej pracy. Należy ją traktować jako teoretyczny przykład postępowania projektowego przy tworzeniu ultralekkich części konstrukcyjnych dla branży motoryzacyjnej (*automotive*). Przedstawiony proces można zaadaptować nie tylko do różnych typów zacisków hamulcowych, lecz także do szerszej grupy podzespołów motocyklowych i samochodowych.

Literatura

- Cameron K., 2021, *How Motorcycles Carry Their Weight*, <https://www.cycleworld.com/story/blogs/ask-kevin/how-motorcycles-carry-their-weight/> [dostęp: 22.12.2024].
- Corno M., Savaresi S.M., Tanelli M., Fabbri L., 2008, *On optimal motorcycle braking*, Control Engineering Practice, vol. 16, iss. 6, pp. 644–657. <https://doi.org/10.1016/j.conengprac.2007.08.001>.
- Cossalter V., Lot R., Maggio F., 2004, *On the Braking Behaviour of Motorcycles*, SAE Technical Papers. <https://doi.org/10.4271/2004-32-0018>.
- Croccolo D., Agostinis M., Olmi G., Tizzanini A., 2013, *Analysis of the Stress State in Brake Caliper Mounts of Front Motorbike Suspensions*, Advances in Mechanical Engineering, vol. 5. <https://doi.org/10.1155/2013/525010>.
- Herbe N., Gnäupel-Herold T., Quinn T., 2017, *Fatigue Properties of a Titanium Alloy (Ti-6Al-4V) Fabricated Via Electron Beam Melting (EBM): Effects of Internal Defects and residual Stress*, International Journal of Fatigue, vol. 94, iss. 2, pp. 202–210. <https://doi.org/10.1016/j.ijfatigue.2016.04.022>.

- Liu S., Shin Yung C., 2019, *Additive manufacturing of Ti6Al4V alloy: A review*, Materials & Design, vol. 164. <https://doi.org/10.1016/j.matdes.2018.107552>.
- Moto Guzzi, *Service Station Manual Breva V1100*. [https://guzzitek.org/gb/ma_us_uk/1100/Breva1100_052007_Atel\(GB\).pdf](https://guzzitek.org/gb/ma_us_uk/1100/Breva1100_052007_Atel(GB).pdf) [dostęp: 20.12.2024].
- Nguyen H.D., Pramanik A., Basak A.K., Dong Y., Prakash C., Debnath S., Shankar S., Jawahir I.S., Dixit S., Buddhi D., 2022, *A critical review on additive manufacturing of Ti-6Al-4V alloy: microstructure and mechanical properties*, Journal of Materials Research and Technology, vol. 18, pp. 4641–4661. <https://doi.org/10.1016/j.jmrt.2022.04.055>.

Pozyskiwanie i analiza danych o cyberatakach z wykorzystaniem systemu honeypot

Miłosz Gaszyna 

Akademia Górniczo-Hutnicza, Wydział Informatyki, Elektroniki i Telekomunikacji, Kraków

Streszczenie: Niniejszy artykuł koncentruje się na analizie złośliwej aktywności w sieci z wykorzystaniem środowiska honeypotowego T-Pot. Celem badania było przechwycenie i zrozumienie rzeczywistych działań cyberprzestępców, ich technik, wektorów ataku oraz wykorzystywanego oprogramowania. Eksperyment umożliwił zebranie danych z ponad 160 krajów, identyfikację najczęściej atakowanych usług, przechwycenie próbek złośliwego oprogramowania oraz analizę ataków typu brute force i słownikowych. Szczególny nacisk położono na poznanie charakterystyki generowanego złośliwego ruchu i przeanalizowanie go w sposób umożliwiający wyciągnięcie praktycznych wniosków oraz sformułowanie zaleceń z zakresu cyberbezpieczeństwa.

Słowa kluczowe: cyberbezpieczeństwo, honeypot, T-Pot, złośliwe oprogramowanie, botnet, SMB, SSH, automatyzacja ataków, analiza ruchu sieciowego, ataki brute force, ataki słownikowe

ACQUISITION AND ANALYSIS OF CYBERATTACK DATA USING A HONEYPOT SYSTEM

Abstract: This article focuses on the analysis of malicious network activity using the T-Pot honeypot environment. The aim of the study was to capture and understand real-world actions of cybercriminals, their techniques, attack vectors, and the software they use. The experiment enabled the collection of data from over 160 countries, the identification of the most frequently targeted services, the capture of malware samples, and the analysis of brute force and dictionary attacks. Emphasis was placed on examining the characteristics of generated malicious traffic and analyzing it in a way that allows for drawing practical conclusions and formulating cybersecurity recommendations.

Keywords: cybersecurity, honeypot, T-Pot, malware, botnet, SMB, SSH, attack automation, net work traffic analysis, brute force attacks, dictionary attacks

https://doi.org/10.7494/978-83-68219-80-7_2

1. Wprowadzenie

Współczesny świat w coraz większym stopniu opiera się na technologiach cyfrowych. Praktycznie każda dziedzina życia – od gospodarki, przez administrację publiczną, aż po codzienne funkcjonowanie jednostek – ulega cyfryzacji. Ten dynamiczny rozwój branży IT ma jednak swoją ciemną stronę, równoległe bowiem i równie szybko rośnie skala oraz złożoność cyberprzestępczości. Cyberbezpieczeństwo stało się jednym z kluczowych wyzwań globalnych. Zarówno sektor prywatny, jak i publiczny nieustannie mierzą się z coraz bardziej zaawansowanymi technologicznie atakami, co nieprzerwanie podnosi poziom trudności ich wykrywania i zapobiegania im. Skala tego zjawiska jest ogromna – w samym 2024 roku straty wywołane cyberprzestępczością szacowane były na około 9,5 biliona dolarów (Experis 2025), co stanowi niemal 10% światowego PKB. Dla zobrazowania: gdyby cyberprzestępczość była państwem, jej „gospodarka” plasowałaby się na trzecim miejscu na świecie, ustępując jedynie Stanom Zjednoczonym i Chinom.

Obecne trendy wskazują, że problem będzie jedynie narastać. Postęp w zakresie sztucznej inteligencji, automatyzacji działań ofensywnych, technik typu deepfake czy klonowania głosu w celu realizowania wyrafinowanego vishingu, zapowiadają nową erę zagrożeń. Cyberataki coraz rzadziej są masowe i przewidywalne, a stają się bardziej spersonalizowane, szybkie i trudne do wykrycia. W tym kontekście niezwykle istotne staje się nie tylko reagowanie na incydenty, ale przede wszystkim ich wczesne wykrywanie, analiza i rozumienie mechanizmów działania atakujących. Jednym z narzędzi wspierających te działania są systemy typu honeypot, które pozwalają na symulowanie podatnych usług i środowisk w celu monitorowania i analizowania aktywności napastników bez narażania kluczowej infrastruktury na niebezpieczeństwo. Dzięki nim możliwe jest nie tylko identyfikowanie nowych zagrożeń w czasie rzeczywistym, ale również doskonalenie mechanizmów obronnych oraz budowanie świadomości zagrożeń wśród specjalistów i organizacji.

Celem niniejszej pracy było przeprowadzenie realistycznego eksperymentu badawczego z wykorzystaniem systemu honeypot, który umożliwił zarejestrowanie, zidentyfikowanie i przeanalizowanie rzeczywistych prób cyberataków. Głównym założeniem było nie tylko wykrycie aktualnych zagrożeń, lecz także głębsze zrozumienie mechanizmów działania atakujących dzięki ocenie stosowanych technik i schematów postępowania. Układ artykułu został zaplanowany w sposób umożliwiający stopniowe zagłębianie się w tematykę oraz logiczne przeprowadzenie czytelnika przez wszystkie etapy badania. Po krótkim przeglądzie literatury przedstawiono środowisko badawcze oraz zastosowaną metodologię. Kolejny punkt poświęcono procesowi pozyskiwania danych oraz ich przygotowaniu do analizy. Najobszerniejsza część pracy zawiera szczegółową analizę zebranych informacji – rozpoczynając od geolokalizacji źródeł ataków, przez

analizę poszczególnych usług, ze szczególnym uwzględnieniem SMB i SSH, aż po ocenę ataków brute force i słownikowych. Następnie na podstawie zgromadzonej wiedzy scharakteryzowano ogólny profil aktywności atakujących. Całość zamyka rozdział z wnioskami, w którym wskazano najważniejsze obserwacje, trendy oraz praktyczne rekomendacje w zakresie cyberbezpieczeństwa.

1.1. Przegląd literatury

Najnowsze badania potwierdzają, że honeypoty pozostają niezwykle wartościowym narzędziem w procesie pozyskiwania danych o zagrożeniach. Honeypoty dostarczają nie tylko danych o źródłach i wektorach ataków, ale umożliwiają również szczegółową analizę technik wykorzystywanych przez napastników, w tym komend systemowych, metod eskalacji uprawnień, używanych exploitów oraz rozpowszechnianych próbek złośliwego oprogramowania. Obecnie obserwuje się rosnące znaczenie integracji honeypotów z systemami analitycznymi opartymi na sztucznej inteligencji i uczeniu maszynowym, które umożliwiają automatyczną klasyfikację i detekcję podejrzanej aktywności na podstawie zebranych danych. W niektórych implementacjach skuteczność wykrywania zagrożeń przekracza 99% (Firmansyah i Zahra 2023), co czyni honeypoty istotnym źródłem danych treningowych dla nowoczesnych systemów przeciwdziałających cyberatakowi. W kontekście dynamicznie rozwijających się zagrożeń, ich zastosowanie rozszerza się również na środowiska chmurowe, internet rzeczy (IoT), a także przemysłowe systemy sterowania (ICS), gdzie stanowią unikalne uzupełnienie tradycyjnych metod detekcji i reagowania na incydenty.

1.2. Środowisko badawcze

Na potrzeby eksperymentu badawczego przygotowano dedykowane środowisko testowe w chmurze obliczeniowej Microsoft Azure, którego celem było jak najwierniejsze odwzorowanie rzeczywistego systemu produkcyjnego dostępnego z poziomu internetu. Uruchomiono maszynę wirtualną z systemem operacyjnym Ubuntu 24.04 LTS, wyposażoną w dwa rdzenie CPU, 8 GB pamięci RAM oraz 256 GB przestrzeni dyskowej. Maszyna posiadała publiczny adres IP, umożliwiający pełny dostęp do usług z zewnątrz i tym samym otwarcie jej na działania potencjalnych napastników (rys. 1).

Na etapie przygotowania środowiska zainstalowano platformy Docker oraz Docker Compose, stanowiące podstawę do uruchomienia frameworka T-Pot (Telekom Security 2024). W celu zachowania bezpieczeństwa systemu zarządzającego i uniknięcia konfliktu portów, domyślny port SSH (22), służący do zdalnego zarządzania maszyną, został przeniesiony na port o niestandardowym numerze, chroniony silnym certyfikatem i hasłem. Pozwoliło to zarezerwować port 22 dla honeypota emulującego SSH i jednocześnie ograniczyć ryzyko przejęcia dostępu do systemu. Środowisko zostało

również zabezpieczone przez restrykcyjne reguły zapory systemowej, a także przez konfigurację sieciowych grup bezpieczeństwa z poziomu platformy Microsoft Azure. Dostęp do wszystkich nieużywanych portów został jednoznacznie zablokowany, dzięki czemu możliwa była precyzyjna kontrola nad ruchem sieciowym i wykluczenie przypadkowych połączeń.

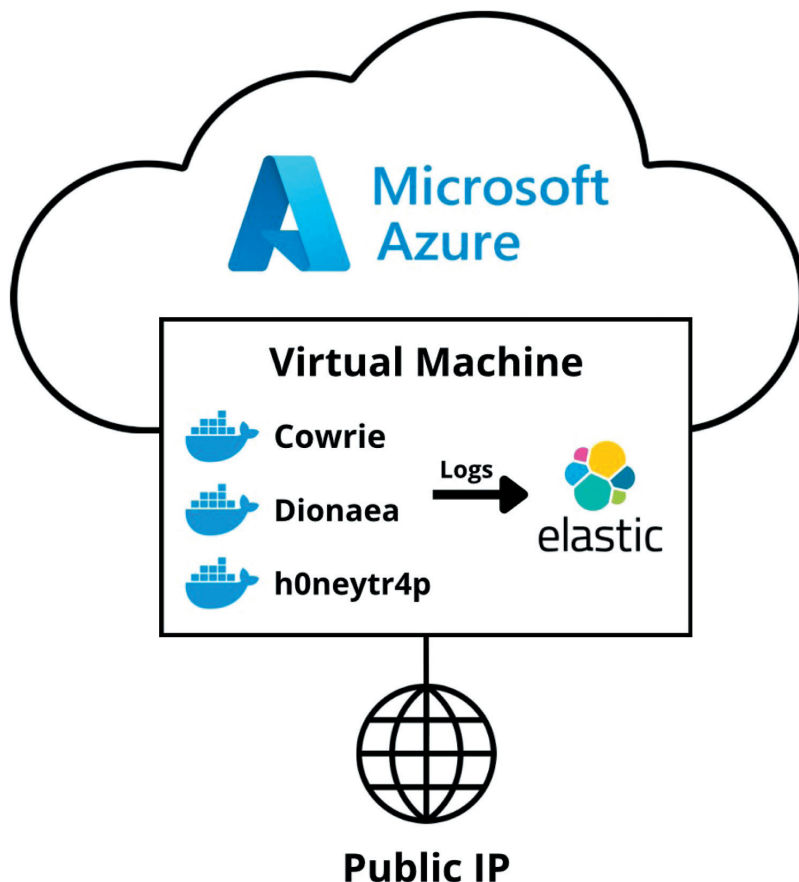
Zgodnie z założeniami badawczymi celem nie było stworzenie tzw. czarnej dziury, która bezrefleksyjnie przechwytuje cały przychodzący ruch, lecz odzwierciedlenie realistycznego środowiska produkcyjnego, w którym atakujący mają kontakt jedynie z rzeczywiście udostępnionymi usługami na konkretnych portach. W przygotowanym środowisku wdrożono framework T-Pot, będący zaawansowanym, wieloskładnikowym systemem typu honeypot, opracowanym przez Deutsche Telekom Security. Jego głównym celem jest kompleksowe pozyskiwanie, przetwarzanie i wizualizacja danych o zagrożeniach cybernetycznych w sposób skalowalny i zautomatyzowany. T-Pot wykorzystuje skonteneryzowaną infrastrukturę Dockera, umożliwiając uruchomienie każdego komponentu honeypotowego w osobnej, odizolowanej przestrzeni. Dzięki temu system charakteryzuje się dużą modularnością, łatwością aktualizacji oraz odpornością na błędy pojedynczych usług. Domyślna instalacja zawiera zestaw ponad 20 pułapek, między innymi:

- Cowrie – emulująca protokoły SSH oraz Telnet, rejestrujący próby brute force i eksploracji systemu,
- Dionaea – nastawiona na zbieranie złośliwego oprogramowania dostarczanego przez atakujących za pośrednictwem różnych protokołów (FTP, SMB, SQLServer),
- Glastopf – honeypot webowy wykrywający ataki na aplikacje internetowe,
- Conpot – symulująca infrastrukturę przemysłową ICS/SCADA,
- Elasticsearch – emulująca środowiska Elasticsearch.

Do analizy i wizualizacji pozyskanych danych wykorzystano narzędzia stosu ELK (Elasticsearch, Logstash, Kibana), które umożliwiają korelację zdarzeń oraz identyfikację wzorców zachowań. Dodatkowo T-Pot zintegrowany jest z narzędziem typu IDS/IPS Suricata, które umożliwia wykrywanie i klasyfikowanie anomalii w ruchu sieciowym. Zastosowano również narzędzia wspomagające pasywną identyfikację źródeł ruchu, tj. p0f i FATT. W ramach eksperymentu zdecydowano się na minimalistyczną, lecz reprezentatywną konfigurację, obejmującą trzy honeypoty:

- Cowrie – emulacja usług Telnet i SSH, rejestrująca próby logowania i interakcje z systemem,
- Dionaea – symulacja usług FTP, SMB oraz Microsoft SQL Server, umożliwiająca przechwytywanie malware i exploitów,
- h0neytr4p – pasywny honeypot, który odpowiada na próby połączenia do portów HTTP oraz HTTPS, rejestrując między innymi próby skanowania i fingerprintingu.

Dzięki zróżnicowaniu usług i protokołów środowisko umożliwiało rejestrację szerokiego spektrum aktywności atakujących – od klasycznych prób brute force, przez skanowanie portów i fingerprinting, aż po realne próby infekcji systemów.

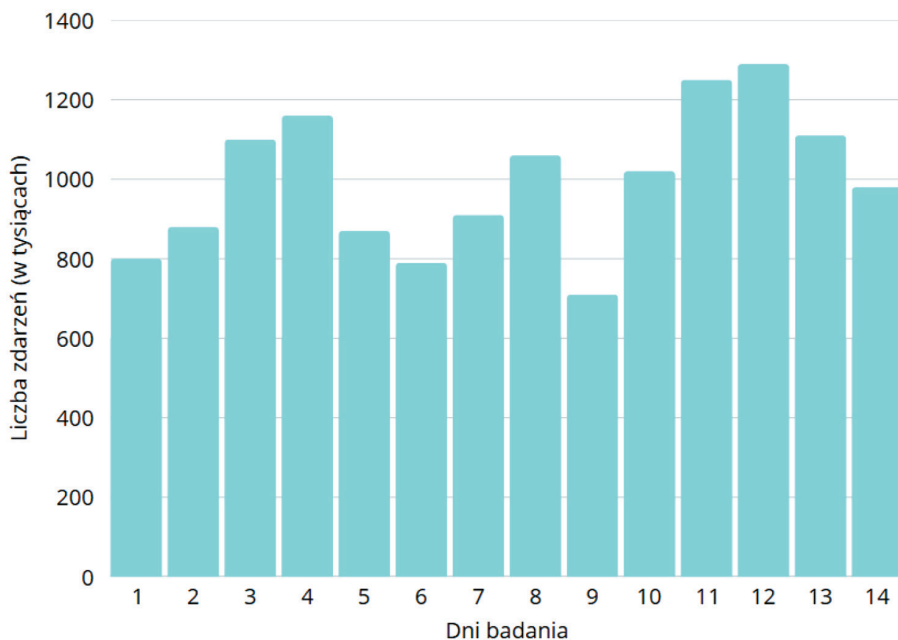


Rys. 1. Wizualizacja infrastruktury sieciowej

1.3. Pozyskanie danych

Eksperyment trwał nieprzerwanie przez 14 dni, a uruchomione honeypoty przez cały czas przechwytywały ruch sieciowy ukierunkowany na uprzednio wymienione usługi. Dzięki temu możliwe było zarejestrowanie pełnego spektrum ataków, zarówno masowych skanów, jak i ręcznych, starannie przygotowanych ingerencji. W trakcie działania honeypota zarejestrowano ponad 14 milionów logów, pochodzących z różnych źródeł. Po wstępnym, automatycznym przefiltrowaniu zdarzeń udało się wyodrębnić ponad 600 tysięcy przypadków, które można sklasyfikować jako rzeczywiste ataki. W ciągu każdego dnia rejestrowano nie mniej niż 700 tysięcy do maksymalnie 1,3 miliona logów, co wskazuje

na nieustającą, równomierną presję ataków. Analiza osi czasu nie wykazała wyraźnych zależności związanych z porą dnia ani dniem tygodnia – ataki były prowadzone i w nocy, i w dzień, zarówno w weekendy, jak i w dni robocze. Pewne chwilowe skoki aktywności można było jednak zaobserwować w okresach nasilonych prób brute force głównie na usługę SSH oraz przy atakach ukierunkowanych na protokół SMB (rys. 2).



Rys. 2. Diagram aktywności według dni

Zebrane dane były analizowane w kilku etapach. W pierwszym kroku posłużono się systemem wykrywania włamań (*intrusion detection system suricata*), który na bieżąco rozpoznawał wzorce ataków i odpowiednio je klasyfikował. W kolejnym etapie dane trafiały do środowiska ELK (Elasticsearch, Logstash, Kibana), gdzie były filtrowane, korelowane i poddawane wizualizacji – co pozwoliło na zidentyfikowanie dominujących kierunków i typów zagrożeń.

W celu uzupełnienia detekcji opartej na sygnaturach zastosowano również dodatkowe narzędzia pozwalające lepiej zrozumieć, kim byli napastnicy i z jakich środowisk operowali. Jednym z nich było narzędzie FATT, które umożliwia analizę przechwyconego ruchu sieciowego w czasie rzeczywistym, rozpoznając obsługiwane protokoły (takie jak HTTP, FTP czy SMB) oraz charakterystyczne wzorce zachowań typowe dla znanych technik ataku lub nietypowych klientów. W kontekście stworzonego środowiska FATT pozwalało wydobywać konkretne informacje o metodach i narzędziach używanych przez atakujących bez potrzeby dalszej interakcji z nimi. Równolegle wykorzy-

stywano także narzędzie p0f, również działające w pełni pasywnie, które na podstawie cech segmentów TCP – takich jak wartość TTL, rozmiar okna czy użyte flagi – umożliwiało rozpoznanie systemu operacyjnego hosta próbującego nawiązać połączenie. Dzięki temu możliwe było określenie, czy dany ruch pochodził np. z systemu Windows, Linux czy innej platformy, a także wykrycie obecności NAT-ów, zapór sieciowych lub serwerów proxy po stronie nadawcy. Zastosowanie tych narzędzi pozwoliło lepiej profilować źródła ataku, odróżnić przypadkowe połączenia od celowych działań oraz zyskać szerszy obraz środowiska, z którego przeprowadzano próby kompromitacji systemu.

2. Analiza pozyskanych danych

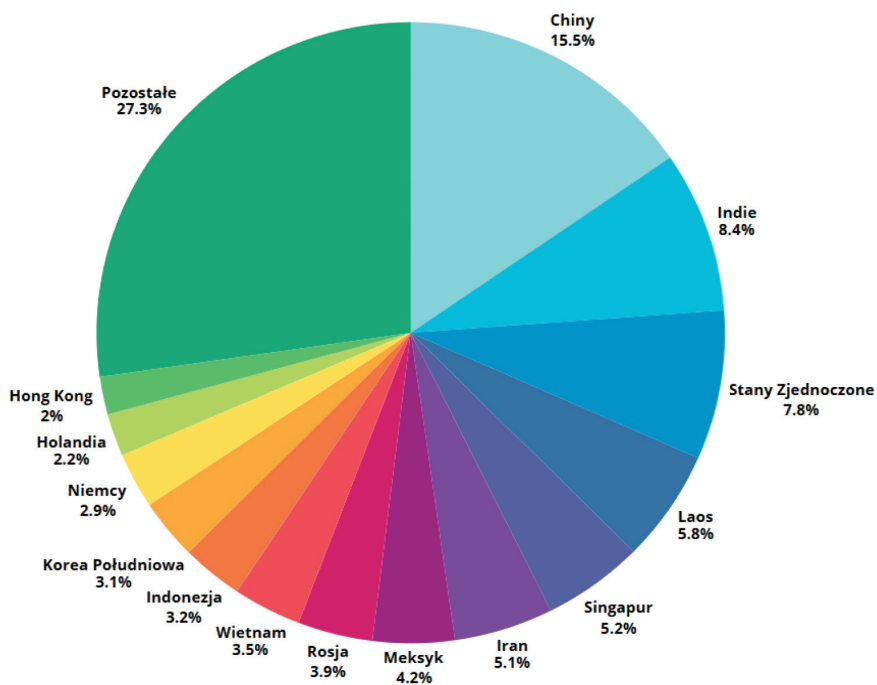
2.1. Geolokalizacja źródeł ataków

Podczas dwutygodniowego działania honeypota zarejestrowano ruch pochodzący z co najmniej 160 różnych państw. Tak szeroki zasięg geograficzny świadczy o globalnej skali zagrożeń i powszechnej aktywności botów oraz zautomatyzowanych narzędzi skanujących.

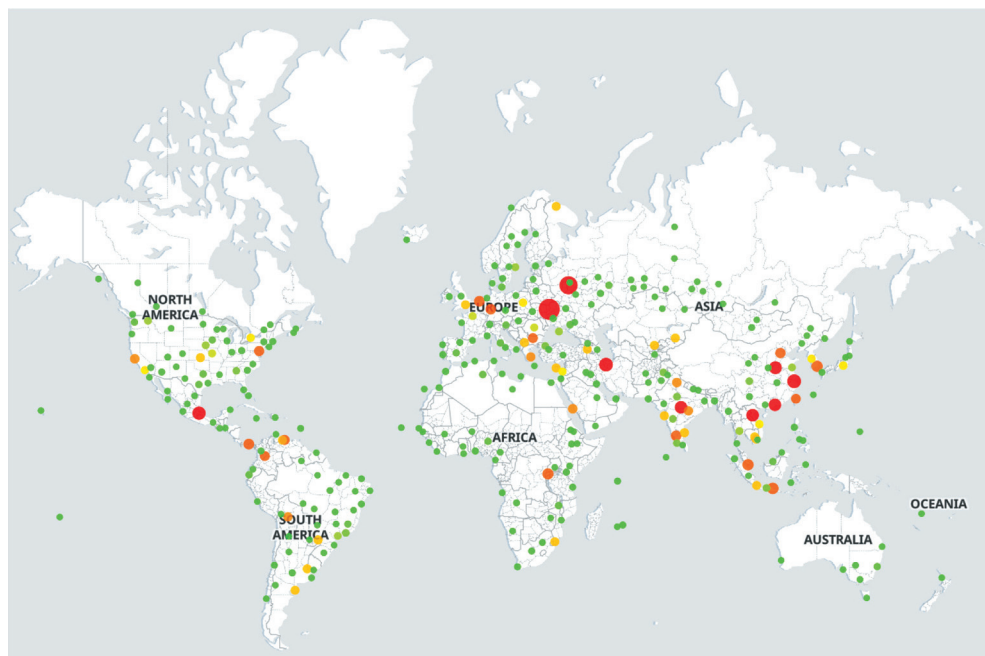
Warto jednak zaznaczyć, że ustalenie geolokalizacji adresów IP nie jest procesem w pełni precyzyjnym. Część napastników korzysta z infrastruktur pośredniczących – takich jak VPN, sieć TOR, serwery proxy czy platformy chmurowe – co utrudnia jednoznaczne przypisanie konkretnego adresu do fizycznej lokalizacji. Mimo tych ograniczeń dane geolokalizacyjne pozwalają uzyskać ogólny obraz najbardziej aktywnych regionów i kierunków, z których pochodził złośliwy ruch.

Jak wynika z analizy adresów IP, wśród państw generujących największy wolumen złośliwych połączeń znalazły się kolejno: Chiny (15,5%), Indie (8,4%), Stany Zjednoczone (7,8%), a następnie Laos, Singapur, Iran, Meksyk oraz Rosja (rys. 3). Aby przedstawić intensywność ataków w ujęciu przestrzennym, opracowano heatmapę, która pozwala na szybkie rozpoznanie głównych ognisk aktywności. Wśród najczęściej występujących lokalizacji znalazły się: Moskwa, Kijów, Szanghaj, Teheran, Zhengzhou, Meksyk, Chhindwara (Indie), Hongkong i Hanoi (rys. 4). Szczególnie interesującym spostrzeżeniem była wysoka aktywność przypisana do adresów IP pochodzących z Ukrainy, co – w kontekście trwającego konfliktu zbrojnego z Rosją – może wskazywać zarówno na działania ofensywne, jak i wykorzystanie lokalnej infrastruktury przez podmioty trzecie. Pewnym zaskoczeniem okazała się również intensywność ruchu z Meksyku, który rzadziej bywa uwzględniany w analizach najaktywniejszych źródeł cyberataków.

Pozostałe lokalizacje – takie jak Iran, Rosja czy Chiny – ze względu na swoją dotychczasową reputację w analizach cyberzagrożeń nie budzą większego zdziwienia i potwierdzają ogólnoświatowe obserwacje dotyczące źródeł złośliwej aktywności sieciowej.



Rys. 3. Procentowy udział poszczególnych krajów w złośliwym ruchu

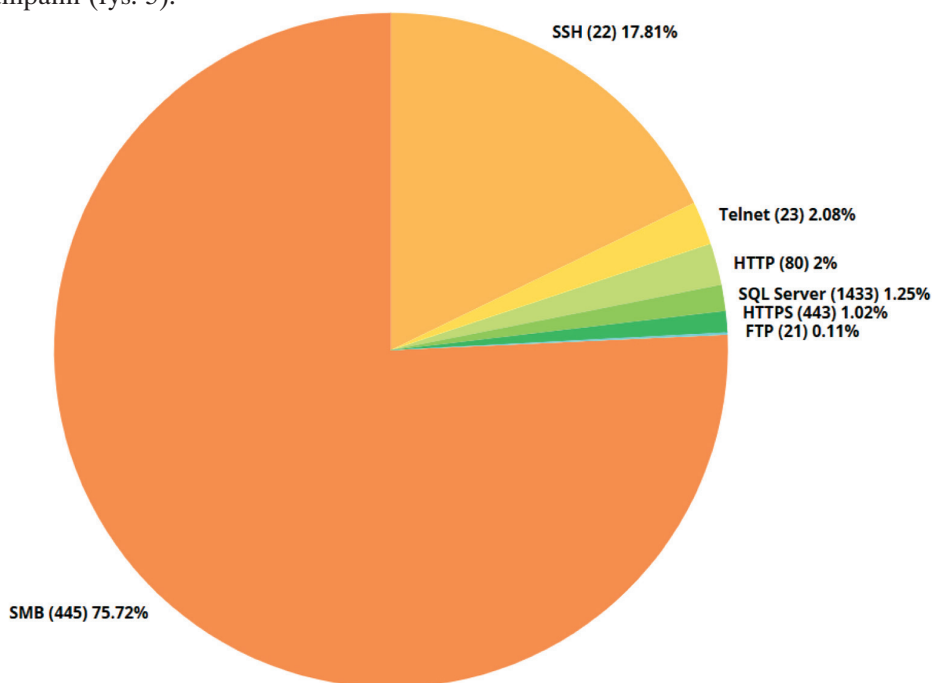


Rys. 4. Heatmapa prezentująca rozkład źródeł ataków

2.2. Atakowane usługi

Wyniki eksperymentu wykazały wyraźną koncentrację złośliwego ruchu na dwóch usługach: SMB i SSH. Protokół SMB, wdrożony celowo w przestarzałej i podatnej wersji SMBv1, przyciągnął ponad 75% całkowitego ruchu sklasyfikowanego jako złośliwy, stając się głównym celem ataków. Drugim najczęściej atakowanym punktem była usługa SSH, która odpowiadała za około 18% całkowitej aktywności.

Z kolei ruch związany z pozostałymi usługami – takimi jak FTP, Telnet, HTTP, HTTPS i SQL Server – stanowił niespełna 7% całości i miał marginalny charakter. W jego obrębie odnotowano głównie automatyczne próby skanowania portów oraz pobieżne testy usług, bez oznak skoordynowanych lub ukierunkowanych złośliwych kampanii (rys. 5).



Rys. 5. Liczba ataków według usługi

2.2.1. SMB

W ramach działania honeypota emulującego usługę SMB przechwycono 94 próbki złośliwego oprogramowania przesyłanego przez napastników w trakcie prób wykorzystania znanych podatności. Spośród nich aż 92 zostały sklasyfikowane jako różne warianty ransomware WannaCry. W obrębie tej grupy udało się zidentyfikować 39 unikalnych odmian, co pokazuje, że mimo upływu czasu zagrożenie to nadal występuje w sieci internet.

Warto przypomnieć, że pierwotny atak WannaCry z maja 2017 roku, choć bardzo skuteczny, został dość szybko powstrzymany dzięki istnieniu tzw. killswitcha – domeny, której obecność dezaktywowała szkodliwe działanie programu. Niedługo później kod źródłowy tego ransomware zaczął krążyć w przestrzeni publicznej, stając się bazą do tworzenia zmodyfikowanych, trudniej wykrywalnych wersji przez społeczność cyberprzestępczą. To właśnie ta aktywność – oparta na adaptacjach oryginalnego kodu – tłumaczy tak dużą liczbę różniących się od siebie technicznie próbek, mimo że sam mechanizm infekcji nie uległ znaczącej ewolucji. Obecnie WannaCry nie stanowi realnego zagrożenia dla większości nowoczesnych systemów, ponieważ jego skuteczność opiera się na wykorzystaniu starej luki w protokole SMBv1, która została dawno załatwana. W rezultacie ruch związany z tym złośliwym oprogramowaniem ma dzisiaj raczej charakter ciekawostki niż aktywnego zagrożenia – stanowi przykład długowieczności kodu i automatyzacji działań botnetów, które wciąż próbują infekować przypadkowe cele w sieci. W kontekście tej obserwacji warto zauważyć, że eksperyment był realizowany na systemie Linux, który nie jest podatny na ataki z wykorzystaniem EternalBlue w tym WannaCry, co dodatkowo podkreśla relatywnie wyższy poziom bezpieczeństwa tego systemu operacyjnego w środowiskach serwerowych w porównaniu ze starszymi wersjami Windows.

Poza rodziną WannaCry zidentyfikowano również dwie inne próbki złośliwego oprogramowania. Pierwszą z nich był ZombieBoy, znany również pod nazwą Midie – złośliwe oprogramowanie zaliczane do klas trojan i miner kryptowalut. Głównym celem tego oprogramowania jest potajemne wykorzystanie zasobów obliczeniowych zainfekowanej maszyny do wydobywania kryptowaluty Monero, przy jednoczesnym zapewnieniu operatorowi zdalnego dostępu do systemu ofiary. ZombieBoy wyposażony jest w funkcje umożliwiające utrzymanie się w systemie, nawiązywanie połączenia z serwerem Command & Control oraz rozpoznanie środowiska operacyjnego. W niektórych wersjach zaobserwowano również komponenty umożliwiające lateralne przemieszczanie się po sieci lokalnej. Program ten jest trudny do wykrycia, ponieważ nie wykazuje agresywnych symptomów w działaniu, a jego głównym celem jest długotrwała eksploatacja zasobów systemu bez wzbudzania podejrzeń użytkownika.

Drugą zidentyfikowaną próbkę stanowiło złośliwe oprogramowanie znane jako Mikey/CryptInject/MulDrop7. Oprogramowanie to charakteryzuje się modułarną budową, pozwalającą na dynamiczne dostarczanie kolejnych komponentów w trakcie infekcji. Po początkowym uruchomieniu malware instaluje dodatkowy kod wykonywalny, który może pełnić funkcje droppera, backdoora lub narzędzia do dalszego rozprzestrzeniania infekcji w sieci. Jedną z jego najgroźniejszych cech jest trwałość – potrafi przetrwać ponowne uruchomienie systemu, unikać wykrycia przez typowe skanery antywirusowe oraz tworzyć kopie samego siebie w różnych lokalizacjach systemu plików.

Wykorzystuje również techniki zaciemniania kodu i niestandardowe schematy komunikacji z serwerami C2, co czyni go wysoce problematycznym w usuwaniu.

Oba zagrożenia, choć mniej powszechne niż WannaCry, świadczą o rosnącej różnorodności i technicznej złożoności szkodliwego oprogramowania wykorzystywanego w atakach sieciowych. Ich obecność w środowisku honeypotowym wskazuje, że nawet niewielkie, niepozorne wektory ataku mogą być wykorzystywane do testowania nowych form ataków lub rozprzestrzeniania złośliwego kodu na większą skalę.

Warto nadmienić, że ataki na protokół SMB obserwowane podczas eksperymentu charakteryzowały się wysokim poziomem automatyzacji – począwszy od skanowania portów i wykrywania podatnych usług, aż po próbę przesyłania złośliwego kodu. Często odbywały się one w ramach dłuższych sesji TCP pochodzących z tych samych adresów IP, co może wskazywać na działanie zautomatyzowanych botów.

2.2.2. SSH

Drugim istotnym wektorem ataku był protokół SSH, który przyciągał napastników próbujących uzyskać dostęp do systemu za pomocą zautomatyzowanych ataków typu brute force i ataków słownikowych. W środowisku testowym wdrożono specjalnie przygotowaną emulację tej usługi – bazującą na honeypocie Cowrie – która została skonfigurowana w taki sposób, aby po kilku nieudanych próbach logowania umożliwić atakującemu dostęp do pozorowanego terminala systemu Linux. Niezależnie od wprowadzonych danych uwierzytelniających system przyznawał dostęp po określonej liczbie błędnych prób, symulując skuteczne zalogowanie się, co pozwalało obserwować autentyczne zachowanie napastników w środowisku przypominającym rzeczywisty system.

Po uzyskaniu dostępu atakujący przystępowali zazwyczaj do serii poleceń rozpoznawczych, które miały na celu identyfikację środowiska i możliwości systemu. Najczęściej były to takie komendy jak: whoami, uname -a, ifconfig, ls czy ps aux, których celem było ustalenie uprawnień użytkownika, wersji systemu, dostępnych interfejsów sieciowych oraz uruchomionych procesów. W kolejnym kroku napastnicy podejmowali próbę pobrania złośliwego oprogramowania z zewnętrznych źródeł, korzystając z narzędzi takich jak wget i curl, wskazując konkretne adresy URL. W wyniku tych działań zarejestrowano 45 unikalnych hashy plików, będących odpowiedzią na próby pobierania i uruchamiania złośliwego kodu. Wśród nich 10 próbek nie znajdowało się w publicznych bazach danych i mogły reprezentować nowe lub niestandardowe warianty złośliwego oprogramowania. Kolejne 7 plików okazało się nieszkodliwymi skryptami, a pozostałe 28 hashy odpowiadało znanemu złośliwemu oprogramowaniu, należącemu głównie do rodzin botnetów oraz narzędzi do kopania kryptowalut, których działanie opiera się na przejmowaniu kontroli nad urządzeniem i wykorzystywaniu go do dalszych ataków lub generowania zysków.

Wśród najczęściej identyfikowanych zagrożeń pojawiły się cztery dobrze znane botnety, regularnie wykorzystywane w zautomatyzowanych kampaniach:

1. Mirai – jeden z najbardziej rozpoznawalnych botnetów w historii ataków typu DDoS, zaprojektowany z myślą o infekowaniu urządzeń IoT (takich jak kamery, telewizory, routery) wykorzystujących domyślne lub słabe dane logowania. Po przejęciu urządzenia Mirai rejestruje się na serwerze Command & Control, z którego odbiera instrukcje do przeprowadzania ataków. Jego kod źródłowy został upubliczniony w 2016 roku, co doprowadziło do powstania wielu odmian, z których część nadal pozostaje aktywna.
2. Multiverze – mniej znany, ale wysoce funkcjonalny modularny botnet, zdolny do pełnienia wielu ról jednocześnie. Może działać jako agent DDoS, serwer proxy, sniffer sieciowy, a także pobierać dodatkowe komponenty złośliwego oprogramowania. Znany jest z dynamicznego adresowania serwerów C2, wykorzystywania szyfrowanej komunikacji i dużej elastyczności, dzięki czemu trudniej go wykryć i zablokować.
3. Bashlite (znany również jako Gafgyt, Qbot) – botnet celujący w urządzenia z systemem Linux, zwłaszcza te z otwartymi portami Telnet i SSH, gdzie możliwe jest logowanie się przy użyciu fabrycznych danych uwierzytelniających. Zainfekowane urządzenia wykorzystywane są głównie do przeprowadzania ataków DDoS, a Bashlite często występuje w kombinacji z innymi narzędziami malware, stanowiąc część większych kampanii.
4. XORDDoS – zaawansowany botnet, który stosuje szyfrowanie kodu za pomocą operacji XOR, co znacznie utrudnia jego detekcję. Oferuje funkcjonalności root-kita, umożliwiające ukrywanie jego obecności w systemie, a także rozbudowany zestaw modułów do przeprowadzania ataków DDoS, infekcji oraz zdalnego zarządzania hostem.

Wszystkie ataki na SSH miały wysoce zautomatyzowany charakter – cały proces, od wyszukania dostępnego hosta i otwartego portu, przez próbę logowania, aż po instalację malware'u, odbywał się bez udziału człowieka. Obserwowane zachowanie wskazuje na użycie gotowych zestawów skryptów oraz powiązanie z infrastrukturą typu Command & Control, która sterowała aktywnością botów i dostarczała im instrukcje. W zdecydowanej większości przypadków działania te odbywały się w ramach pojedynczych, krótkich sesji, co świadczy o ich w pełni automatycznym charakterze.

Szczególnie interesująca była dynamika ataków w czasie. W początkowej fazie eksperymentu liczba prób logowania i infekcji była relatywnie niska. Jednak około dziesiątego dnia obserwacji nastąpił wyraźny wzrost intensywności ruchu. Może to świadczyć o tym, że adres IP honeypota został zaindeksowany w jednej z ogólnodostęp-

nych baz danych, z których korzystają botnety oraz systemy masowego skanowania sieci w poszukiwaniu nowych celów. Od tego momentu system był regularnie atakowany z różnych źródeł, co znacząco zwiększyło liczbę rejestrowanych zdarzeń.

2.3. Ataki brute force i słownikowe

W czasie trwania eksperymentu zaobserwowano dużą aktywność związaną z atakami typu brute force i słownikowymi. Ich celem było uzyskanie dostępu do systemu przez zgadywanie kombinacji loginu i hasła. Na podstawie zarejestrowanych prób logowania opracowano wizualizacje przedstawiające najczęściej wykorzystywane nazwy użytkowników oraz hasła.

Wśród loginów zdecydowanym liderem był „root”, który pojawiał się najczęściej w atakach – co nie dziwi, biorąc pod uwagę, że jest to domyślna nazwa konta administracyjnego w systemach Linux. Na dalszych pozycjach uplasowały się również inne powszechnie spotykane nazwy kont, takie jak „admin”, „user”, „ftpuser”, „guest” oraz konta przypisane do konkretnych usług i aplikacji np. „gitlab”, „nginx”, „hadoop”, „postgres”, „esuser”, „zabbix”, „oracle” (rys. 6).

W zakresie haseł dominowały trywialne, często stosowane i łatwe do zgadnięcia kombinacje, w tym przede wszystkim „1234”, „123456”, „admin”, „admin123”, „password”, „test123”, a także warianty z użyciem popularnych wzorców klawiaturowych, jak „qwerty123”, „1qaz@WSX”, „qwe123” czy „123qaz”. Pojawiły się również próby logowania z pustym hasłem ((empty)) oraz bardziej złożone ciągi znaków, które najprawdopodobniej pochodziły z automatycznie generowanych list haseł (rys. 7).

Dane te jednoznacznie potwierdzają, że większość ataków opiera się na gotowych, ogólnodostępnych słownikach i koncentruje się na standardowych nazwach kont oraz słabych hasłach – co czyni niezabezpieczone systemy wyjątkowo podatnymi. Ich skuteczność byłaby wysoka w środowiskach, w których pozostawiono domyślne dane logowania. Co istotne, podczas eksperymentu udało się przechwycić kompletne słowniki haseł, zawierające setki, a niekiedy nawet tysiące unikalnych pozycji. Było to możliwe dzięki niedoskonałości w sposobie przeprowadzania ataków przez niektóre boty. Mechanizm brute force działał w taki sposób, że każda próba logowania kończyła się natychmiastowym zerwaniem sesji SSH, po czym uruchamiana była nowa próba z kolejną kombinacją loginu i hasła. W wyniku takiego działania nie dochodziło do pozorowanego zalogowania się do emulowanej powłoki, co oznacza, że jedyną rejestrowaną aktywnością były właśnie ciągi uwierzytelniające przesyłane w każdej sesji.

Dzięki temu możliwe było zrekonstruowanie pełnych słowników wykorzystywanych przez boty, co stanowi cenne źródło wiedzy o stosowanych metodach ataku, a także może posłużyć jako materiał referencyjny dla analityków bezpieczeństwa czy

Szczegółowe przykłady tego typu aktywności zostały już omówione w kontekście usług SMB oraz SSH, gdzie działania atakujących przebiegały od skanowania, przez próby logowania aż po instalację złośliwego oprogramowania – bez udziału człowieka.

Analogicznie również ruch skierowany na pozostałe emulowane usługi – FTP, Telnet, HTTP, HTTPS oraz SQL Server – nosił wyraźne znamiona automatyzacji. W tych przypadkach dominowały szybkie i powierzchowne próby rozpoznawcze, obejmujące skanowanie portów, identyfikację wersji usług oraz ocenę ich dostępności. Dla protokołów webowych dodatkowo zarejestrowano próby enumeracji katalogów oraz testowania reakcji serwera na standardowe zapytania HTTP, co sugeruje użycie narzędzi skanujących, takich jak Nikto czy Gobuster.

Nie odnotowano jednak żadnych przypadków dalszych działań czy prób faktycznego wykorzystania wykrytych usług – aktywność ograniczała się wyłącznie do szybkiego „rozpoznania terenu”. Tęgo typu operacje charakteryzowały się krótkimi, losowo rozłożonymi w czasie sesjami, których celem było katalogowanie dostępnych zasobów pod kątem przyszłych ataków. Choć skala tych prób była zauważalna, to brak jakiegokolwiek głębszej interakcji czyni je dużo mniej istotnymi z punktu widzenia zagrożeń niż działania zaobserwowane odnośnie do SMB i SSH. Dzięki zastosowanym narzędziom do pasywnej analizy ruchu sieciowego – FATT oraz p0f – możliwe było określenie, z jakich systemów operacyjnych pochodziły zarejestrowane próby ataków. Analiza metadanych wykazała, że aż 69% zarejestrowanych działań inicjowanych było z systemów opartych na jądrze Linux, co odzwierciedla powszechne stosowanie tej platformy w środowiskach botnetowych i automatycznych skanerach. Kolejne 30% pochodziło z systemów Windows i dotyczyły głównie ataków na usługę SMB. Zaledwie 1% przypadków przypisano do innych platform, takich jak macOS i FreeBSD. Struktura ta dobrze obrazuje preferencje środowisk przestępczych – systemy Linux są chętnie wykorzystywane ze względu na dostępność darmowych narzędzi, łatwość automatyzacji i możliwość uruchamiania lekkich botów bez potrzeby dostępu do interfejsu graficznego.

Analiza źródeł ruchu wskazuje, że większość atakujących nie podejmowała prób ukrywania swojej tożsamości ani śladów działań. Ponad 99,9% adresów IP biorących udział w atakach było już wcześniej zidentyfikowanych i zaindeksowanych w publicznie dostępnych bazach zagrożeń, takich jak VirusTotal czy AbuseIPDB. Spośród tysięcy prób zarejestrowanych w trakcie eksperymentu zaledwie 48 połączeń pochodziło z zewnętrznych węzłów sieci TOR, co oznacza, że nacisk na anonimizację i ukrycie nie był priorytetem dla atakujących.

Brak zastosowania zaawansowanych technik maskujących jednoznacznie sugeruje, że głównym celem była szeroko zakrojona enumeracja oraz próby kompromitacji systemów o znanych podatnościach, bez konieczności inwestowania zasobów w unikanie wykrycia i ukrywanie tożsamości. Takie podejście cechuje najczęściej działania

botów, których zadaniem jest przesiewowe skanowanie sieci publicznej w poszukiwaniu najłatwiejszych celów.

4. Wnioski i rekomendacje

W przeprowadzonym eksperymencie udało się nie tylko uchwycić skalę i intensywność współczesnych ataków cybernetycznych, ale również dokonać ich wieloaspektowej analizy – obejmującej źródła, charakter, techniki, czas występowania, wykorzystywane narzędzia i cele atakujących. Dzięki integracji wielu komponentów monitorujących możliwe było spojrzenie na złośliwą aktywność z różnych perspektyw: zarówno ilościowej, jak i jakościowej, co pozwoliło uzyskać pełniejszy obraz rzeczywistych zagrożeń, z jakimi mierzy się publicznie dostępną infrastrukturą sieciową.

Ruch sieciowy zarejestrowany w trakcie badania pochodził z ponad 160 państw świata, co unaocznia globalny charakter zagrożeń i skalę działania współczesnych, zautomatyzowanych kampanii cyberataków. Szczególnie intensywne ogniska aktywności wykryto w krajach takich jak Chiny, Indie, Stany Zjednoczone, Rosja, Ukraina i Iran – co pokrywa się z wynikami licznych raportów branżowych i akademickich dotyczących źródeł złośliwej aktywności. Tego rodzaju zgodność potwierdza, że nawet w badaniach eksperymentalnych możliwe jest uchwycenie realnych trendów występujących w globalnym krajobrazie cyberzagrożeń. Analiza czasowa nie wykazała wyraźnych zależności w rozkładzie ataków – rejestrowano je zarówno w dzień, jak i w nocy, bez względu na konkretny dzień tygodnia. Taki brak regularnych cykli jednoznacznie świadczy o globalnym zasięgu i pełnej automatyzacji działań, realizowanych przez złożone sieci botnetów i narzędzia skryptowe działające w sposób ciągły i niezależny od stref czasowych czy aktywności użytkowników. Obserwacja ta ukazuje, że cyberprzestrzeń pozostaje nieustannie penetrowana przez zautomatyzowane mechanizmy rozpoznawcze i ofensywne, wykorzystywane przez różne grupy wpływu w celu uzyskania przewagi – technicznej, informacyjnej lub ekonomicznej – niezależnie od miejsca i czasu.

Analiza usług wskazała na silną koncentrację ataków na dobrze znanych, powszechnie wykorzystywanych protokołach – takich jak SSH i SMB. Pozwoliło to zidentyfikować najbardziej narażone punkty infrastruktury oraz sposoby ich wykorzystywania przez atakujących. Na podstawie przechwyconych sesji możliwe było wyodrębnienie najczęstszych technik wykorzystywanych w początkowej fazie ataku, co stanowi istotną wiedzę z punktu widzenia praktyki bezpieczeństwa sieciowego.

W ramach przeprowadzonego badania udało się zidentyfikować wiele aktywnych zagrożeń w postaci złośliwego oprogramowania oraz botnetów. Szczególne miejsce zajmowały tutaj zmodyfikowane warianty znanych rodzin malware, takich jak ransomware

WannaCry, którego pierwotna wersja pojawiła się już w maju 2017 roku, oraz Mirai – botnet zainicjowany w sierpniu 2016 roku, odpowiedzialny za jedno z największych ataków DDoS w historii. Pomimo upływu lat i powszechnej wiedzy o tych zagrożeniach nadal są one aktywnie wykorzystywane, co świadczy o niepokojącej żywotności starych wektorów ataku. O ile teraz podatne na WannaCry systemy operacyjne stanowią już zdecydowaną mniejszość, ich obecność wciąż daje napastnikom możliwość wywoływania poważnych szkód – zwłaszcza tam, gdzie nie wdrożono odpowiednich aktualizacji zabezpieczeń. Lekcja płynąca z tej obserwacji jest jednoznaczna: systemy informatyczne powinny być regularnie aktualizowane, a korzystanie z przestarzałych, znanych jako podatne środowisk powinno być całkowicie wyeliminowane. W sytuacjach, gdzie z różnych względów (np. wymagań sprzętowych lub aplikacyjnych) nie można tego uniknąć, konieczne jest ich odizolowanie od internetu, np. za pomocą strefy zdemilitaryzowanej (DMZ) lub rozwiązań typu odwrotne proxy, ograniczających powierzchnię ataku i uniemożliwiających bezpośredni dostęp do takiego systemu z zewnątrz. Dodatkowo warto wdrażać wielowarstwowe systemy zabezpieczeń, obejmujące m.in. zapory sieciowe, systemy IDS/IPS, filtry złośliwego oprogramowania oraz polityki regulujące zasady korzystania z aplikacji i systemów. Eksperyment ukazał także istotną rolę polityki haseł w kontekście cyberbezpieczeństwa. Liczne próby logowania, w tym ataki brute force oraz słownikowe, pozwoliły na przeanalizowanie najczęściej używanych poświadczeń. Uwagę zwracała powszechność prób z wykorzystaniem loginów takich jak „root” czy „admin” w parze z banalnymi hasłami typu „123456”, „password” czy „qwerty”. Wyniki te nie pozostawiają złudzeń – domyślne lub zbyt proste dane logowania pozostają głównym celem automatycznych ataków. Oznacza to, że każda organizacja, niezależnie od skali działania, powinna wdrożyć świadomą i konsekwentną politykę haseł, która obejmuje zmianę domyślnych danych uwierzytelniających, stosowanie haseł o odpowiedniej długości i złożoności, cykliczne ich aktualizowanie, a także rozważenie implementacji mechanizmów, takich jak uwierzytelnianie wieloskładnikowe czy blokowanie wielokrotnych nieudanych prób logowania. Kluczowe jest także edukowanie użytkowników, że hasła oparte wyłącznie na samych cyfrach lub literach, popularnych słowach, imionach czy częstych kombinacjach, stanowią realne zagrożenie w obliczu masowych i automatycznych kampanii ataków.

Warto jeszcze raz zdecydowanie podkreślić skalę oraz poziom automatyzacji współczesnych zagrożeń w cyberprzestrzeni. Obecnie aż 51% całkowitego ruchu internetowego generowanego na świecie pochodzi od botów (Imperva 2024), a 37% stanowią boty złośliwe – czyli te, które odpowiadają za działania podobne do zarejestrowanych podczas eksperymentu: masowe skanowanie zasobów sieciowych, ataki typu brute force, kampanie DDoS, a także rozprzestrzenianie złośliwego oprogramowania. Do tego należy

dodać boty spammerskie, rozprzestrzeniające phishing czy tzw. ad fraud bots, które klikają reklamy w celu wyłudzenia środków od reklamodawców. Zjawisko to ma charakter masowy, zautomatyzowany i ciągły, co przekłada się na jego ogromną skalę i częstotliwość – ataki tego typu są dziś codziennością, a ich wykrycie wcale nie wymaga zaawansowanych narzędzi. Co istotne, napastnicy operują bez większych prób ukrywania się: adresy IP wykorzystywane przez boty są często publicznie znane, skatalogowane jako źródła szkodliwej aktywności i łatwe do zidentyfikowania. Brak potrzeby anonimizacji wynika właśnie z bezosobowego, masowego charakteru tych działań, które nie mają na celu ukierunkowanego ataku, lecz polegają na nieustannym, globalnym skanowaniu sieci w poszukiwaniu dowolnych słabych punktów. Ta nowa rzeczywistość cyfrowa, w której większość ruchu generują nie ludzie, lecz automaty, stanowi ogromne wyzwanie dla administratorów systemów i całych organizacji – pokazując, że zagrożenia nie są incydentalne, lecz trwałe i wszechobecne. Choć przeważająca część zarejestrowanej aktywności miała charakter zautomatyzowany i masowy – co samo w sobie stanowi potwierdzenie bezprecedensowej skali współczesnych zagrożeń – eksperyment nie powinien być traktowany wyłącznie jako demonstracja prostych, powtarzalnych działań botów. Właśnie dzięki temu, że środowisko honeypotowe umożliwia rejestrację rzeczywistego, nieocenzurowanego ruchu, możliwe było pozyskanie unikalnych, cennych z punktu widzenia analityki cyberbezpieczeństwa danych. Dobrym przykładem są zidentyfikowane przypadki pobierania przez napastników nieznanych wcześniej skryptów, których sumy kontrolne nie występowały w publicznych bazach złośliwego oprogramowania. Tego rodzaju znaleziska, w przypadku gdy faktycznie okażą się złośliwe, dają olbrzymią przewagę zespołom reagowania – mogą one wcześniej niż reszta świata poznać mechanizmy nowego zagrożenia, a tym samym dostroić swoje systemy obronne, zanim ataki przybiorą na sile i staną się powszechnie znane. Podobnie, przechwytywanie pełnych słowników haseł wykorzystywanych w atakach brute force mogłoby – przy odpowiednio zaawansowanej analizie – ujawnić poziom przygotowania i wiedzy napastników o atakowanym podmiocie. Gdyby okazało się, że słownik jest ukierunkowany i zawiera dane specyficzne dla danej organizacji (np. imiona i nazwiska pracowników, skróty nazw projektów czy loginy tworzone w charakterystyczny dla organizacji sposób), świadczyłoby to o precyzyjnym charakterze ataku oraz o wcześniejszym rozpoznaniu środowiska przez atakujących. Informacje te mogłyby zostać wykorzystane do rewizji procedur bezpieczeństwa, wdrożenia dodatkowych zabezpieczeń oraz do zwiększenia świadomości zagrożeń wśród pracowników.

Przeprowadzony eksperyment potwierdził, że honeypoty stanowią niezwykle użyteczne narzędzie zarówno w kontekście komercyjnym, badawczym, jak i edukacyjnym. Ich największą zaletą jest możliwość pozyskiwania realistycznych danych o rzeczywistych zagrożeniach występujących w sieci. W przeciwieństwie do symulacji czy testów laboratoryjnych, honeypoty pozwalają obserwować działania napastników w czasie rze-

czywistym – przechwytywać techniki skanowania, identyfikację usług, próby włamań, a także próby infekcji złośliwym oprogramowaniem. W trakcie projektu udało się zarejestrować wszystkie etapy ataków: od automatycznego wykrycia dostępnych usług, przez próbę wykorzystania znanych podatności, aż po transfer i próbę uruchomienia złośliwego oprogramowania. Honeypoty takie jak T-Pot umożliwiają emulację różnorodnych środowisk, dzięki czemu badacz ma pełną kontrolę nad scenariuszem eksperymentu, a jednocześnie pozyskuje dane rzeczywistego ruchu pochodzącego z globalnych kampanii ataków cybernetycznych. Przechwycone próbki ransomware WannaCry, botnetów XORDDoS, Mirai czy koparek kryptowalut są tego najlepszym przykładem. Z punktu widzenia edukacyjnego i operacyjnego, honeypoty stanowią znakomite narzędzie do nauki analizy incydentów, pisania reguł detekcji oraz budowy systemów klasyfikujących zagrożenia z użyciem technik uczenia maszynowego. Pozyskiwane w ten sposób autentyczne, zróżnicowane dane umożliwiają praktyczne szkolenie w zakresie analizy ruchu sieciowego, inżynierii wstecznej, korelacji logów, a także testowania i dostrajania systemów wykrywania zagrożeń (EDR, SIEM). Dzięki odizolowanemu środowisku honeypoty pozwalają prowadzić te działania bez ryzyka dla zasobów produkcyjnych, znacząco podnosząc świadomość i kompetencje zarówno studentów, jak i specjalistów ds. bezpieczeństwa.

Eksperyment z użyciem T-Pot okazał się nie tylko udanym przedsięwzięciem z zakresu praktycznej analizy cyberbezpieczeństwa, ale również swoistym oknem na świat zagrożeń, które każdego dnia dotykają globalną sieć. Umożliwił zebranie unikalnych danych i przeprowadzenie wieloaspektowej analizy, a także podkreślił znaczenie tego typu rozwiązań w badaniach, edukacji oraz w realnym wsparciu działań defensywnych.

Literatura

- Experis, 2025, *Prognozy CIO na 2025 rok: Kluczowe priorytety dla liderów technologii*. <https://www.experis.pl/pl/cio-2025-outlook-key-priorities-for-tech-leaders> [dostęp: 7.06.2025].
- Firmansyah D.A., Zahra A., 2023, *Honeypot-Based Thread Detection using Machine Learning Techniques*. <https://ijettjournal.org/Volume-71/Issue-8/IJETT-V71I8P221.pdf> [dostęp: 8.06.2025].
- Imperva, 2024, *2024 Bad Bot Report*. <https://www.imperva.com/resources/resource-library/reports/2024-bad-bot-report/> [dostęp: 4.06.2025].
- Telekom Security, 2024, *T-Pot – The All in One Multi Honeypot Platform* [dokumentacja online]. <https://github.security.telekom.com/2024/04/honeypot-tpot-24.04-released.html> [dostęp: 20.03.2025].
- Telekom Security, 2024, *T-Pot – The All in One Multi Honeypot Platform* [repozytorium kodu źródłowego]. <https://github.com/telekom-security/tpotce> [dostęp: 20.03.2025].

Optimalizacja algorytmu *wave function collapse* pod kątem czasu wykonania

Filip Kowalski 

Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej, Kraków

Streszczenie: Algorytm *wave function collapse* (WFC) ma zalety, dzięki którym stanowi narzędzie przydatne w generowaniu proceduralnym, jednak często jest wymagający obliczeniowo, co ogranicza możliwości jego praktycznego stosowania. W niniejszej pracy przeanalizowano czas wykonania algorytmu WFC i zaproponowano metody jego optymalizacji. Z przeprowadzonych testów wynika, że wąskim gardłem jest krok obserwacji, w szczególności obliczanie minimalnej entropii. Aby skrócić czas wykonania algorytmu, wykorzystano dwie techniki optymalizacyjne: zapamiętywanie wartości entropii poszczególnych komórek oraz buforowanie listy indeksów komórek o aktualnie najniższej entropii. Testy przeprowadzone na siatkach o różnych rozmiarach wykazały, że zaproponowane usprawnienia pozwoliły na ponad 100-krotne przyspieszenie algorytmu, co znacząco poprawia jego użyteczność w zastosowaniach wymagających większej wydajności.

Słowa kluczowe: *wave function collapse*, WFC, optymalizacja, generowanie proceduralne

OPTIMIZATION OF THE WAVE FUNCTION COLLAPSE ALGORITHM FOR EXECUTION TIME

Abstract: The wave function collapse (WFC) algorithm, despite its advantages in procedural generation, is often computationally expensive, limiting its practical applications. The following paper analyzes the execution time of WFC and proposes methods for its optimization. It identifies the observation step, specifically the calculation of minimum entropy, as a key bottleneck. Two successive optimization techniques are introduced: memoization of individual cell entropy values and caching of the list of cell indices with the currently lowest entropy. Tests conducted on grids of varying sizes demonstrate that the proposed improvements achieve speedups exceeding 100×, significantly enhancing the algorithm's utility in performance-demanding applications.

Keywords: wave function collapse, WFC, optimization, procedural generation

https://doi.org/10.7494/978-83-68219-80-7_3

1. Wprowadzenie

Proceduralne generowanie treści (ang. *procedural content generation* – PCG) to technika tworzenia danych za pomocą algorytmów celem zautomatyzowania tego procesu tak, aby wymagał on minimalnego udziału człowieka (Lazaridis i Fragulis 2024). Od lat stanowi ona istotne narzędzie w wielu obszarach informatyki i jest powszechnie wykorzystywana do generowania terenów i miast oraz systemów zadań w grach komputerowych, jak również tekstur, obrazów czy wszelakich animacji (Smith 2021).

Klasyczne przykłady implementacji PCG opierają się na wykorzystaniu technik redukcji szumu, L-systemów lub licznych algorytmów przeszukiwania (Togelius i in. 2011, Maleki i Zhao 2024). Spośród tego rodzaju technik w ostatnich latach na popularności zyskał algorytm *wave function collapse* (WFC) zaprezentowany przez Maxima Gumina (2020). Wyróżnia się on na tle innych metod tym, że daje gwarancję lokalnej spójności generowanych struktur, a także prostotą implementacji i możliwością generowania zasobów na podstawie niewielkiej próbki wejściowej, która niekoniecznie musi być narzucona z góry przez programistę.

Mimo tych zalet algorytm WFC nie zawsze nadaje się do zastosowania w bardziej złożonych zadaniach, ponieważ jest on wymagający obliczeniowo. Czasochłonność algorytmu staje się wyzwaniem szczególnie w przypadku generowania obiektów składających się z dużej liczby pól lub potencjalnych stanów. W niniejszym artykule przedstawiono analizę działania algorytmu WFC oraz propozycje jego optymalizacji pod kątem czasu wykonania, co ma kluczowe znaczenie dla możliwości jego praktycznego wykorzystania przez producentów gier, dla których bardzo istotnym aspektem jest ich płynne działanie, bez konieczności stosowania ekranów ładowania.

1.1. Zasada działania algorytmu WFC

Fundamentem algorytmu WFC jest reprezentująca generowany system siatka wypełniona komórkami, z których każda ma przypisany jeden lub więcej stanów, jakie może przyjąć. Wpływają one na entropię komórki stanowiącą miarę używaną do wyznaczania komórek, które zostaną przetworzone w danym kroku algorytmu. Każdy ze stanów ma zestaw złącz (odpowiadający liczbie kierunków i zwrotów w danej przestrzeni – dla przestrzeni dwuwymiarowej są to cztery złącza) definiujący reguły sąsiedztwa danego stanu z każdym innym. Algorytm rozpoczyna się od inicjalizacji systemu w stanie maksymalnej entropii – każda z komórek może osiągnąć każdy z możliwych stanów. Wraz z wykonywaniem kolejnych kroków algorytmu entropia systemu maleje, aż stanie się on stabilny.

Każdy krok algorytmu polega na wyborze komórki z najniższą entropią i poddaniu jej dwóm procesom: obserwacji i propagacji. Szczegóły tych procedur zostały opisane w kolejnej sekcji¹.

1.2. Stan wiedzy na temat optymalizacji WFC

Problematyka czasu wykonania algorytmu WFC jest przedmiotem badań i tematem licznych publikacji. Istniejące podejścia do optymalizacji można podzielić na usprawnienia algorytmiczne kroku propagacji oraz modyfikacje heurystyk obserwacji. Pierwszy rodzaj podejścia reprezentują Fehr i Courant (2018), którzy wykazali, że adaptacja algorytmu AC-4 (z licznikami wsparcia) pozwala znacząco zredukować złożoność obliczeniową względem naiwnych metod sprawdzania spójności.

Podjęmowane są także próby skalowania algorytmu przez jego dekompozycję na warstwy abstrakcji, co w metodzie hierarchicznej zaproponowali Beukman i współpracownicy (2023). Inni autorzy zauważają jednak, że proste metody przyspieszania obliczeń przez zrównoleglenie mają istotne ograniczenia. Jak wskazują López i Chover (2025), silne zależności skutkujące tym, że zmiana w jednej komórce może kaskadowo wpłynąć na całą mapę, czynią dekompozycję na niezależne wątki szczególnie skomplikowanym zadaniem. Wobec trudności z efektywnym zrównolegleniem kluczowa pozostaje optymalizacja sekwencyjna kroku obserwacji. W tym obszarze standardem jest wykorzystanie kopców, jednak w niniejszym artykule zaproponowano mniej wymagającą obliczeniowo technikę buforowania indeksów, dostosowaną do charakterystyki WFC.

2. Cel i metodyka badań

Celem prezentowanego artykułu jest: 1. analiza czasu wykonania algorytmu dla różnych zestawów parametrów wejściowych; 2. zidentyfikowanie kroków algorytmu, które są wymagające obliczeniowo; 3. zaproponowanie sposobów optymalizacji poszczególnych operacji.

2.1. Założenia algorytmu

Algorytm WFC wyróżnia się uniwersalnością. Może on działać w dowolnej liczbie wymiarów, wspierać symetryczne i niesymetryczne złącza poszczególnych stanów, a także pozwala na analizę próbki wejściowej zamiast wykorzystania określonego zestawu

¹ Więcej na temat działania algorytmu zob. Karth i Smith 2017.

zasad. Przygotowując implementację algorytmu na potrzeby badania, przyjęto następujące założenia:

- Algorytm działa w środowisku dwuwymiarowym.
- Algorytm wspiera tylko symetryczne połączenia poszczególnych stanów.
- Szansa wystąpienia każdego ze stanów może być inna.
- Reguły sąsiedztwa są niezmiennie i określone przed uruchomieniem algorytmu.

Przedstawione ograniczenia implementacji algorytmu pozwoliły uprościć proces analizy i optymalizacji, jednocześnie zachowując kluczowe aspekty i kroki algorytmu.

2.2. Metodyka badań

W celu uzyskania obiektywnych wyników pomiar czasu wykonania algorytmu przeprowadzono w następujący sposób:

- Czas wykonania każdego z kluczowych etapów pojedynczego kroku został zmierzony za pomocą funkcji bibliotecznych udostępnionych w ramach języka C#.
- Po zakończeniu działania algorytmu WFC dla danego systemu zmierzono czas wykonania algorytmu 1, który został uznany za bazową jednostkę czasu dla tego pomiaru.
- Czas wykonania każdego z kluczowych etapów pojedynczego kroku został podzielony przez jednostkę bazową czasu.
- Przed analizą dane zostały poddane przetworzeniu celem usunięcia błędów pomiarowych. Pomiar był uznawany za poprawny, jeśli mieścił się w przedziale:

$$Q_1 - 5 \cdot IQR \leq x \leq Q_3 + 5 \cdot IQR \quad (1)$$

gdzie x to wartość pomiaru, Q_1 oraz Q_3 to odpowiednio dolny i górny kwartył zestawu danych, a IQR to rozstęp ćwiartkowy wyrażony wzorem:

$$IQR = Q_3 - Q_1 \quad (2)$$

Opisany proces pomiarowy gwarantuje otrzymanie wyników, które są niezależne od parametrów jednostki obliczeniowej oraz stanu urządzenia w momencie uruchomienia algorytmu. Nawet jeśli maszyna zostanie chwilowo obciążona, kalkulacje bazowej jednostki czasu zostaną spowolnione w tym samym stopniu co wygenerowanie treści.

```
int[] baselineArr = new int[10000];
int baselineCount = baselineArr.Length;

for (int i = 0; i < baselineCount; i++) {
    baselineArr[i] = i;
}
```

Algorytm 1. Obliczenie bazowej jednostki czasu

2.3. Przeprowadzanie testów

Każdy test szybkości działania algorytmu został przeprowadzony z wykorzystaniem takiego samego zestawu stanów oraz zasad sąsiedztwa. Zestaw składa się z 32 różnych możliwych do osiągnięcia stanów oraz 8 unikalnych identyfikatorów złącz. Generowanie treści na potrzeby badań wykonano dla siatek o następujących rozmiarach:

- 50×50 komórek (w sumie 2500),
- 100×100 komórek (w sumie 10 000),
- 200×200 komórek (w sumie 40 000).

3. Implementacja algorytmu w formie bazowej

Algorytm WFC został zaimplementowany z wykorzystaniem języka programowania C#, w środowisku Unity. Wybór tego środowiska pozwolił na wykorzystanie gotowych metod wizualizacji rezultatów działania algorytmu bez dodatkowego wydłużania czasu jego działania. Na potrzeby badań stworzony został modularny i prosty w obsłudze system implementujący określone założenia algorytmu WFC.

3.1. Kluczowe kroki algorytmu

Algorytm WFC w istocie sprowadza się do wykonywania naprzemiennie dwóch procedur, aż układ osiągnie stan stabilny. Warunek ten jest spełniony wtedy i tylko wtedy, gdy wszystkim komórkom w układzie przypisany został konkretny stan. Główna pętla algorytmu została przedstawiona za pomocą algorytmu 2.

```
while(true) {  
    Cell cell = Collapse();  
  
    if(cell != null) {  
        Propagate(cell);  
    } else {  
        break;  
    }  
}
```

Algorytm 2. Główna pętla przetwarzania algorytmu WFC

3.1.1. Obserwacja (*collapse*)

Krok obserwacji rozpoczyna się od wytypowania komórki układu o najmniejszej entropii. W przypadku istnienia kilku komórek spełniających ten warunek algorytm wybiera ją z tego zbioru losowo. Następnie przypisywany jest jej losowy dostępny stan

(na podstawie rozkładu prawdopodobieństwa ich wystąpienia). Jeśli cały system osiągnął już w równowagę, krok ten zwraca wartość pustą.

W kroku obserwacji często wyróżnia się również podkrok związany z obliczaniem minimalnej entropii i wyborem komórki, która zostanie poddana dalszej części procesu obserwacji.

3.1.2. Propagacja (*propagate*)

W tym kroku komórki sąsiadujące (w znaczeniu sąsiedztwa von Neumanna) z komórką wybraną na wejściu poddawane są procesowi redukcji stanów, które mogą przyjąć. Nowy zestaw stanów jest wynikiem aplikacji reguł sąsiedztwa w ujęciu stanów komórki, z której następuje propagacja, oraz komórki docelowej.

Jeśli w wyniku przeprowadzenia tego kroku komórka sąsiednia zmniejszy swoją liczbę dostępnych stanów, ona również zostaje poddana procesowi propagacji. Oznacza to, że redukcja, która właśnie zaszła, może wymusić redukcję stanów kolejnych granicznych komórek.

Aby możliwe było zakończenie procesu propagacji w jak najmniejszej liczbie kroków, stosowana jest pomocnicza tablica określająca, czy dana komórka została w tym kroku algorytmu odwiedzona. Ze względu na specyfikę tej operacji odwiedzenie każdej komórki więcej niż raz nie spowoduje dodatkowej redukcji stanów.

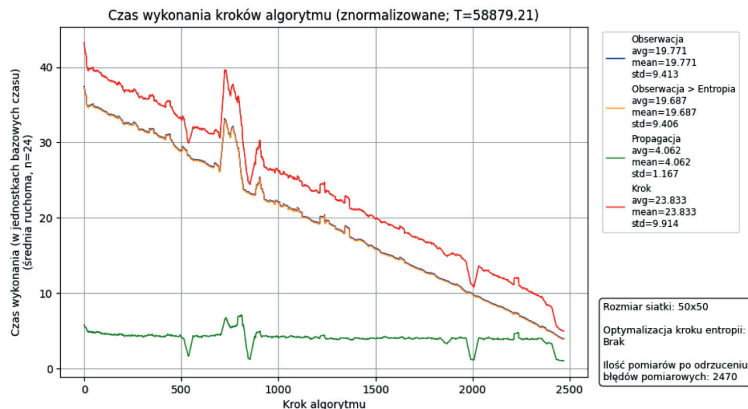
3.2. Czas wykonania

Algorytm uruchomiony bez żadnych dodatkowych optymalizacji pozwala na generowanie treści niewielkich rozmiarów niemalże w czasie rzeczywistym. Widoczne jest jednak znaczne pogorszenie wydajności w przypadku większych rozmiarów generowanej planszy, co zostało ujęte w tabeli 1.

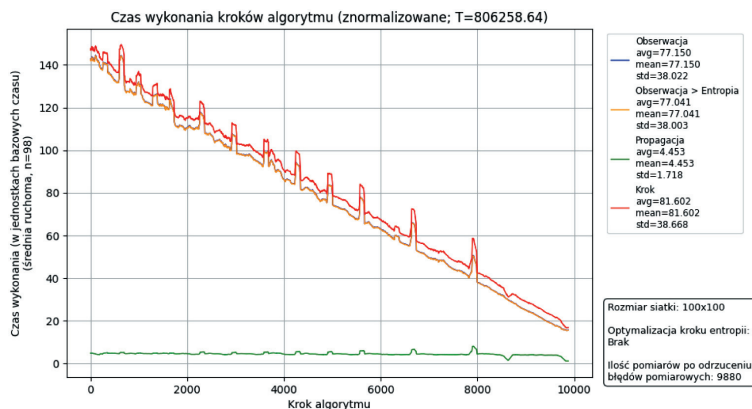
Tabela 1
Czas wykonania algorytmu WFC bez optymalizacji

Wysokość siatki	Szerokość siatki	Czas wykonania (w jednostkach bazowych czasu)
50	50	58 879,21
100	100	806 258,64
200	200	17 580 450,19

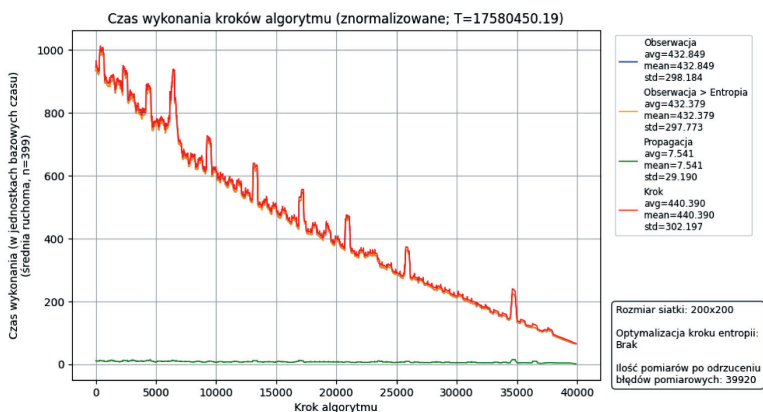
Celem dokładniejszej analizy przebiegu algorytmu oraz znalezienia sposobów optymalizacji procesu na rysunkach 1–3 przedstawiono czas wykonania poszczególnych kroków algorytmu z rozbiciem na etapy obserwacji (z wyszczególnieniem podkroku wyliczania minimalnej entropii) i propagacji.



Rys. 1. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 50×50 , brak optymalizacji kroku entropii)



Rys. 2. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 100×100 , brak optymalizacji kroku entropii)



Rys. 3. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 200×200 , brak optymalizacji kroku entropii)

Z rysunków 1–3 wynika, że w przypadku kroku obserwacji niemal cały czas wykonania przypada na podkrok obliczania entropii. Świadczą o tym znacznie zbliżone do siebie parametry wartości środkowej danych oraz odchylenia standardowego.

Po przeanalizowaniu sposobu działania algorytmu łatwo zauważyć, że to właśnie krok związany z wyliczaniem minimalnej entropii oraz wybieraniem komórki będzie najbardziej wymagający obliczeniowo. Do przeprowadzenia procedury konieczne jest sprawdzenie entropii każdego z pól celem określenia wartości minimalnej, następnie wyznaczenie komórek, które cechują się taką entropią, oraz ostatecznie losowe wybranie tej z nich, która zostanie poddana obserwacji.

Warto również zauważyć, że czas wykonania kroku entropii jest krótszy wraz z kolejnymi krokami algorytmu. Dzieje się tak dlatego, że w procesie obliczania minimalnej entropii i wybierania kandydata do obserwacji uczestniczą komórki, którym nie został przypisany żaden określony stan. Ponieważ w każdym kroku algorytm przypisuje jednej z komórek konkretny stan (podczas kroku obserwacji), naturalne jest, że liczba komórek uwzględnionych w kroku entropii zacznie spadać. Krok ten jest zatem kluczowy dla optymalizacji algorytmu.

4. Optymalizacja algorytmu

Rozdział ten poświęcono przedstawieniu technik optymalizacji algorytmu WFC z wykorzystaniem danych zebranych w wyniku uruchomienia algorytmu w wersji bazowej.

4.1. Zapamiętywanie wartości entropii komórek

Entropia w ujęciu teorii informacji jest opisywana następującym wzorem:

$$H(X) = \sum_{i=1}^n p(x_i) \log \frac{1}{p(x_i)} = - \sum_{i=1}^n p(x_i) \log p(x_i) \quad (3)$$

gdzie x_i odpowiada konkretnemu stanowi, a $p(x_i)$ określa prawdopodobieństwo jego wystąpienia.

Obliczanie tej wartości na nowo dla każdej komórki jest czasochłonne, szczególnie że nie może się ona zmienić, jeśli komórka nie przeszła procesu redukcji osiągalnych stanów. Pierwszym sposobem na optymalizację podkroku entropii jest zatem ograniczenie liczby wywołań procedur związanych z obliczaniem entropii poszczególnych pól.

Implementacja tego usprawnienia może zostać wykonana na kilka sposobów. W przypadku prezentowanego algorytmu została wykorzystana pomocnicza tablica przechowująca obliczoną wartość entropii każdej komórki.

4.1.1. Implementacja

W celu wdrożenia opisanej optymalizacji wprowadzono następujące zmiany w kodzie:

- Stworzono jednowymiarową tablicę pomocniczą, która przechowuje wartości entropii poszczególnych komórek.
- Zmodyfikowano krok obserwacji, wprowadzając w tablicy pomocniczej aktualizację wartości odpowiadającej obserwowanej komórce. W momencie dokonania obserwacji do tablicy zostaje wpisana maksymalna wartość osiągnięta przez zmienianą typ *float*².
- Rozszerzono krok propagacji o logikę, która aktualizuje w tablicy wartość odpowiadającą obecnie przetwarzanej komórce. Dzieje się to wtedy i tylko wtedy, gdy liczba stanów osiągalnych przez komórkę ulegnie zmniejszeniu.
- Zmieniono krok wyliczania entropii, implementując użycie tablicy pomocniczej według algorytmu 3.

```
float lowestEntropy = entropyCache.Min();

if (lowestEntropy == float.MaxValue) {
    // Zwrócenie wartości pustej oznacza osiągnięcie stabilnego systemu
    return null;
}

int randomIndex = Enumerable.Range(0, elementCount).Where(i =>
    entropyCache[i] == lowestEntropy).Random();
return cells[randomIndex];
```

Algorytm 3. Optymalizacja kroku obliczania entropii i wyboru komórki

4.1.2. Czas wykonania

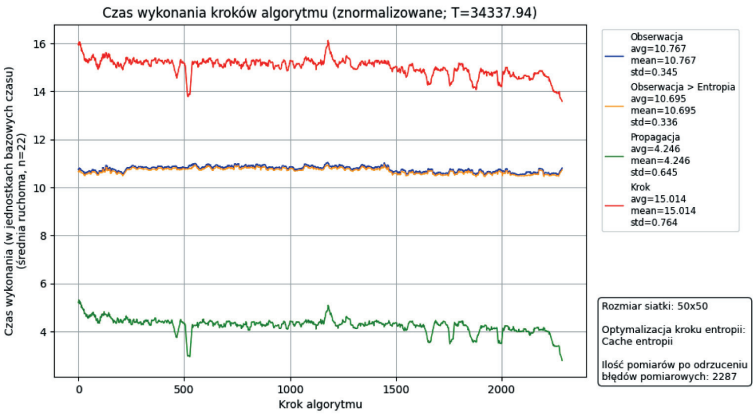
Podobnie jak w przypadku algorytmu w wersji bazowej algorytm z zastosowaniem tablicy pomocniczej do przechowywania obliczonych wartości entropii poszczególnych komórek został poddany testom generowania treści, których czasy wykonania zostały zebrane w tabeli 2.

² Warto zadać pytanie, dlaczego nadawana jest taka wartość entropii, jeśli obserwowana komórka ma określony stan, a więc jej entropia równa jest 0. Jest to uproszczenie, które pozwala na łatwiejsze znalezienie komórki z minimalną entropią. Komórki, które zostały poddane obserwacji, są efektywnie usuwane z puli poprzez nadanie im wartości entropii, która jest nie mniejsza niż wartość entropii minimalnej w całym systemie.

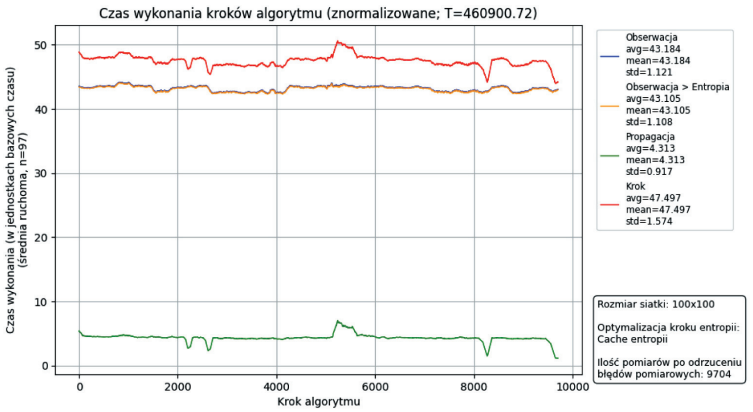
Tabela 2
Czas wykonania algorytmu WFC z zastosowaniem zapamiętywania entropii komórek

Wysokość siatki	Szerokość siatki	Czas wykonania (w jednostkach bazowych czasu)
50	50	34 337,94
100	100	460 900,72
200	200	6 354 468,99

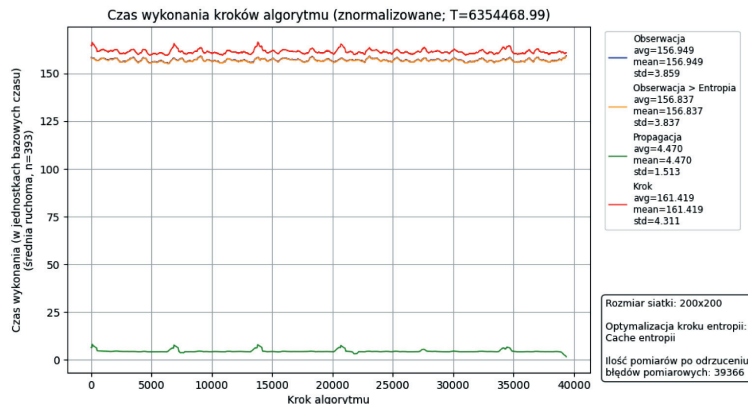
Dodatkowo zmierzono czas wykonania poszczególnych kroków algorytmu, a rezultaty przedstawiono na rysunkach 4–6.



Rys. 4. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 50 × 50, optymalizacja z zastosowaniem tablicy pomocniczej do przechowywania wartości entropii komórek)



Rys. 5. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 100 × 100, optymalizacja z zastosowaniem tablicy pomocniczej do przechowywania wartości entropii komórek)



Rys. 6. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 200×200 , optymalizacja z zastosowaniem tablicy pomocniczej do przechowywania wartości entropii komórek)

Na rysunkach 4–6 widoczne jest wyraźne zmniejszenie czasu wykonania kroku obliczania entropii, jak również tendencja do utrzymywania stałego czasu egzekucji tej procedury wraz z postępem algorytmu dla danego systemu. Warto jednak zauważyć, że rozwiązanie to nadal jest obciążone istotną wadą: udział obliczeń związanych z entropią w całkowitym czasie iteracji algorytmu diametralnie wzrasta wraz ze zwiększaniem liczby komórek w systemie.

4.2. Zapamiętywanie indeksów komórek z najniższą entropią

Warto ponownie zastanowić się nad zagadnieniem entropii komórki oraz krokiem propagacji. Należy podkreślić, że entropia komórki zależna jest wyłącznie od liczby stanów, które może ona przyjąć, a liczba ta może zostać zredukowana tylko podczas procedury propagacji. Dzięki temu możliwe jest wykonanie dodatkowych optymalizacji, z których najistotniejsze to:

- zapamiętywanie minimalnej entropii w systemie;
- zapisywanie komórek, które osiągają minimalną entropię.

4.2.1. Implementacja

Wprowadzenie optymalizacji polegało na zdefiniowaniu dwóch nowych zmiennych pomocniczych oraz zmodyfikowaniu dwóch procedur:

- Procedura propagacji została rozszerzona o logikę zapamiętywania minimalnej wartości entropii oraz komórek, które ją osiągają, według algorytmu 4.
- Procedura obliczania minimalnej entropii oraz wyboru komórki została zastąpiona nowymi instrukcjami według algorytmu 5.

```

float entropy = entropyCache[index];

if (entropy < systemLowestEntropy) {
    systemLowestEntropy = entropy;
    lowestEntropyIndices.Clear();
}

if (entropy != float.MaxValue && entropy == systemLowestEntropy) {
    lowestEntropyIndices.Add(index);
}

```

Algorytm 4. Optymalizacja kroku propagacji

```

if (lowestEntropyIndices.Count == 0) {
    return GetLowestEntropyOld();
}

float lowestEntropy = systemLowestEntropy;

if (lowestEntropy == float.MaxValue) {
    return null;
}

int randomIndex = lowestEntropyIndices.Random();
lowestEntropyIndices.Remove(randomIndex);

if (lowestEntropyIndices.Count == 0) {
    systemLowestEntropy = float.MaxValue;
}

return cells[randomIndex];

```

Algorytm 5. Optymalizacja procedury obliczania minimalnej entropii i wyboru komórki

Warto zwrócić uwagę na dwa ważne elementy nowej logiki. Po pierwsze, w przypadku pustego zbioru indeksów o minimalnej entropii algorytm zwraca taki sam wynik, jaki zwracał przed zmianą. Po drugie, opróżnienie tablicy pomocniczej skutkuje przywróceniem minimalnej entropii do maksymalnej wartości typu *float*. Gwarantuje to, że każda nowa wartość obliczona w kroku propagacji zostanie uznana za globalne minimum. Podczas implementacji tej optymalizacji wykorzystano oczywiście elementy wprowadzone w ramach poprzedniego usprawnienia.

4.2.2. Czas wykonania

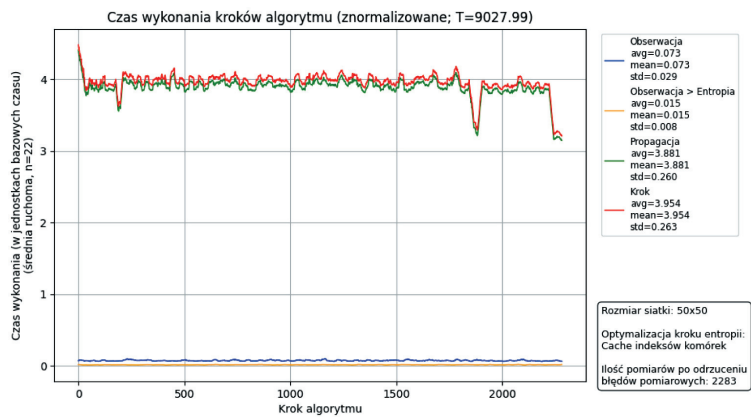
W celu weryfikacji finalnej propozycji usprawnienia (z wykorzystaniem listy do zapamiętywania indeksów komórek osiągających minimalną entropię) algorytm został poddany testom generowania treści, których czasy egzekucji zostały zebrane w tabeli 3.

Przedstawione na rysunkach 7–9 czasy wykonania poszczególnych kroków wyraźnie pokazują, skąd wzięło się tak znaczne przyspieszenie algorytmu.

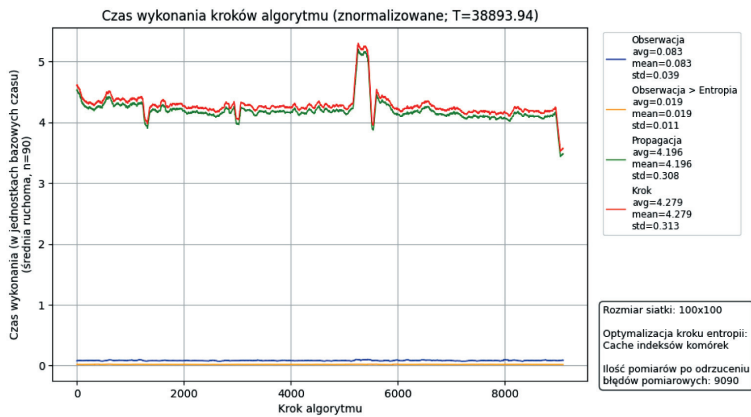
Tabela 3

Czas wykonania algorytmu WFC z zapamiętywaniem indeksów komórek osiagających minimalną wartość entropii

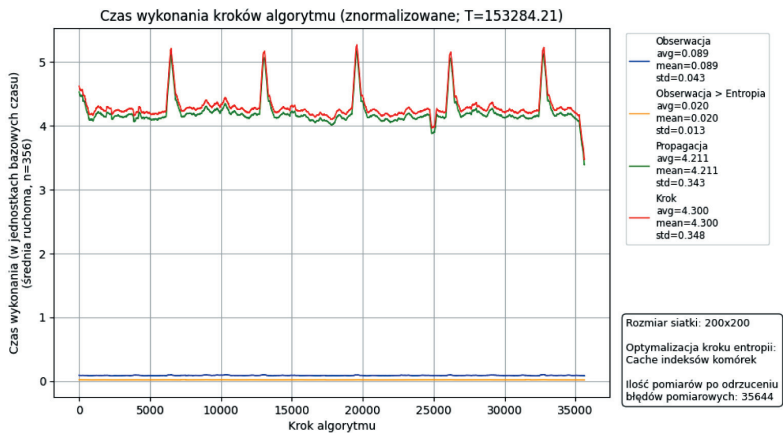
Wysokość siatki	Szerokość siatki	Czas wykonania (w jednostkach bazowych czasu)
50	50	9 027,99
100	100	38 893,94
200	200	153 284,21



Rys. 7. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 50 × 50, optymalizacja z zapamiętywaniem indeksów komórek osiagających minimalną wartość entropii)



Rys. 8. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 100 × 100, optymalizacja z zapamiętywaniem indeksów komórek osiagających minimalną wartość entropii)



Rys. 9. Czas wykonania poszczególnych kroków algorytmu WFC (siatka 200 × 200, optymalizacja z zapamiętywaniem indeksów komórek osiągających minimalną wartość entropii)

Po raz pierwszy zdarzają się przypadki, w których to krok propagacji zajmuje większą część czasu obliczeń w danym kroku algorytmu, a wykonanie kroku obserwacji wymaga nieznacznej ilości czasu.

5. Analiza wyników

Wyniki działania algorytmu w formie bazowej oraz po zaimplementowaniu zaproponowanych optymalizacji zostały zestawione w tabeli 4, która dokładnie obrazuje skrócenie czasu wykonania procedur w każdym z wariantów zmodyfikowanego algorytmu. Wyraźnie widać, że wariant z finalną optymalizacją charakteryzuje się najwyższą wydajnością, a jego przewaga tylko powiększa się wraz ze zwiększaniem rozmiaru siatki.

Tabela 4
Porównanie czasu wykonania poszczególnych wariantów algorytmu WFC

Rozmiar siatki	Wariant algorytmu	Czas wykonania (w jednostkach bazowych czasu)	Krotność poprawy szybkości działania (względem algorytmu bazowego)
50 × 50	podstawowy	58 879,21	1,000
	cache entropii	34 337,94	1,715
	cache indeksów	9 027,99	6,522

Tabela 4 cd.

100 × 100	podstawowy	806 258,64	1,000
	cache entropii	460 900,72	1,749
	cache indeksów	38 893,94	20,730
200 × 200	podstawowy	17 580 450,19	1,000
	cache entropii	6 354 468,99	2,767
	cache indeksów	153 284,21	114,692

Warto również przeanalizować zestawienie średniego czasu wykonania pojedynczego kroku algorytmu dla każdego z jego wariantów przedstawione w tabeli 5. W przypadku finalnej optymalizacji widoczna jest tendencja do osiągania niemalże stałego czasu wykonania pojedynczego kroku algorytmu niezależnie od liczby komórek. Oznacza to, że możliwe jest oszacowanie czasu wykonania algorytmu dla siatki dowolnych rozmiarów.

Tabela 5

Porównanie czasu wykonania pojedynczego kroku dla różnych wariantów algorytmu

Wariant algorytmu	Rozmiar siatki	Czas wykonania (w jednostkach bazowych czasu)	Średni czas wykonania jednego kroku (w jednostkach bazowych czasu)
Podstawowy	50 × 50	58 879,21	23,838
	100 × 100	806 258,64	81,605
	200 × 200	17 580 450,19	440,392
Cache entropii	50 × 50	34 337,94	15,014
	100 × 100	460 900,72	47,496
	200 × 200	6 354 468,99	161,420
Cache indeksów	50 × 50	9 027,99	3,954
	100 × 100	38 893,94	4,279
	200 × 200	153 284,21	4,300

6. Podsumowanie i dalsze badania

Analiza czasu wykonania algorytmu WFC z rozbiciem na poszczególne procedury pozwoliła zaproponować dwie techniki optymalizacji oraz zbadać ich skuteczność.

Zoptymalizowany algorytm (wykorzystujący tablicę pomocniczą do przechowywania entropii poszczególnych komórek oraz zapamiętujący indeksy komórek osiągających minimalną jej wartość) cechował się o wiele wyższą wydajnością niż algorytm w podstawowej wersji, a ponadto wykazywał tendencję do utrzymywania stałego czasu wykonania, który był wprost proporcjonalny do liczby komórek w systemie.

Choć przedstawione techniki optymalizacji pozwalają na wielokrotne przyspieszenie algorytmu, nadal możliwe jest dodatkowe skrócenie czasu obliczeń wymaganych do wygenerowania treści. Można to osiągnąć, modyfikując sposób przeprowadzania procedury propagacji, na którą w finalnym wariancie algorytmu przypadła największa ilość czasu wykonania (o wiele więcej niż na krok obserwacji). Wartym zbadania sposobem optymalizacji jest wykorzystanie programowania dynamicznego w celu przechowywania rezultatów procesu redukcji stanów w zależności od stanów wejściowych.

Literatura

- Beukman M., Ingram B., Liu I., Rosman B., 2023, *Hierarchical WaveFunction collapse*, Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, vol. 19, no. 1, s. 23–33. <https://www.doi.org/10.1609/aiide.v19i1.27498>.
- Farrokhi Maleki M., Zhao R. (2024). *Procedural content generation in games: A survey with insights on emerging LLM integration*, Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, vol. 20, no. 1, s. 167–178. <https://doi.org/10.1609/aiide.v20i1.31877>.
- Fehr M., Courant N., 2018, *fast-wfc*, GitHub. <https://github.com/math-fehr/fast-wfc> [dostęp: 30.11.2025].
- Gumin M., 2020, *WaveFunctionCollapse*, GitHub. <https://github.com/mxgmn/WaveFunctionCollapse> [dostęp: 4.06.2025].
- Karth I., Smith A.M., 2017, *WaveFunctionCollapse is constraint solving in the wild*, [w:] Canossa A., Hartevelde C., Zhu J., Sicart M., Deterding S. (eds.), *Proceedings of the 12th International Conference on the Foundations of Digital Games. FDG '17. Cape Cod, MA, August 14–17, 2017*, Association for Computing Machinery. <https://doi.org/10.1145/3102071.3110566>.
- Lazaridis L., Fragulis G.F., 2024, *Creating a Newer and Improved Procedural Content Generation (PCG) Algorithm with Minimal Human Intervention for Computer Gaming Development*, Computers, vol. 13, iss. 11, art. 304. <https://www.doi.org/10.3390/computers13110304>.
- Smith G., 2021, *The future of procedural content generation in games*, Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, vol. 10, no. 3, s. 53–57. <https://www.doi.org/10.1609/aiide.v10i3.12748>.

- Togelius J., Yannakakis G.N., Stanley K.O., Browne C., 2011, *Search-based procedural content generation: a taxonomy and survey*, IEEE Transactions on Computational Intelligence and AI in Games, vol. 3, iss. 3, s. 172–186. <https://www.doi.org/10.1109/tciaig.2011.2148116>.
- Villar López M.B.V., Chover M., 2025, *Procedural generation of 3D maps with Wave Function Collapse: optimization and advanced constraints*, [w:] Fellner D.W. (ed.), *XXXIV Spanish Computer Graphics Conference, Jaén, Spain, June 2–4, 2025*, The Eurographics Association. <https://www.doi.org/10.2312/ceig.20251107>.

Rozwój sterowników napędów elektrycznych i ich implementacja w łożyku planetarnym Kalman

Jakub Kutek¹ , Łukasz Kwinta² , Mariusz Gibiec¹ 

¹ Akademia Górniczo-Hutnicza, Wydział Inżynierii Mechanicznej i Robotyki, AGH Space Systems, Kraków

² Akademia Górniczo-Hutnicza, Wydział Informatyki, AGH Space Systems, Kraków

Streszczenie: Artykuł przedstawia proces projektowania oraz wdrażania uniwersalnych sterowników napędów elektrycznych przeznaczonych do łożyka planetarnego Kalman, opracowanego przez zespół AGH Space Systems. Celem było zastąpienie przestarzałych kontrolerów napędów nowoczesnymi, wydajnymi jednostkami wspierającymi zarówno silniki szczotkowe prądu stałego DC (*direct current*), jak i bezszczotkowe BLDC (*brushless direct current*) oraz silniki synchroniczne z magnesem trwałym PMSM (*permanent magnet sync motor*).

Wyzwania konstrukcyjne obejmowały opracowanie uniwersalnego i modułowego rozwiązania sprzętowego opartego na nowoczesnym układzie gate drivera (Infineon MOTIX™ 6EDL7141), który zapewnia elastyczną konfigurację przez interfejs SPI oraz liczne mechanizmy zabezpieczające. Projekt zakładał stworzenie jednej, powtarzalnej płytki wykonawczej kontrolera, współpracującej z wymiennymi płytkami matkami dopasowanymi do konkretnych aplikacji w łożyku.

Oprogramowanie zostało zaprojektowane w sposób modułarny, z niezależnymi bibliotekami dla każdego układu peryferyjnego. Zastosowano również zaawansowany algorytm sterowania FOC (*field oriented control*), który zapewnia płynne i wydajne sterowanie silnikami trójfazowymi, znacząco przewyższające tradycyjną metodę komutacji sześciokrokowej (*six step commutation*).

Rozwiązanie znacząco poprawia niezawodność, skalowalność i łatwość konserwacji łożyka, a jego uniwersalność umożliwia wykorzystanie w różnych komponentach mobilnych, w tym napędzie głównym, modułach skrętnych czy manipulatorach. Artykuł podkreśla znaczenie modularności w projektowaniu złożonych systemów mechatronicznych oraz znaczenie precyzyjnego sterowania w aplikacjach terenowych o wysokim stopniu nieprzewidywalności.

Słowa kluczowe: sterowniki napędów elektrycznych, interfejsy komunikacyjne, algorytm FOC, łożyk planetarny

DEVELOPMENT OF ELECTRIC DRIVE CONTROLLERS AND THEIR IMPLEMENTATION IN THE KALMAN PLANETARY ROVER

Abstract: The article presents the process of designing and implementing universal electric drive controllers intended for the “Kalman” planetary rover, developed by the AGH Space Systems team. The goal was to replace outdated drive controllers with modern, efficient units supporting both brushed DC (Direct Current) motors and brushless BLDC (BrushLess Direct Current) motors, as well as permanent-magnet synchronous motors PMSM (Permanent Magnet Synchronous Motor).

The design challenges included developing a universal and modular hardware solution based on a modern gate-driver system (Infineon MOTIX™ 6EDL7141), which provides flexible configuration through an SPI interface and numerous protection mechanisms. The project assumed creating a single, repeatable controller power board that works with interchangeable motherboard modules tailored to specific rover applications. The software was designed in a modular manner, with independent libraries for each peripheral system. An advanced Field Oriented Control (FOC) algorithm was also implemented, providing smooth and efficient control of three-phase motors, significantly outperforming the traditional Six-Step Commutation method. The solution significantly improves the rover’s reliability, scalability, and ease of maintenance, while its universality enables its use across various mobile components, including the main drive, steering modules, and manipulators. The article highlights the importance of modularity in designing complex mechatronic systems and the significance of precise motor control in field applications characterized by a high degree of unpredictability.

Keywords: electric drive controllers, communication interfaces, FOC algorithm, planetary rover

https://doi.org/10.7494/978-83-68219-80-7_4

1. Wprowadzenie

Napęd łazika jest z definicji jednym z najważniejszych podsystemów w całej konstrukcji. Jest to cecha, która odróżnia pojazdy planetarne od stacjonarnych sond. Napęd musi być niezawodny i efektywny, aby umożliwić poruszanie się konstrukcji w trudnym i często nieznanym terenie.

Łazik Kalman jest robotem zbudowanym ściśle pod wymagania zawodów konstrukcji studenckich symulujących teren marsjański. Przykładem takich zawodów jest University Rover Challenge cechujący się najtrudniejszym terenem spośród wszystkich zawodów z tej serii (The Mars Society 2025). Zawody odbywają się na terenie ośrodka badawczego Mars Desert Research Station w Utah (Stany Zjednoczone), gdzie na co dzień m.in. odbywają się misje analogowe mające symulować rzeczywiste warunki panujące na Marsie (rys. 1).



Rys. 1. Trudny teren jednej z misji na zawodach University Rover Challenge 2024
Źródło: KSAF (fot. Katarzyna Sadowy)

Taki teren stanowi duże wyzwanie dla systemu jezdnego łazika. Przy poruszaniu się wśród przeszkód, takich jak duże kamienie i wydmy, występują duże wahania momentu siły obciążającej napęd oraz ryzyko uszkodzenia systemu w wyniku zablokowania się koła po zaczepieniu o przeszkodę.

Pierwotne kontrolery napędów elektrycznych wraz z silnikami DC w łaziku dobrze spełniały swoją funkcję, jednak obecnie jest to jeden z najstarszych komponentów w całej konstrukcji. Słabo udokumentowane i obciążone licznymi wadami, utrudniają-

cymi konserwację, zaczęły jednak sprawiać coraz większe trudności, co w połączeniu z rosnącymi wymaganiami zawodów z serii *rover challenge* przyczyniło się do narodzin pomysłu na zaprojektowanie nowych kontrolerów wspierających także silniki bezszczotkowe prądu stałego BLDC (*brushless direct current*), które miałyby zastąpić wysłużone silniki szczotkowe w modułach skrętów łazika Kalman.

2. Uniwersalne podejście

W łaziku Kalman wykorzystywane są obecnie dwa rodzaje silników – synchroniczne z magnesem trwałym PMSM (*permanent magnet sync motor*) w ramieniu robotycznym oraz DC w modułach jezdnych – zarówno dla systemu napędowego, jak i jezdnego. W najbliższych planach rozwoju konstrukcji zakładane jest wprowadzenie nowego systemu skrętnego wykorzystującego silniki BLDC, a także zaprojektowanie nowego elementu chwytneho manipulatora, również opartego na tym typie silników, co przekłada się na odpowiednie kontrolery do każdego komponentu. Stąd narodził się pomysł na uniwersalny kontroler do silników DC i BLDC (a także PMSM). Założeniem jest stworzenie kompaktowego modułu, który w zależności od zastosowania będzie wymagał jedynie lekkiej modyfikacji kształtu płytki lub dedykowanego adaptera do wyprowadzenia potrzebnego wejścia/wyjścia, unifikujące bazę komponentów dla zastosowań sterowania silnikami. Korzystanie z jednego rozwiązania, składającego się z takich samych komponentów elektronicznych, w wielu miejscach znacząco ułatwia utrzymanie konstrukcji i potencjalne naprawy. Redukuje także ilość potrzebnej wiedzy do przekazywania kolejnym pokoleniom studentów pracujących nad konstrukcją w przyszłości. Uniwersalny kontroler pozwala także na bezproblemową integrację zewnętrznych modułów, jak na przykład moduł wiertła powstający na zawody European Rover Challenge 2025 w Krakowie. Dzięki takiemu modułowi nie będzie potrzeby projektowania specjalnie dostosowanego kontrolera silników, ponieważ zostanie wykorzystane istniejące rozwiązanie.

Projektowanie uniwersalnego modułu nie jest prostym zadaniem. Pierwszym problemem było wybranie układu sterującego bramkami tranzystorów (ang. *gate driver*), który funkcjonowałby poprawnie i deterministycznie zarówno przy wykorzystaniu dwufazowych silników DC, jak i trójfazowych BLDC (oraz PMSM). Do realizacji tego zadania wybrano układ firmy Infineon MOTIX™ 6EDL7141. Układ ten zapewnia szeroką konfigurowalność dzięki wbudowanemu interfejsowi SPI oraz ponad pięćdziesięciu możliwym do zaprogramowania parametrom. Umożliwia to precyzyjne dostosowanie charakterystyki działania sterownika do konkretnego zastosowania, a tym samym optymalizację sprawności i niezawodności systemu. Dodatkowo układ zawiera zestaw narzędzi pozwalających na monitorowanie napięć i prądów silnika, temperatury urządzenia czy też wykrywanie blokady rotora oraz definiowanie poziomów zabezpieczeń, co zapewnia możliwość ich dostosowania do warunków aplikacyjnych. Wbudowane

wzmacniacze pomiarowe pozwalają na precyzyjne pomiary prądu, co jest niezbędne do prawidłowego działania algorytmów sterowania (Infineon 2025).

Dzięki tym cechom układ 6EDL7141 idealnie wpisuje się w koncepcję uniwersalnego sterownika, pozwalając na osiągnięcie wysokiej gęstości mocy, elastyczności oraz niezawodności przy minimalnym nakładzie projektowym na poziomie hardware'u.

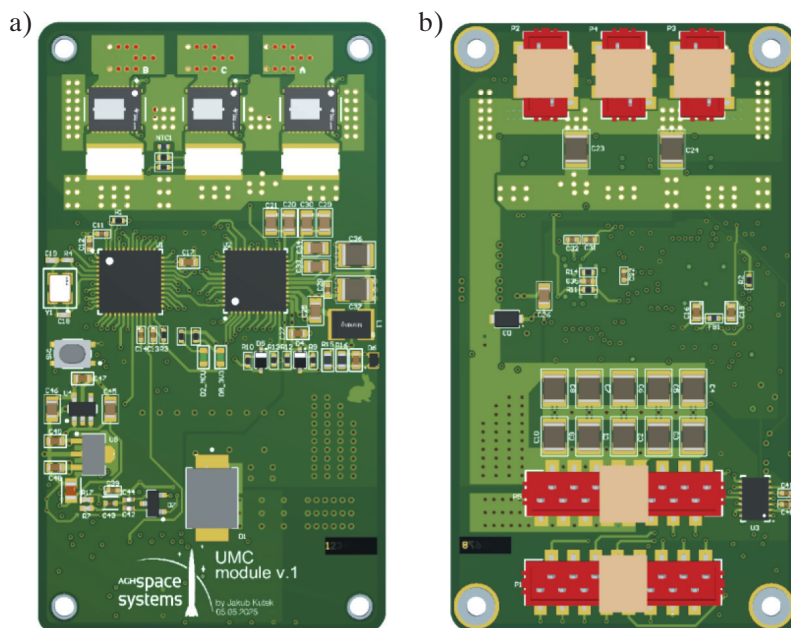
Kolejnym wyzwaniem jest zaplanowanie i obsługa wielu możliwych konfiguracji urządzeń pomiarowych – koniecznych do obsługi pętli kontroli dla sterowania silnikami. W zależności od zastosowania wymagane są różne urządzenia zapewniające informację zwrotną, np. dla napędu wystarczy informacja o aktualnej rzeczywistej prędkości obrotowej silnika, która może być realizowana przez prosty enkoder inkrementalny, ale już obsługa silnika w module skrętnym wymaga także informacji o aktualnej pozycji rotora, co może zapewnić tylko enkoder absolutny.

Uniwersalność jest wyzwaniem zarówno dla oprogramowania wbudowanego, jak i dla samego projektu płytki drukowanej. Główne założenia projektu płytki to zastosowanie jednego modułu dla jednego silnika oraz korzystanie z tego samego układu w różnych implementacjach, a także niezmienność układu pomiędzy różnymi implementacjami. Założenia te zrealizowano przez zaprojektowanie jednego rodzaju płytki drukowanej z pojedynczym trójfazowym sterownikiem silników, która może pracować tylko w połączeniu z płytką matką wyposażoną w odpowiednie złącza płytka-płytką. Dzięki temu pomiędzy poszczególnymi implementacjami zachowuje się jeden projekt płytki sterującej silnikami, a wymagana jest tylko zmiana płytki matki, do której można przyłączyć wiele pojedynczych kontrolerów i wejść dla urządzeń pomiarowych w zależności od wymagań implementacyjnych. Uniwersalność uzyskano dzięki modularnej budowie oprogramowania – dla każdego układu cyfrowego, z którym należy komunikować się z poziomu mikroprocesora, stworzono niezależną bibliotekę udostępniającą intuicyjny interfejs do sterowania urządzeniem. W ten sposób zbudowano kolejne warstwy abstrakcji, którymi łatwo zarządzać na najwyższym poziomie. Innym wyzwaniem było obsługa wielofunkcyjnych wyprowadzeń procesora, jednak dzięki bibliotekom HAL udostępnianym przez firmę STMicroelectronics (STMicroelectronics 2025) stworzono mechanizmy dynamicznie konfigurujące wejścia i wyjścia mikroprocesora w pożądanym sposób.

3. Projekt elektroniki

Projekt płytek drukowanych został opracowany zgodnie z założeniami modułowości oraz uniwersalności, co pozwala na elastyczne dostosowanie systemu sterowania do różnych typów silników. Głównym założeniem było stworzenie jednolitego modułu kontrolera silników różnych typów, który może współpracować z dedykowaną płytką matką (*carrier board*) zależnie od konkretnego zastosowania implementacyjnego.

Moduł uniwersalnego kontrolera (rys. 2) został zaprojektowany jako minimalna jednostka funkcjonalna zawierająca wszystkie wymagane elementy do sterowania napędem. W jego skład wchodzi: mikrokontroler (STM32G474CEU6), tranzystory mocy, układ nadawczo-odbiorczy (*transceiver*) magistrali CAN (TCAN1141), rezystory do pomiaru prądu faz, a także sekcja własnego zasilania.

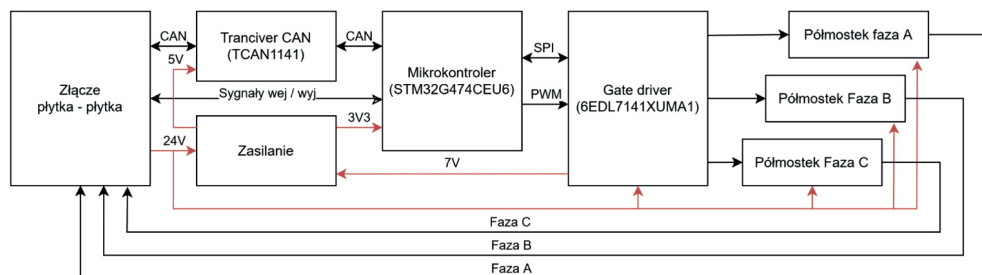


Rys. 2. Widok modułu kontrolera silników z góry (a) i dołu (b), gdzie widać złącza płytka-płytką

Dzięki takiemu podejściu możliwe jest tworzenie płytek matek do konkretnych aplikacji, z zachowaniem identycznej płytki wykonawczej dla silnika. Zmniejsza to złożoność projektową, umożliwia standaryzację produkcji i upraszcza testowanie.

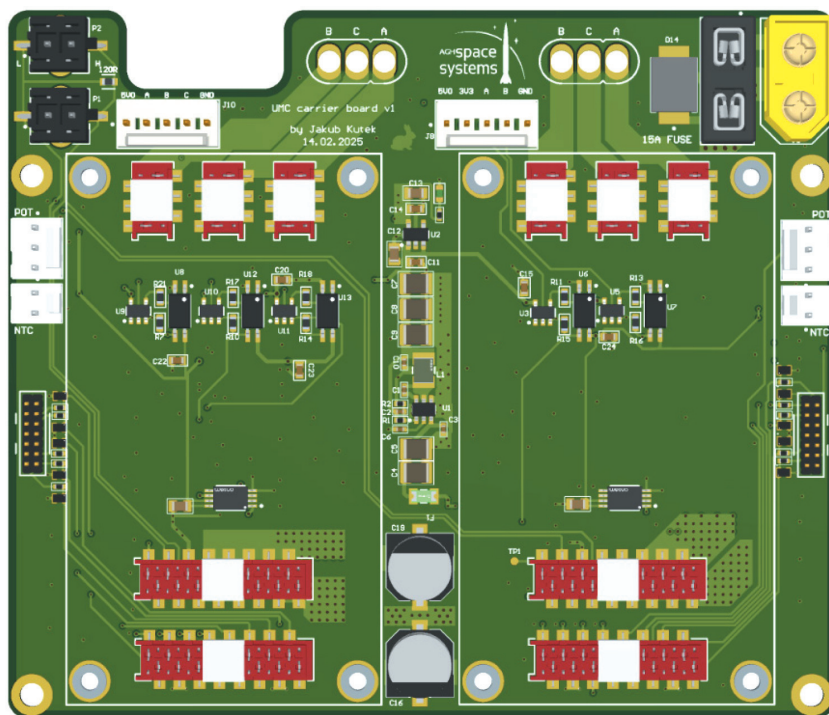
Dodatkowo na płycie bazowej znajduje się układ pamięci EEPROM dla każdego modułu kontrolera, w którym zapisywane są jego ustawienia. Rozwiązanie to umożliwia ujednolicenie procesu wersjonowania oprogramowania oraz szybką identyfikację i naprawę błędów konfiguracyjnych podczas uruchamiania systemu.

Schemat blokowy (rys. 3) przedstawia podział funkcjonalny modułu, w tym jego połączenia z płytą matką za pomocą odpowiednich konektorów. Połączenie to przenosi wszystkie sygnały wejściowe i wyjściowe (zasilanie, sygnały pomiarowe, magistrala CAN, wyjścia fazowe do silnika, magistrale danych) na płytkę bazową. Na niej realizowane są wszystkie zewnętrzne funkcjonalności, takie jak zabezpieczenia (ograniczające napięcia przejściowe TVS, bezpieczniki przetężeniowe PTC), wejścia dla enkoderów, termistorów oraz dedykowane zasilanie do wszystkich zewnętrznych elementów.



Rys. 3. Schemat blokowy modułu kontrolera

W obecnie rozwijanym zastosowaniu do modułów napędowych łazika potrzebne jest sterowanie silnikiem DC odpowiedzialnym za napęd koła, a także silnikiem bezszczotkowym odpowiedzialnym za skręt koła. Do odpowiedniego sterowania wymagana jest także informacja zwrotna od elementów wykonawczych. W tym celu została zaprojektowana odpowiednio dostosowana płytki bazowa (*baseboard*) (rys. 4). Posiada ona wyprowadzenia do podłączenia wymaganych komponentów, takich jak enkoder magnetyczny, enkodery optyczne, termistory do pomiaru temperatur silników czy złącza dla magistrali CAN. Wyposażono ją w niezbędne zabezpieczenia, które gwarantują bezawaryjną pracę modułów kontrolera silnika.



Rys. 4. Obecnie rozwijana płytki bazowa z miejscem na dwa moduły wykonawcze

4. Algorytm FOC

Sterowanie silnikami trójfazowymi to złożony problem. Rolę komutatora odgrywa w tym przypadku układ elektroniczny. Najpopularniejsze metody sterowania opierają się na estymacji pozycji rotora w czasie i odpowiednim sterowaniu prądem płynącym do cewek silnika. W układach cyfrowych sterowanie prądem można uzyskać przez zmianę napięcia skutecznego – wypełnienia sygnału PWM sterującego cewką silnika, zgodnie ze wzorami (1) i (2):

$$U_{sk} = \eta V \quad (1)$$

$$I_c = \frac{U_{sk}}{R_c} = \eta \frac{V}{R_c} = \eta \cdot const \quad (2)$$

$$I_c \propto \eta \propto U_{sk}$$

gdzie:

- U_{sk} – napięcie skuteczne,
- V – napięcie wejściowe,
- η – procent wypełnienia sygnału PWM,
- I_c – prąd uzwojenia silnika,
- R_c – rezystancja uzwojenia silnika.

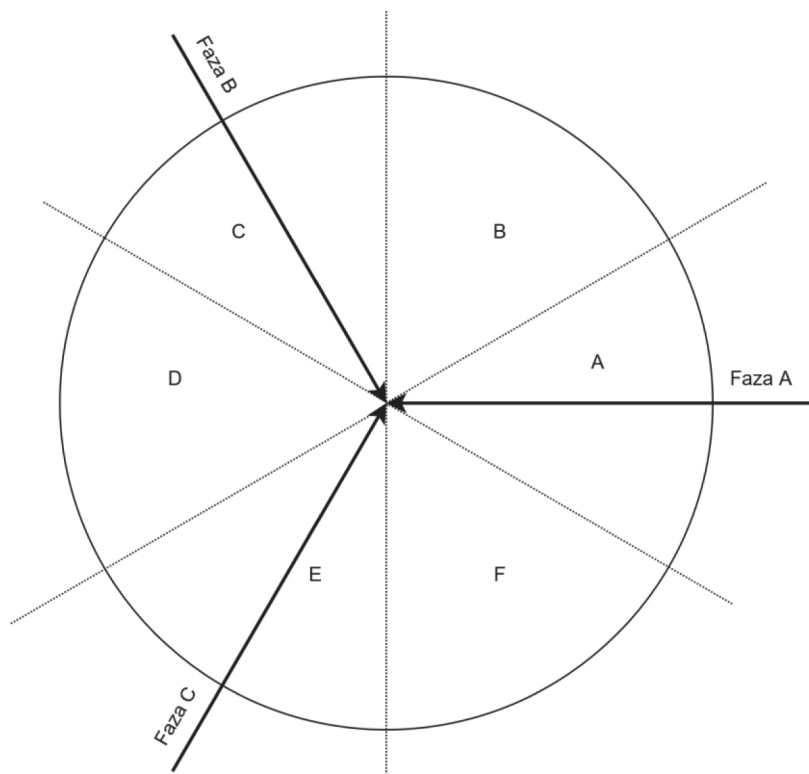
Podstawową metodą komutacji jest sześciokrokowy algorytm (ang. *six step commutation*), zawierający sześć przedziałów pozycji, w których może znajdować się rotor silnika – są one wyznaczone przez trzy proste poprowadzone przez każdą z trzech cewek silnika (MathWorks 2025a).

W każdym z obszarów określić dwie fazy (półmostki), które biorą udział w przyciąganiu rotora.

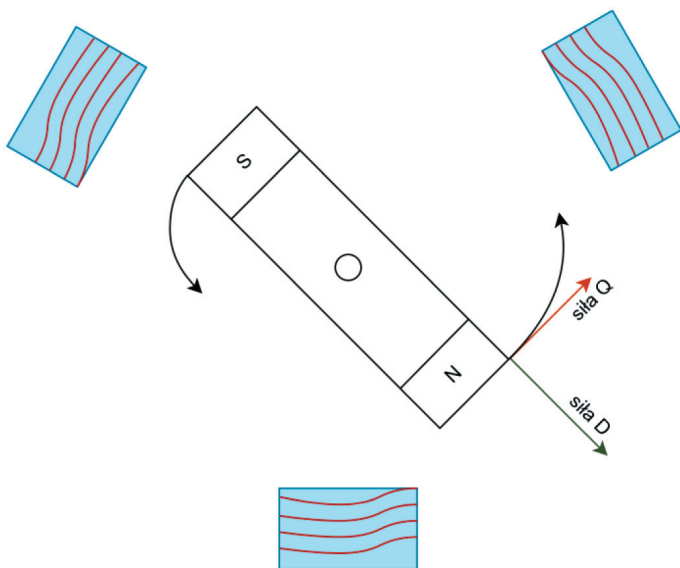
Jednak ten sposób komutacji nie jest optymalny. Dyskretne przejścia pomiędzy kolejnymi sektorami powodują duże straty momentu oraz brak płynności obrotu. Dodatkowo następuje utrata precyzji sterowania oraz obniżenie maksymalnej prędkości obrotowej.

Biorąc pod uwagę powyższe, w projekcie uniwersalnego kontrolera silników zaimplementowano algorytm FOC (*field oriented control*), który przez zaawansowany układ regulacji oraz przekształcenia matematyczne stara się optymalizować siły działające na rotor (MathWorks 2025b, Performance Motor Devices 2025).

Algorytm FOC wprowadza koncept prostopadłych sił działających w danej chwili na rotor. Są to siły D (ang. *direct*) i Q (ang. *quadrature*), które zaznaczono na rysunku 5. Siła D działa równolegle z osią rotora i próbuje go „rozciągać”. Ta siła jest niepożądana, więc w układzie regulacji jako jej wartość zadaną ustawia się 0. Natomiast siła Q obraca rotor i odpowiada za docelowy moment silnika.

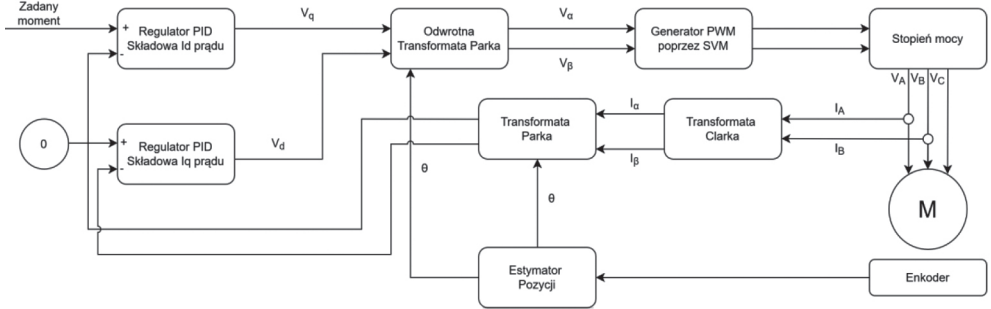


Rys. 4. Obszary w algorytmie *six step*



Rys. 5. Demonstracja sił działających na rotor silnika trójfazowego

Na rysunku 6 przedstawiono schemat pętli kontroli realizującej algorytm FOC. Aby uzyskać zadaną wartość momentu, wykorzystuje się dwa regulatory do kontroli wcześniej opisanych sił. Ich wyjście stanowią prądy (napięcia, zgodnie z równaniami powyżej zachodzi zależność wprost proporcjonalna) w ruchomym układzie odniesienia rotora. Optymalizowane są siły działające w układzie odniesienia rotora, lecz sterowanie polem magnetycznym odbywa się poprzez układ statora. W tym celu wykorzystuje się odwrotną transformatę Parka, która przekształca ruchomy układ odniesienia (siły – napięcia działające na rotor), na stacjonarny układ odniesienia – siły – napięcia działające na stator – uzwojenie silnika (MathWorks 2025c).



Rys. 6. Schemat działania algorytmu FOC

Na rysunku 7 pokazano oba układy odniesienia. Kąt obrotu zależy od czasu, zgodnie z zależnością (3), gdzie ω to prędkość kątowna rotora.

$$\theta = \omega t \quad (3)$$

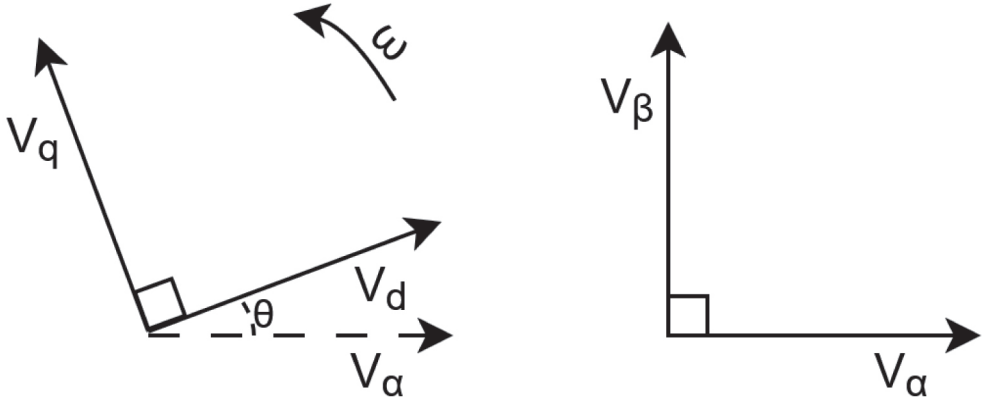
Odwrotna transformata Parka przekształca układy odniesienia zgodnie ze wzorem:

$$\begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} V_d \\ V_q \end{bmatrix} \quad (4)$$

Następnie wynik z odwrotnej transformaty Parka jest przekształcany na sygnały PWM przez modulację wektora przestrzennego (ang. *space vector modulation*) (Wikipedia 2025).

W ramach sprzężenia zwrotnego prąd z trójfazowego układu abc statora jest przekształcany na dwufazowy układ $\alpha\beta$ przez transformatę Clarka (MathWorks 2025d).

$$\begin{bmatrix} I_\alpha \\ I_\beta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix} \quad (5)$$



Rys. 7. Ruchomy układ odniesienia dq (po prawej) i stacjonarny $\alpha\beta$ po lewej

Ostatecznie przekształca się układ $\alpha\beta$ na układ ruchomy dq transformatą Parka (MathWorks 2025e).

$$\begin{bmatrix} I_d \\ I_q \end{bmatrix} = \begin{bmatrix} \sin \theta & -\cos \theta \\ \cos \theta & \sin \theta \end{bmatrix} \begin{bmatrix} I_\alpha \\ I_\beta \end{bmatrix} \quad (6)$$

Zastosowanie takiej metody sterowania pozwala uzyskać większą dokładność sterowania oraz wydajność.

5. Podsumowanie

Przedstawiony projekt uniwersalnego sterownika napędów wspiera zarówno silniki szczotkowe DC, jak i bezszczotkowe BLDC oraz PMSM. Rozwiązanie sprzętowe oparto na nowoczesnym układzie gate drivera (Infineon MOTIX™ 6EDL7141), który zapewnia elastyczną konfigurację przez interfejs SPI oraz liczne mechanizmy zabezpieczające. Pozwala to na wielokrotne wykorzystanie powtarzalnej płytki wykonawczej kontrolera, współpracującej z wymiennymi płytkami matkami dopasowanymi do konkretnych aplikacji. Zaimplementowane oprogramowanie zawiera biblioteki dla układów peryferyjnych oraz algorytm FOC zapewniający płynne i wydajne sterowanie silnikami trójfazowymi, znacząco przewyższający tradycyjną metodę komutacji. Uzyskana modularność rozwiązania pozwala na szybki rozwój złożonych systemów, takich jak łańcuk planetarny, przez jego wykorzystanie w napędzie głównym, modułach skrętnych, manipulatorach czy pokładowym laboratorium. Jej efektem jest także precyzja sterowania oraz niezawodność, skalowalność i łatwość konserwacji łańcika, co ma duże znaczenie w aplikacjach pracujących w trudnych i nieprzewidywalnych warunkach.

Literatura

- Infineon, 2025, MOTIX™ 6EDL7141 – *Dokumentacja techniczna*. https://www.infineon.com/dgdl/Infineon-6EDL7141-DataSheet-v01_20-EN.pdf?fileId=8ac78c8c8e-7ead30018ec7bd6aac46f3 [dostęp: 16.06.2025].
- MathWorks, 2025a, *Six Step Commutation – Documentation*. <https://www.mathworks.com/help/mcb/ref/sixstepcommutation.html> [dostęp: 17.06.2025].
- MathWorks, 2025b, *Field Oriented Control – Documentation*. <https://www.mathworks.com/help/mcb/gs/implement-motor-speed-control-by-using-field-oriented-control-foc.html> [dostęp: 17.06.2025].
- MathWorks, 2025c, *Inverse Park Transform – Documentation*. <https://www.mathworks.com/help/mcb/ref/inverseparktransform.html> [dostęp: 23.06.2025].
- MathWorks, 2025d, *Clarke Transform – Documentation*. <https://www.mathworks.com/help/mcb/ref/clarketransform.html> [dostęp: 23.06.2025].
- MathWorks, 2025e, *Park Transform – Documentation*. <https://www.mathworks.com/help/sps/ref/parktransform.html> [dostęp: 23.06.2025].
- Performance Motion Devices, 2025, *Field Oriented Control (FOC) – A Deep Dive*. <https://www.pmdcorp.com/resources/type/articles/get/field-oriented-control-foc-a-deep-dive-article> [dostęp: 17.06.2025].
- STMicroelectronics, 2025, *HAL – Hardware Abstraction Layer*. https://www.st.com/resource/en/user_manual/um2570-description-of-stm32g4-hal-and-lowlayerdrivers%E2%80%93stmicroelectronics.pdf [dostęp: 17.06.2025].
- The Mars Society, 2025, *University Rover Challenge*. <https://urc.marssociety.org/> [dostęp: 15.06.2025].
- Wikipedia, 2025, *Space Vector Modulation*. https://en.wikipedia.org/wiki/Space_vector_modulation [dostęp: 23.06.2025].

Oil and gas industry challenges in light of the new EU methane law

Łukasz Łukańko , Joanna Wartak 

AGH University of Krakow, Faculty of Drilling, Oil and Gas

Abstract: The implementation of the new European Union methane regulation presents significant challenges for oil and gas companies, requiring compliance with stricter emission control measures. The regulation mandates systematic monitoring, transparent reporting, and the reduction of methane leaks throughout the supply chain. Adherence to these requirements necessitates substantial investments in advanced detection technologies, infrastructure upgrades, and operational modifications to minimize emissions effectively. Moreover, companies must contend with increased compliance costs and the risk of financial penalties for non-compliance. Beyond the financial implications, these regulatory changes may influence corporate strategies, driving a transition toward more sustainable practices, such as methane capture and utilization. This article examines the primary challenges and strategic responses of oil and gas companies in adapting to the evolving regulatory framework during oil and gas exploration.

Keywords: methane emission, methane regulations, drilling, production test

WYZWANIA PRZEMYSŁU NAFTOWEGO I GAZOWNICZEGO W ŚWIELE NOWEGO ROZPORZĄDZENIA METANOWEGO

Streszczenie: Wdrożenie nowego rozporządzenia Unii Europejskiej dotyczącego emisji metanu stanowi poważne wyzwanie dla przedsiębiorstw sektora naftowego i gazowniczego, ponieważ wymaga dostosowania się do restrykcyjnych wymogów w zakresie kontroli emisji. Rozporządzenie nakłada obowiązek systematycznego monitorowania, transparentnego raportowania oraz redukcji wycieków metanu na wszystkich etapach łańcucha dostaw. Realizacja tych wymagań wymaga znaczących nakładów inwestycyjnych w zaawansowane technologie detekcji, modernizację infrastruktury oraz dostosowanie procesów operacyjnych w celu skutecznego ograniczenia emisji metanu. Ponadto przedsiębiorstwa muszą uwzględnić wzrost kosztów związanych z zapewnieniem zgodności regulacyjnej oraz ryzyko sankcji finansowych w przypadku jej naruszenia. Oprócz wymiaru ekonomicznego, nowe uwarunkowania regulacyjne mogą wpływać na kierunki strategiczne przedsiębiorstw, stymulując transformację w kierunku bardziej zrównoważonych praktyk, w tym technologii wychwytywania i wykorzystania metanu. Niniejszy artykuł analizuje kluczowe wyzwania oraz strategiczne reakcje przedsiębiorstw sektora naftowego i gazowego na zmieniające się ramy regulacyjne w obszarze poszukiwania i wydobywania węglowodorów.

Słowa kluczowe: emisja metanu, regulacje dotyczące metanu, wiercenie, testy produkcyjne

https://doi.org/10.7494/978-83-68219-80-7_5

1. Introduction

Methane is a potent greenhouse gas, with a global warming potential approximately 80 times greater than carbon dioxide over a 20-year period (IPCC 2021). Recognizing its significant role in climate change, the European Union (IEA 2023) has introduced stringent regulations to reduce methane emissions in the oil and gas sector. These measures aim to curb fugitive emissions, eliminate routine venting and flaring, and establish comprehensive monitoring systems to enforce compliance (European Commission 2024).

Methane emissions in the oil and gas sector come from several sources (Saunois et al. 2020). One major source is fugitive emissions, which result from leaks in pipelines, valves, and storage tanks (IEA 2022). Additionally, operational releases occur during various stages of exploration, extraction, processing and transportation (UNEP 2021). Another significant contributor is venting and flaring, where excess gas is either released directly into the atmosphere or burned (World Bank 2023).

The exploration, production, processing, and transportation of hydrocarbons, require a comprehensive regulatory framework to ensure environmental protection, operational safety, and resource management. This regulation applies to various aspects of the oil, gas defining their scope and limitations.

2. New EU methane regulations

The regulatory framework covers several key areas related to methane management in the oil and gas sector. It applies to the exploration and production of hydrocarbons, including oil and natural gas extraction, as well as the gathering and processing of natural gas. These activities involve geological surveys, drilling, and extraction techniques that must adhere to strict environmental and safety standards.

Additionally, the regulation addresses the management of wells, extending to idle wells, temporarily plugged wells, and permanently plugged and abandoned wells. Proper closure and ongoing monitoring of these wells are essential to prevent environmental hazards such as groundwater contamination and methane emissions (Jackson et al. 2013).

The framework also includes oversight of the transmission and distribution of natural gas. While metering systems at end-use points and connections between the distribution network and metering systems on land owned by end-users are excluded, underground storage and LNG facility operations remain subject to industry standards and governmental regulations (IEA 2021).

The new regulation enforces strict methane management measures across the EU's oil and gas sector, requiring companies to comply with four key obligations. First, operators involved in exploration, production, processing, transmission, distribution, underground storage, and liquefied natural gas operations must implement rigorous

measurement, reporting, and verification systems. They are also required to submit annual methane emissions reports aligned with standardized frameworks such as the Oil and Gas Methane Partnership 2.0 – OGMP 2.0 (EY 2024).

Second, the regulation mandates leak detection and repair programs. Operators must conduct regular methane leak detection surveys using advanced technologies like infrared cameras and satellite monitoring.

Third, restrictions on venting and flaring are imposed, prohibiting routine methane releases except in emergency situations or when operationally unavoidable. To minimize emissions, operators are encouraged to adopt alternative gas utilization strategies, such as reinjection or on-site energy use (Mondaq 2024).

Finally, the regulation introduces specific requirements for monitoring and mitigating emissions from inactive, temporarily plugged, and permanently abandoned wells, ensuring that even non-operational sites adhere to methane reduction standards (Schlumberger 2024).

The regulation introduces a phased approach to compliance, gradually increasing obligations for operators and importers. It officially comes into force in July 2024, marking the beginning of the transition period. By 2025, operators must start submitting annual methane emissions reports, while importers will be required to report emissions data from their foreign suppliers.

From January 2027, all new fossil fuel import contracts must comply with EU measurement, reporting, and verification standards. By 2028, importers must report the methane intensity of the fossil fuels they bring into the EU. Finally, by 2030, maximum methane intensity values for both domestic production and imports will become legally enforceable, ensuring stricter emissions control across the sector (Mondaq 2024).

3. Emission from drilling operations

The process of exploring and extracting oil and gas begins with well identification, drilling, and testing. Drilling, a crucial step in well construction, involves excavating circular wells and is carried out during exploration and field development.

During drilling, gases – particularly from the drilling mud – are vented into the atmosphere. As the hole is drilled, the drilling mud flows through the mud treatment system, releasing methane into the atmosphere. In the first stage of cleaning, when the mud shakers are used, most methane is vented into the atmosphere due to the shaking process. Next, the mud flows through the mud tanks to the mud pump and back into the well.

During drilling operations amount of methane vented from drilling mud to atmosphere depends mainly from following factors:

- drilling rate,
- thickness of the reservoir rock,
- natural gas structuration's.

Other activities where methane is venting to the atmosphere is well testing. During well test activates natural gas flows from well and mainly it burn using the flare. The methane emission from well testing will be depends on:

- gas flow,
- gas content,
- time of well test,
- flare efficiency.

4. Reporting of methane emissions from drilling operations

According to the methane regulation, operators must submit reports to competent authorities detailing methane emissions at the source level (European Parliament 2024). The quantification must be conducted using, at minimum, overall emission factors for all identified sources. This reporting requirement aims to standardize emissions measurement and facilitate regulatory oversight (European Commission 2023).

Until the adoption of relevant implementing acts, operators and companies are required to follow the OGMP 2.0 technical guidelines and reporting templates. These guidelines, developed by the United Nations Environment Programme and the Climate & Clean Air Coalition, provide a comprehensive framework for methane emissions reporting across upstream, midstream, and downstream activities (UNEP 2022).

Due to the OGMP 2.0 vented methane emissions along the mud separator train including mud degassing using the emission factors shown, as an example, in the Table 1.

Table 1
Methan emission from drilling operations (UNEP 2022)

Emission sources	Methane Emission Factor
Offshore wells (tCH ₄ /drilling day)	
Water-based mud	0.2605
Oil-based mud	0.0586
Synthetic mud	0.586
Onshore wells (tCH ₄ /drilling day)	
Water-based mud	0.0458
Oil-based mud	0.0103
Synthetic mud	0.0103

For drilling day calculations, only the days when drilling operations are conducted in reservoir rocks should be considered, rather than the entire duration of drilling activities.

To calculate the amount of methane flared during well testing, OGMP outlines specific steps to follow. The first step in methane emissions calculation is determining flare efficiency. Measured combustion efficiency factors may also be used, recognizing that they provide a more conservative reported value compared to destruction efficiency. The following is a non-exhaustive list of emerging measurement techniques:

- passive optical gas measurement systems:
 - video imaging spectral radiometry,
 - passive fourier transform infrared;
- active optical gas measurement systems:
 - open-path OP-FTIR,
 - differential absorption lidar.

Additionally, process simulation models based on representative flare systems and operating/environmental conditions, validated through direct measurements and engineering calculations, can be employed to determine flare destruction efficiency. Experimental studies, including wind tunnel tests, have demonstrated that waste-gas velocity at the flare tip and external environmental factors (such as wind speed, atmospheric pressure, and relative humidity) significantly impact destruction efficiency.

A parametric model can be applied to estimate destruction efficiency in flares, incorporating these environmental and operational parameters to enhance accuracy (UNEP 2021).

$$\eta = 1 - \frac{A^3 \cdot \exp(B \cdot \frac{U_\infty}{(g \cdot V_j \cdot d_o)^{\frac{1}{3}}})}{(LHV_{\text{mass}})^3} \quad (1)$$

where:

- η – carbon conversion efficiency,
- LHV_{mass} – lower heating value (MJ/kg),
- A – fuel-dependent (for natural gas: 156.4 MJ/kg),
- B – fuel-dependent coefficient (for natural gas: 0.318),
- g – gravitation acceleration (9.81 m/s^2),
- V_j – flare tip exit velocity (m/s),
- d_o – stack outside diameter (m),
- U_∞ – wind speed (m/s).

Having calculated the carbon conversion efficiency, the total amount of methane emission can be determined using the following equation (UNEP 2021):

$$F_{\text{CH}_4} = \eta \cdot V_g \cdot T \quad (2)$$

where:

- F_{CH_4} – methane emission from flaring operations (m^3),
- η – carbon conversion efficiency,
- V_g – mean flow/flux amount of flared of the natural gas (m^3/h),
- T – time of flaring (h).

The new EU methane regulations also allow the use of emission factors for flaring activities during well test operations to determine methane emissions. For this purpose, OGMP recommends using API factors – Tables 2 and 3 (API 2009, 2021).

Table 2
Flaring emission factors for specific operations (API 2021)

Segment/source	CH ₄ emission factor	CO ₂ emission factor	N ₂ O emission factor
	(kg/event)		
Flaring emission factors for petroleum systems			
Well testing	384.0	125,682.3	0.2
Well testing (hydraulic fracturing)	1,851.2	347,390.4	1.2
Flaring emission factors for natural gas systems			
Well testing	8,246.7	1,808,851.2	1.3
Well testing (hydraulic fracturing)	3,004.2	360,024.3	3.9

Simplified emission factors are provided in Table 3 for use when only total throughput data is available (API 2009). IPCC also provides CO₂ and CH₄ emission factors for the same flare sources. These flare emission factors are based on the volume of production or throughput for different types of petroleum operations.

Flare combustion efficiency measures the conversion of hydrocarbons to CO₂, expressed as the percentage of carbon in the flare feed gas that is fully oxidized to CO₂. Flare destruction efficiency, on the other hand, represents the percentage of a specific compound in the flare feed gas that is converted into another compound, particularly methane conversion to CO₂.

Table 3
GHG emission factors for gas flares (API 2009)

Flare sources	Emission factors		
	CO ₂ [*]	CH ₄ [*]	N ₂ O ^{**}
	Gg/10 ⁶ m ³ gas production		
Flaring – gas production	1.2·10 ⁻³	7.6·10 ⁻⁷	2.1·10 ⁻⁸
Flaring sweet gas processing	1.8·10 ⁻³	1.2·10 ⁻⁶	2.5·10 ⁻⁸
Flaring sour gas processing	3.6·10 ⁻³	2.4·10 ⁻⁶	5.4·10 ⁻⁸
[*] Uncertainty (%) ±25			
^{**} Uncertainty from –10 to 1000			

Destruction efficiency is generally higher than overall combustion efficiency. For example, a combustion efficiency of 96.5% is estimated to correspond to a destruction efficiency of 98% (U.S. EPA 2015). Although methane destruction efficiency can exceed overall flare combustion efficiency, in the absence of specific data, it is a conservative approach to assume the same default value for both combustion and destruction efficiency.

5. Summary

The EU’s methane regulations mark a significant step toward reducing methane emissions in the O&G sector. By enforcing strict monitoring and mitigation requirements and extending compliance to fossil fuel imports, the EU is positioning itself as a global leader in methane management. As climate science and technology evolve, continued collaboration between governments, industries, and research institutions will be essential in achieving methane reduction targets.

The new methane regulations mark a significant shift in methane management, presenting both challenges and opportunities for the oil and gas sector. One of the main challenges is compliance costs, as companies will need to invest in advanced leak detection, monitoring, and emissions control technologies to meet the new standards.

At the same time, the regulations are expected to drive innovation in emissions mitigation, increasing the demand for methane capture and utilization technologies. This could lead to the development of more efficient and cost-effective solutions for reducing methane emissions.

Beyond the EU, the regulations also have global implications. By enforcing methane standards on fossil fuel imports, the EU is influencing international methane reduction efforts and encouraging suppliers worldwide to adopt stricter emissions controls.

Acknowledgments

The paper was performed within the frame of AGH University of Krakow statutory research grant No. 16.16.190.779/B509 Faculty of Drilling, Oil and Gas, Department of Petroleum Engineering, 2025.

References

- American Petroleum Industry (API), 2009, *Compendium of greenhouse gas emissions estimation methodologies for the oil and natural gas industry. August 2009*. https://www.api.org/~media/files/ehs/climate-change/2009_ghg_compendium.pdf [dostęp: 30.11.2025].
- American Petroleum Industry (API), 2021, *Compendium of greenhouse gas emissions methodologies for the natural gas and oil industry, November 2021*. <https://www.api.org/~media/files/policy/esg/ghg/2021-api-ghg-compendium-110921.pdf> [dostęp: 30.11.2025].
- European Commission, 2023, *EU Methane Strategy and Regulatory Framework*. https://energy.ec.europa.eu/topics/carbon-management-and-fossil-fuels/methane-emissions_en [dostęp: 30.11.2025].
- European Commission, 2024, *EU Methane Strategy and Regulation*. <https://eur-lex.europa.eu/eli/reg/2024/1787/oj/eng> [dostęp: 30.11.2025].
- European Parliament, 2024, *Legislation to Reduce Methane Emissions in the Energy Sector*. <https://eur-lex.europa.eu/eli/reg/2024/1787/oj/eng> [dostęp: 30.11.2025].
- Ernst & Young (EY), 2024, *European Parliament Adopts New Regulation to Reduce Global Methane Emissions*. <https://taxnews.ey.com/news/2024-0830-european-parliament-adopts-new-regulation-to-reduce-global-methane-emissions-from-energy-sector> [dostęp: 30.11.2025].
- International Energy Agency (IEA), 2021, *World Energy Outlook 2021*. <https://www.iea.org/reports/world-energy-outlook-2021> [dostęp: 30.11.2025].
- International Energy Agency (IEA), 2022, *Global Methane Tracker 2022*. <https://www.iea.org/reports/global-methane-tracker-2022> [dostęp: 30.11.2025].
- International Energy Agency (IEA), 2023, *The Role of Methane Mitigation in Energy Transition*. <https://www.iea.org/reports/global-methane-tracker-2023/overview> [dostęp: 30.11.2025].
- Intergovernmental Panel on Climate Change (IPCC), 2021, *Sixth Assessment Report (AR6): Methane and Climate Change*. <https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/> [dostęp: 30.11.2025].
- Jackson R.B., Vengosh A., Carey J.W., Davies R.J., Darrah T.H., O'Shea S.J., Poreda R.J., (2013), *The environmental costs and benefits of fracking. Annual Review of Environment and Resources*, 38(1), pp. 321–352. <https://doi.org/10.1146/annurev-environ-031113-144051>.

- Mondaq, 2024, *EU Adopts Legislation to Tackle Methane Emissions in the Oil & Gas Sector*. <https://www.mondaq.com/unitedstates/oil-gas-electricity/1547922/questions-unsatisfactory-answers-the-european-commission-publishes-guidance-on-the-importer-requirements-of-the-eu-methane-regulation> [dostęp: 30.11.2025].
- Saunio M., Stavert A.R., Poulter B., Bousquet P., Canadell J.G., Jackson R.B., Raymond P.A. et al., 2020, *The Global Methane Budget 2000–2017*, *Earth System Science Data*, 12(3), s. 1561–1623. <https://doi.org/10.5194/essd-12-1561-2020>.
- Schlumberger, 2024, *EU Methane Regulation Comes into Force*. <https://www.slb.com/resource-library/article/2024/eu-methane-regulation-comes-into-force> [dostęp: 30.11.2025].
- United Nations Environment Programme (UNEP), 2021, *Oil and Gas Methane Partnership (OGMP) 2.0: A New Standard for Methane Emission Reporting*. <https://www.ogmpartnership.org/resources> [dostęp: 30.11.2025].
- United Nations Environment Programme (UNEP), (2022), *Reducing methane emissions from the fossil fuel industry: Strategies and Challenges*. <https://www.unep.org/annualreport/2022> [dostęp: 30.11.2025].
- U.S. Environmental Protection Agency (U.S. EPA), 2015, *Background information for final emissions factors development for flares and certain refinery operations and final determination for no changes to VOC emissions factors for tanks and wastewater treatment systems*. https://www.epa.gov/sites/default/files/2020-11/documents/response_to_comments.pdf [dostęp: 30.11.2025].
- World Bank, 2023, *Zero Routine Flaring by 2030 Initiative*. <https://www.worldbank.org/en/programs/zero-routine-flaring-by-2030> [dostęp: 30.11.2025].

Badania trwałości zmęczeniowej typowych lin stalowych stosowanych w górniczych wyciągach szybowych

Grzegorz Olszyna , Andrzej Tytko 

Akademia Górniczo-Hutnicza, Wydział Inżynierii Mechanicznej i Robotyki, Kraków

Streszczenie: Kryzys energetyczny w Europie wynikający z wprowadzenia surowych reguł dotyczących importu węglowodorów oraz stosowania węgla kamiennego jako głównego źródła energii spalanej w atmosferze tlenowej ciągle się pogłębia. Zjawisko to jest szczególnie widoczne w Polsce, gdzie ponad 40% energii elektrycznej pochodzi z elektrowni węglowych. W pełni uzasadnione jest więc podejmowanie działań w celu utrzymania zdolności wydobywczych istniejących szybów przy jednoczesnym zapewnieniu najwyższego poziomu bezpieczeństwa eksploatacji. W artykule podjęto problematykę trwałości zmęczeniowej konstrukcji lin nośnych stosowanych w górniczych wyciągach szybowych. Przedstawiono w nim aktualne wyniki badań obecnie produkowanych i eksploatowanych typów lin uzyskane podczas prób na maszynie zmęczeniowej. Ich interpretacja pozwala na wyjaśnienie kilku istotnych zagadnień, spośród których najważniejsze dotyczy mechanizmu zużywania się lin. W badaniach zastosowano statystyczne metody opisu zmęczeniowego zużycia lin nośnych w funkcji liczby cykli pracy. Znajomość tego procesu ma kluczowe znaczenie, ponieważ jego charakter jest nieliniowy, a bezpieczeństwo eksploatacji zależy od podjęcia właściwej decyzji we właściwym momencie użytkowania.

Słowa kluczowe: liny stalowe, trwałość, zmęczenie, procesy losowe, niezawodność

FATIGUE LIFE TESTS OF TYPICAL STEEL WIRE ROPES USED IN MINING SHAFT HOISTS

Abstract: The energy crisis in Europe, resulting from the introduction of strict rules on the import of hydrocarbons and the use of hard coal as the main source of energy burned in an oxygen atmosphere is still deepening. This phenomenon is particularly evident in Poland, where over 40% of electricity is generated by coal power stations. Therefore, considerations regarding maintaining the production capacity of existing shafts while ensuring the highest level of operational safety remain fully justified.

This article addresses the fatigue durability of rope structures used as hoisting ropes in mining shaft hoists. It presents current test results for the rope types currently produced and used, obtained during fatigue testing. The obtained results and their interpretation facilitate the clarification of several essential issues, the most significant of which concerns the mechanism of rope wear. The research used statistical methods to describe fatigue wear of hoisting ropes as a function of the number of operating cycles. Knowledge of this process is crucial because its nature is nonlinear, and the safe operation depends on making the right decision at the right moment of use.

Keywords: steel wire ropes, endurance, fatigue life, random processes, reliability

https://doi.org/10.7494/978-83-68219-80-7_6

1. Wprowadzenie

Od kilku lat obserwuje się dalsze nasilenie kryzysu energetycznego, którego skutki są szczególnie mocno odczuwalne w Polsce, gdzie większość energii elektrycznej (ponad 40%) wytwarzana jest z węgla kamiennego, stanowiącego również podstawowe źródło ciepła w gospodarstwach domowych na obszarach wiejskich. W tej sytuacji w perspektywie najbliższych kilku, a nawet kilkunastu lat będzie konieczne zapewnienie odpowiedniego poziomu wydobycia tego surowca w naszym kraju. Należy więc kontynuować prace nad utrzymaniem zdolności wydobywczych istniejących szybów, tak aby jednocześnie zapewnić niezawodność urządzeń i wysoki poziom bezpieczeństwa ich eksploatacji.

Aby osiągnąć ten cel, prowadzone są liczne badania dotyczące bezpiecznej eksploatacji różnego rodzaju urządzeń linowych. Przyczyną katastrof górniczych jest często zerwanie lin nośnych, których zły stan znacząco pogarsza poziom bezpieczeństwa w kopalni (Wang i in. 2014). W maszynach podstawowych górnictwa odkrywkowego liny pracują w układach wielokrażkowych. Znając parametry liny oraz układ krażków w takich systemach, można określić jej obciążenie (Peterka i in. 2020). W przypadku maszyn stosowanych w kopalniach odkrywkowych układy kół linowych stanowią zespoły mechaniczne kluczowe dla utrzymania jakości i bezpieczeństwa pracy urządzeń (Olszyna i Tytko 2022). Tak duże urządzenia wymagają przeprowadzania ekspertyz technicznych w celu określenia możliwości wydłużenia ich żywotności (Radu i in. 2024).

Problematyka trwałości zmęczeniowej stalowych lin nośnych stosowanych w górniczych wyciągach szybowych pozostaje wciąż aktualna i nie w pełni rozpoznana. W użyciu są różne konstrukcje, w tym liny trójkątnosplotkowe (TSR – *triangular strand rope*) oraz o liniowym styku drutów w splotce (WS – *Warrington-Seale*). W środowisku specjalistów zajmujących się eksploatacją wyciągów szybowych w Polsce od wielu lat trwa dyskusja dotycząca praktyki stosowania tego rodzaju lin, jednak oczywiste jest, że muszą one spełniać podstawowe wymagania określone w przepisach. Zostały one szczegółowo omówione w załączniku nr 4 do Rozporządzenia Ministra Energii z dnia 23 listopada 2016 r. w sprawie szczegółowych wymagań dotyczących prowadzenia ruchu podziemnych zakładów górniczych (Dz.U. 2017, poz. 1118).

Liny TSR i WS są dobrze poznane pod względem konstrukcyjnym, jednak ponieważ składają się z wielu drutów o różnych średnicach, ich ocena metodami magnetycznymi i wizualnymi wciąż jest przedmiotem dyskusji. W artykule wykorzystano unikatowe wyniki badań zmęczeniowych wymienionych rodzajów lin. Ich analiza i interpretacja, zgodnie z zamierzeniem autorów, mają na celu wyjaśnienie kilku istotnych problemów budzących zainteresowanie użytkowników. Najważniejszym z nich jest określenie optymalnego momentu wymiany lin – niezbyt wcześnie, aby uniknąć niepotrzebnych kosztów i w pełni wykorzystać potencjał eksploatacyjny liny, ale też niezbyt późno, by nie dopuścić do powstania zagrożenia. Na rysunku 1 przedstawiono fotografie przekroju poprzecznego i powierzchni jednej z lin, które poddano badaniom zmęczeniowym – liny o konstrukcji splotek 6×36WS.



Rys. 1. Przekrój poprzeczny i powierzchnia liny o konstrukcji spletek 6×36WS

2. Metodyka badań

Badaniom poddano liny trójkątnospłotkowe (TSR) o punktowym styku drutów w spletkach oraz liny typu *Warrington-Seale* (WS) o styku liniowym. Wytwarzają je wszyscy wiodący producenci lin stalowych na świecie. Obie konstrukcje charakteryzują się wysoką jakością wykonania oraz trwałością eksploatacyjną, jednak obszary ich zastosowań są całkowicie odmienne. Liny TSR zdominowały wielolinowe górnicze wyciągi szybowe z napędem ciernym Koepgo, natomiast liny WS szczególnie dobrze sprawdzają się jako liny podnoszące w różnego rodzaju urządzeniach dźwigowych i dźwignicowych. W górnictwie znalazły również zastosowanie w wyciągarkach bębnowych z wielowarstwowym nawijaniem, wykorzystywanych w głębokich i bardzo głębokich szybach (Tytka 2021).

Układ z kołem ciernym w takich maszynach składa się z odcinków prostych i giętych. W zewnętrznych warstwach drutów wszystkich lin powstają naprężenia zginające, a w linach o budowie trójkątnospłotkowej w drutach wewnętrznych także tzw. naprężenia zginające II rodzaju. Dodatkowo należy wziąć pod uwagę naprężenia stykowe. Naprężenia w drutach liny są różne w poszczególnych fazach pracy urządzenia (okresy jazdy ze stałą prędkością, przyspieszania oraz hamowania) (Wang i in. 2012), a naprężenia zginające zależą od średnicy kół liniowych oraz rozstawu kątów środkowych (Zhang i in. 2019). Podczas przeprowadzania testów na tempo zużycia zmęczeniowego lin mają także wpływ parametry związane z budową danej maszyny zmęczeniowej (Wang i in. 2013), więc wyniki badań mogą być porównywane jedynie wtedy, gdy uzyskano je w takich samych warunkach.

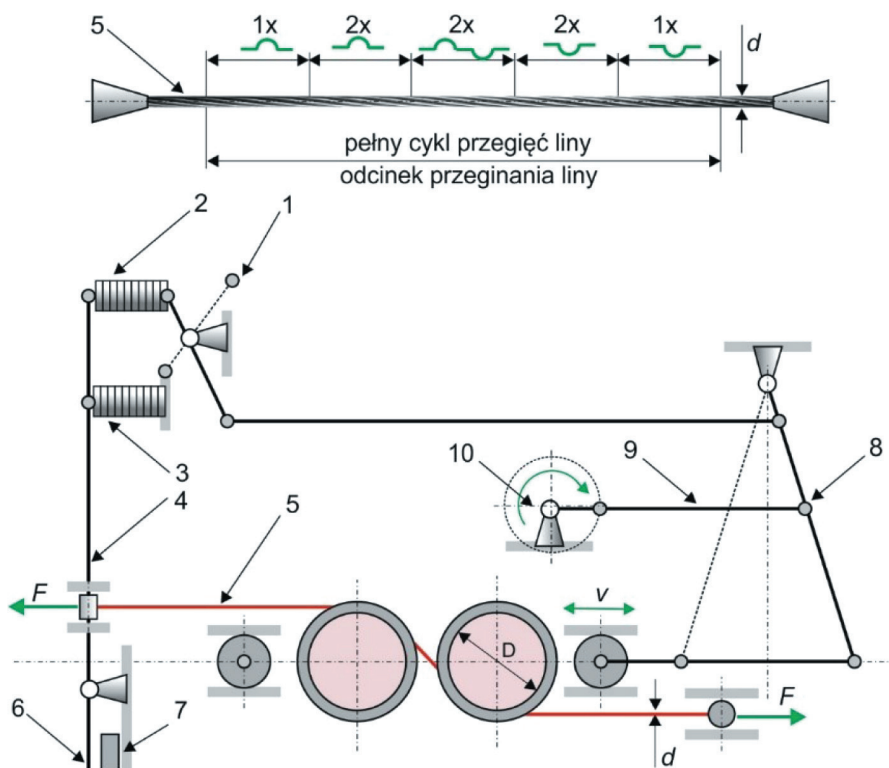
Pomiędzy stykającymi się drutami dochodzi do poślizgu, co prowadzi do lokalnego różnicowania się naprężeń, a w konsekwencji do zmęczenia materiału i zużycia ściernego sąsiadujących drutów (Argatov i in. 2011). Na powierzchni kontaktu wykładziny cierniej z liną występują również uszkodzenia wynikające z odkształceń sprężystych, które mogą się także pojawić wskutek poddawania liny momentowi od skręcania w czasie pracy układu (Peng i in. 2007).

Badania zmęczeniowe lin będących przedmiotem niniejszej analizy przeprowadzono w Laboratorium Lin i Urządzeń Szybowych Głównego Instytutu Górnictwa w Katowicach.

Badania lin, które wszystkie miały średnicę większą niż 40 mm, wykonano na maszynie zmęczeniowej typu P4-GIG-2. W urządzeniu tym liny poddawane są dwustronnemu (dwukierunkowemu) zginaniu przy działaniu siły rozciągającej o charakterze statycznym lub pulsującym. Wyposażone jest ono w mechanizm umożliwiający obrót próbki wokół własnej osi.

Na rysunku 2 przedstawiono schemat kinematyczny maszyny zmęczeniowej P4-GIG-2, natomiast jej podstawowe parametry techniczne zestawiono w tabeli 1. Służy ona do testowania lin o średnicach od 40 mm do 65 mm. Długość badanej próbki liny między uchwytami wynosi 6,5 m, natomiast długość odcinka podlegającego przeginananiu mieści się w zakresie od 4,6 m do 5,0 m w zależności od średnicy kół zginających.

Aby zapobiec przegrzewaniu się liny i wytapianiu smaru podczas pracy urządzenia, wyposażono je w system chłodzenia. Badania zmęczeniowe lin stalowych w warunkach laboratoryjnych mają kluczowe znaczenie dla określenia ich trwałości zmęczeniowej i właściwego doboru rozwiązań do różnego typu urządzeń (Kamarudin i in. 2019).



Rys. 2. Schemat kinematyczny maszyny zmęczeniowej P4-GIG-2

- 1 – wahacz górny; 2 – sprężyna pulsacyjna; 3 – sprężyna napinająca; 4 – wahacz tylny;
5 – badana lina; 6 – wahacz środkowy; 7 – belka oporowa; 8 – wahacz przedni;
9 – łącznik; 10 – napęd

Tabela 1
Podstawowe parametry techniczne maszyny zmęczeniowej P4-GIG-2

Parametr	Jednostka	Wartość
Długość próbki między uchwytami (l_0)	[m]	6,5
Długość przeginanego odcinka	[m]	4,6–5,0
Skok wózka maszyny	[m]	2
Stosunek średnicy koła do średnicy liny (D/d)	–	30
Kąt opasania liny na kole	[°]	62–65
Zakres średnic badanych lin	[mm]	40–65
Prędkość jednostkowa cykli zmęczeniowych	[cykli/min]	10

Zamontowaną w uchwytach linę obciążano siłą statyczną oraz siłą pulsującą używaną za pomocą pary sprężyn – napinającej i pulsacyjnej. Wartość obciążenia była uzależniona od konstrukcji i średnicy badanej liny oraz dobierana na podstawie tablic parametrów obciążeniowych dostępnych w dokumentacji maszyny. Obciążenia pulsacyjne nie mogły przekraczać 25% wartości obciążenia statycznego.

W trakcie badań liny poddawano naciągowi składającemu się ze składowej stałej i zmiennej. Środkowa część liny w jednym cyklu pracy podlegała dwukrotnemu przegięciu na kołach linowych (rys. 2). W celu ograniczenia nagrzewania się próbki spowodowanego tarciem wewnętrznym stosowano chłodzenie liny, tak aby jej temperatura nie przekraczała 50°C.

Badanie każdej próbki kończono w momencie zniszczenia liny, co uznawano za osiągnięcie granicznego stanu zmęczenia. Za całkowite zniszczenie liny uznaje się pełne zerwanie co najmniej jednej splotki lub brak możliwości dalszego obciążania liny siłą poosiową.

Podobne badania lin były prowadzone też m.in. przez Kubisia (2019). Określenie trwałości zmęczeniowej jest kluczowe dla bezpieczeństwa, ponieważ postęp procesu zmęczenia nie jest wykrywalny w fazie eksploatacji – lina nie wykazuje żadnych widocznych uszkodzeń aż do momentu pęknięcia drutu (Lombardi i in. 2013). Trwałość zmęczeniowa lin stalowych jest prognozowana głównie na podstawie wyników badań niszczących oraz niszczących (Zhao i in. 2021), które są prowadzone w celu określenia przewidywanego czasu pracy lin w warunkach zbliżonych do rzeczywistych.

3. Wyniki badań i dyskusja

Badaniom zmęczeniowym poddano liny zestawione w tabeli 2. Wszystkie uzyskane wyniki zostały opracowane statystycznie, natomiast wyniki badań lin oznaczonych jako TSR52A, TSR56A oraz WS56AOL przedstawiono i szczegółowo omówiono w dalszej części pracy.

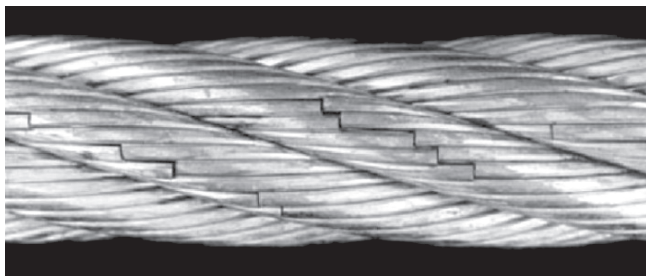
Tabela 2
Rodzaje lin poddanych badaniom zmęzeniowym i ich parametry

Oznaczenie liny	Rodzaj i średnica liny [mm]	Długość próbki między uchwytami l_0 [mm]	Przekrój, S_{Fe} [mm ²]	Siła zrywająca P_w [kN]	Wytrzymałość R_m [MPa]	Siła w linie F_r [kN]	n	Pulsacja [Hz]		
TSR52A	TSR52	6500	1188,0	1905,41	1570	317,57	6,0	0		
TSR52B				1745,34		290,89		43,60		
TSR52C				1891,94		315,32		b.d.		
TSR54A	TSR54		1278,0	2098,82		299,83	7,0	0		
TSR54B				1967,68		281,10		59,97		
TSR54C				2038,76		356,00		56,22		
TSR54D				1967,68		281,10		b.d.		
TSR54E	TSR56		1424,0	2394,46		342,06	7,0	0		
TSR56A				2296,48		328,00		b.d.		
TSR56B				1349,5		2514,53		330,86	7,6	66,17
TSR56C						2257,16		312,84	7,5	62,57
WS56AOL	WS56	6490	1331,0	2002,62	1570	333,77	6,0	66,75		
WS54ALL	WS54	6500	1228,0	b.d.	b.d.	b.d.	b.d.	b.d.		
WS54BOL										
WS50NFC	WS50									
WS60EPIWRC	WS60									
WS60NFC										

gdzie:
TSR – liny trójkątnospłotkowe,
WS – liny *Warrington-Seale*,
LL – lina współzwita,
OL – lina przeciżwita,
NFC – rdzeń włókienny naturalny,
EPIWRC – rdzeń w postaci niezależnej liny pokrytej tworzywem,
 n – współczynnik bezpieczeństwa,
b.d. – brak danych,
A, B, C, D, E – kolejne próbki lin o tej samej konstrukcji

Podczas badań okresowo dokonywano pomiarów średnicy oraz długości skoku liny, zawsze w tym samym miejscu. Ponadto rejestrowano lokalizację i liczbę widocznych pęknięć zmęczeniowych oraz przyrost długości liny. Na podstawie przeprowadzonych badań wyznaczono trwałość zmęczeniową jako liczbę cykli pracy do momentu zniszczenia liny lub zakończenia eksperymentu.

Na rysunku 3 przedstawiono przykładowy fragment liny. Na odcinku, który znajdował się pomiędzy kołami linowymi maszyny zmęczeniowej P4-GIG-2, widoczne są pęknięte druty. Należy podkreślić, że identyfikacja pęknięć na podstawie badań wizualnych (VT – *visual testing*) (zob. Tytko i in. 2023) oraz ich zliczenie nie jest łatwe i ma istotny wpływ na dokładność wykonywanych pomiarów.



Rys. 3. Pęknięcia zmęczeniowe drutów
widoczne na odcinku liny pomiędzy kołami linowymi przewojowymi
maszyny zmęczeniowej P4-GIG-2

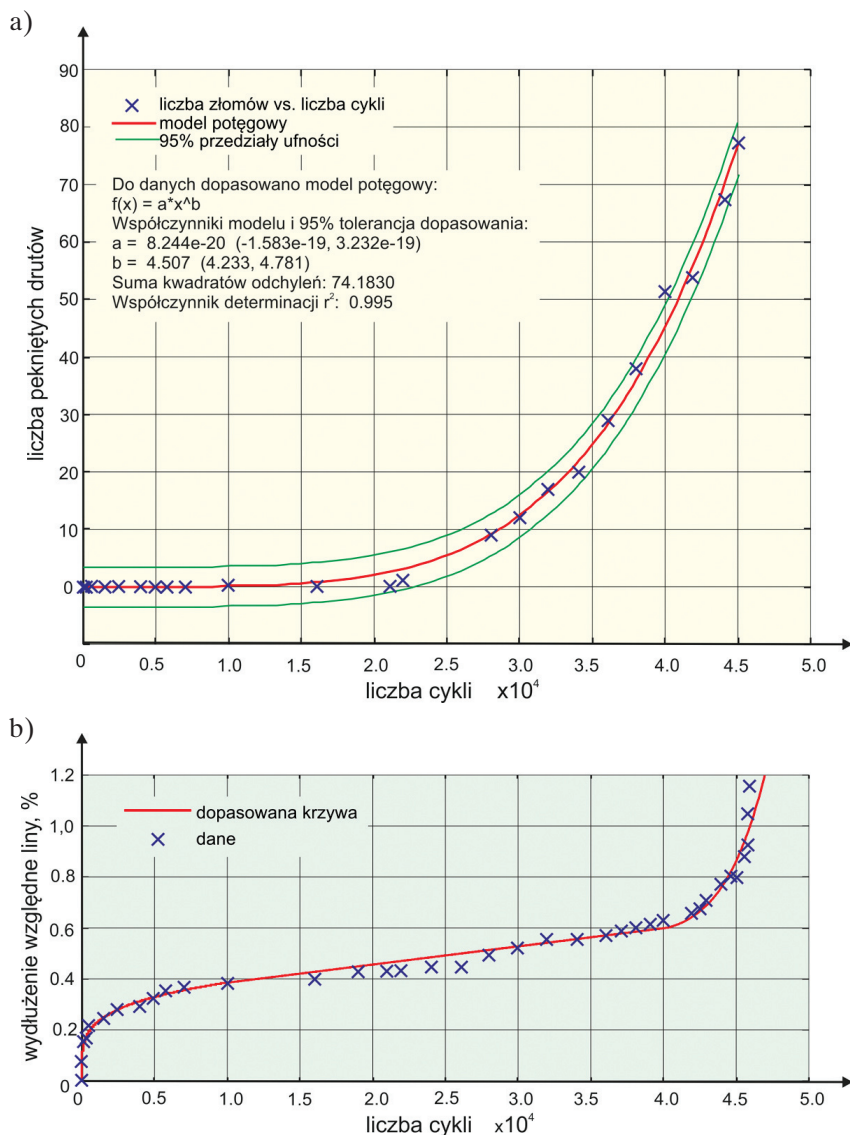
Zmęczenie materiału jest wywołane zmiennym stanem naprężeń od zginania, skręcania i nacisków, które powstają w wyniku przesuwania się liny po kołach linowych, oraz od pulsujących naprężeń rozciągających działających na linę w trakcie pracy (Breim 2007). Badania przeprowadzone na pojedynczych drutach wykazały, że przed wystąpieniem uszkodzenia ulegają one znacznemu odkształceniu plastycznemu (Teissier i in. 2017).

Wybrane wyniki badań trwałości zmęczeniowej dwóch lin typu TSR i jednej typu WS zostały opracowane w formie wykresów obrazujących narastanie pęknięć drutów oraz względne wydłużenie liny w funkcji liczby cykli zmęczeniowych (rys. 4–6). Na rysunku 4a przedstawiono wykres przyrostu liczby pęknięć widocznych na powierzchni liny trójkątnosplotkowej o średnicy 52 mm (TSR52A) w zależności od wykonanej liczby cykli według równania (1), a na rysunku 4b – wydłużenie względne (ϵ) tej liny w zależności od liczby cykli. Wydłużenie względne (ϵ) jest obliczane jako przyrost długości liny w odniesieniu do jej długości początkowej.

$$n = a \cdot N^b \quad (1)$$

gdzie:

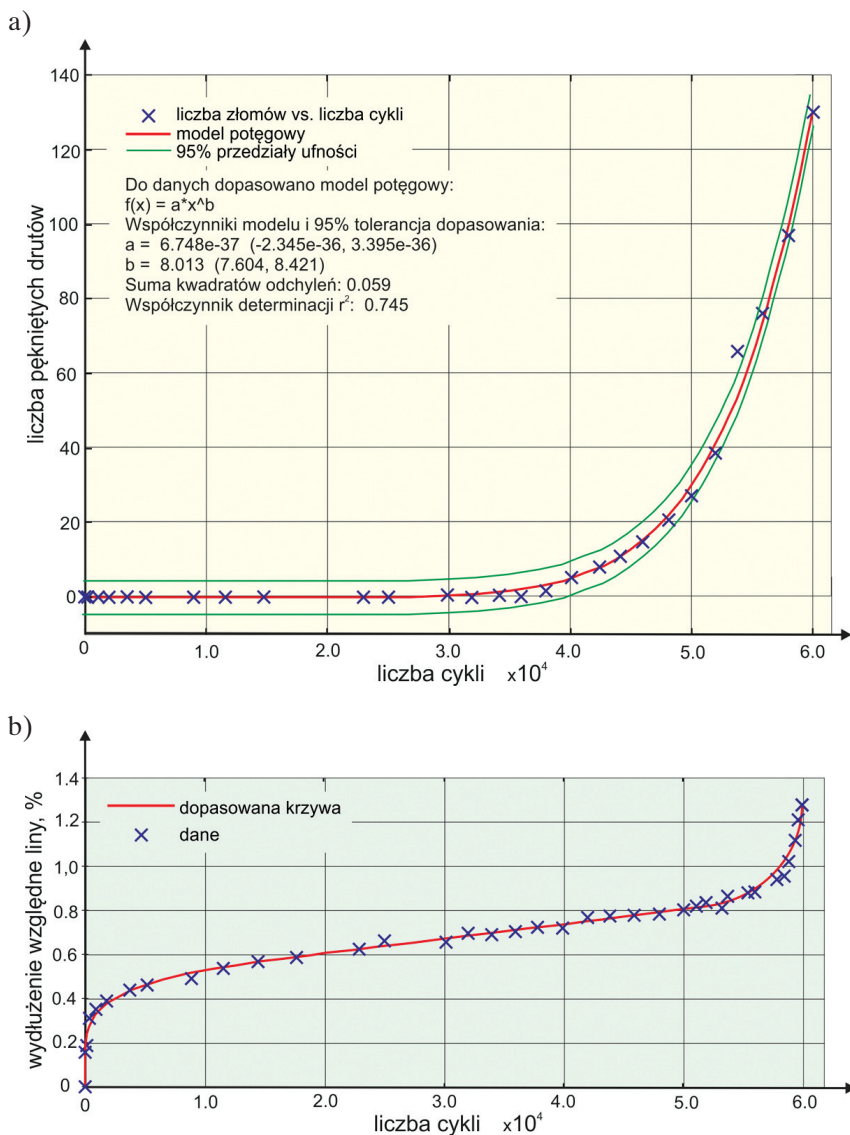
- n – liczba pęknięć zmęczeniowych,
- a, b – współczynniki równania,
- N – liczba cykli zmęczeniowych.



Rys. 4. Wyniki badań na maszynie zmęczeniowej P4-GIG-2 dla liny trójkątnosplotkowej o średnicy 52 mm (TSR52A):

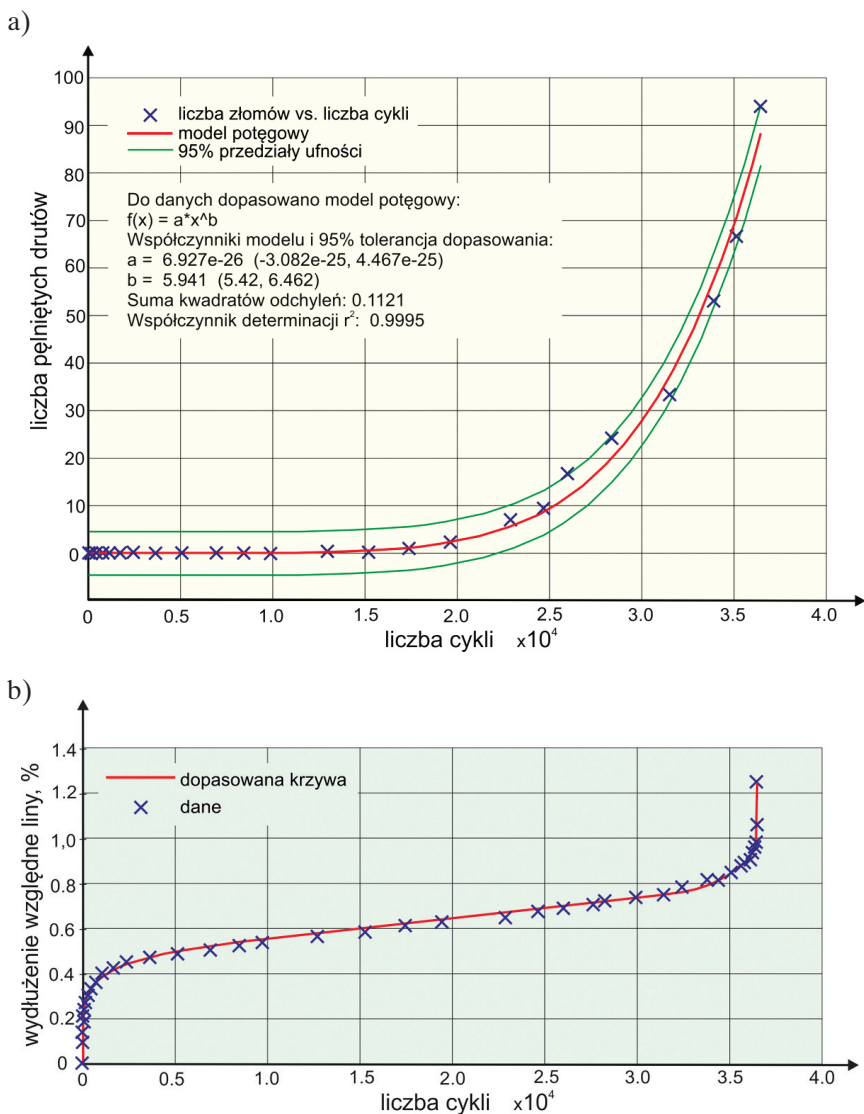
- a) krzywa zależności liczby pękniętych drutów od liczby cykli;
 b) krzywa zależności wydłużenia względnego (ϵ) od liczby cykli

Na rysunku 5a przedstawiono wykres przyrostu liczby pęknięć widocznych na powierzchni liny trójkątnosplotkowej o średnicy 56 mm (TSR56A) w zależności od wykonanej liczby cykli według równania (1), a na rysunku 5b – wydłużenie względne (ϵ) tej liny w zależności od liczby cykli.



Rys. 5. Wyniki badań na maszynie zmęczeniowej P4-GIG-2 dla liny trójkątnopłatkowej o średnicy 56 mm (TSR56A):
 a) krzywa zależności liczby pękniętych drutów od liczby cykli;
 b) krzywa zależności wydłużenia względnego (ϵ) od liczby cykli

Na rysunku 6a przedstawiono wykres przyrostu liczby pęknięć widocznych na powierzchni liny przeciwwziewnej typu WS o średnicy 56 mm (WS56AOL) w zależności od wykonanej liczby cykli według równania (1), a rysunku 6b – wydłużenie względne (ϵ) tej liny w zależności od liczby cykli.



Rys. 6. Wyniki badań na maszynie zmęczeniowej P4-GIG-2 dla liny przeciwwziewnej typu WS o średnicy 56 mm (WS56AOL):

- a) krzywa zależności liczby pękniętych drutów od liczby cykli;
 b) krzywa zależności wydłużenia względnego (ϵ) od liczby cykli

Uzyskane wyniki potwierdzają, że charakterystyki zmęczeniowe badanych lin najlepiej opisuje równanie (1). Na wykresie zmęczenia liny przedstawiono estymowany przebieg dopasowanej funkcji oraz 95-procentowe przedziały istotności tego dopasowania. W każdym analizowanym przypadku przedziały te okazały się bardzo wąskie.

Na wykresach umieszczono również parametry estymowanego modelu oraz dwie statystyki oceny dobroci dopasowania.

Nie zaprezentowano wyników dopasowania modeli do rzeczywistych przebiegów względnego wydłużenia lin, ponieważ dobór modelu w tym przypadku miał charakter arbitralny, co wynikało z przyjęcia wielomianu wysokiego stopnia.

Typowe charakterystyki wydłużenia lin dzielą się na trzy wyraźne okresy (fazy):

- faza I: krótki okres początkowy o szybkim wzroście wydłużenia;
- faza II: długi, niewielki, liniowy wzrost wydłużenia;
- faza III: okres przyspieszonego wydłużenia.

W fazie I dochodzi do wyrównania naciągów poszczególnych drutów i splotek. Pokonują one występujące wewnątrz liny siły tarcia, dzięki czemu zostają skompensowane różne długości drutów, a druty i splotki układają się w linie w taki sposób, że następuje wyrównanie w nich stanu naprężenia. W przeprowadzonych badaniach faza ta trwała do momentu osiągnięcia wydłużenia względnego lin w zakresie od 0,4% do 0,5%.

W fazie II (najdłuższej) obserwuje się niewielki, liniowy wzrost wydłużenia lin o kolejne około 0,4%. Pod koniec tego okresu na linie pojawiają się pierwsze pęknięcia zmęczeniowe.

W fazie III (najkrótszej) następuje nieliniowy wzrost wydłużenia względnego nawet o około 0,5% do momentu odłożenia lub zniszczenia próbki. W całej fazie powstają kolejne pęknięcia zmęczeniowe drutów, przy czym tempo ich pojawiania się narasta potęgowo, a lina traci zwartość. Okres ten jest stosunkowo krótki, dlatego dla bezpieczeństwa liny (zwłaszcza pracujące w górniczych wyciągach szybowych) powinny być odkładane wcześniej, bez względu na zaobserwowany poziom ich zużycia.

4. Wnioski

Prezentowane wyniki badań odnoszą się wyłącznie do lin trójkątnosplotkowych oraz lin o konstrukcji *Warrington-Seale*. Z przedstawionych przebiegów wynika, że liny trójkątnosplotkowe charakteryzują się większą trwałością zmęczeniową.

Proces narastania liczby zidentyfikowanych wizualnie pęknięć zmęczeniowych drutów w linach o typowym wykonaniu splotek (styk punktowy lub liniowy) przedstawiony w funkcji liczby przepracowanych cykli ma charakter funkcji potęgowej o dość łagodnej charakterystyce, co pozwala na śledzenie tempa tego procesu. Obserwacja dwóch procesów zużycia (tj. wydłużenia względnego i narastania liczby powstających pęknięć zmęczeniowych) pozwala trafnie określić właściwy moment wymiany liny. Z analizy przedstawionych wyników można wnioskować, że oba zjawiska są ze sobą ściśle powiązane.

Procedury stosowane obecnie w ramach wspomnianej oceny stanu technicznego liny są związane z procesami zużycia zmęczeniowego, ale nie uwzględniają wyników badań zmęczeniowych. Istnieją jednak metody ciągłego aparaturowego pomiaru zmian długości skoków pracującej liny, np. pomiar metodą laserowego skanowania powierzchni liny.

Literatura

- Argatov I.I., Gómez X., Tato W., Urchegui M.A., 2011, *Wear evolution in a stranded rope under cyclic bending: Implications to fatigue life estimation*, Wear, vol. 271, iss. 11–12, s. 2857–2867. <https://doi.org/10.1016/j.wear.2011.05.045>.
- Breim U., 2007, *Torque cycle fatigue of wire ropes*, [w:] Ridge I.M.L. (ed.), *OIPEEC Conference Proceedings. How to Get the Most out of Your Ropes: Rope-Machine Interaction, Rope Maintenance and Lubrication*, OIPEEC, Reading, UK, s. 99–106.
- Kamarudin M., Jamaluddin M.J., Tamin M.N., 2019, *Fatigue assessment of steel wire rope*, AIP Conference Proceedings, vol. 2068, iss. 1, art. 020109. <https://doi.org/10.1063/1.5089408>.
- Kubiś B., 2019, *Wpływ wstępnej deformacji spletek w linach kompaktowanych na ich trwałość i parametry wytrzymałościowe*, Główny Instytut Górnictwa, Katowice [praca doktorska].
- Lombardi L., Clerici F., 2013, *An innovative method for wire rope fatigue life evaluation*, [w:] Ridge I.M.L. (ed.), *Proceedings of the OIPEEC Conference 2013. Simulating Rope Applications*, OIPEEC, Oxfordshire, UK, s. 115–128.
- Olszyna G., Tytko A., 2022, *A tool for determining the number of bends and places of accumulation of potential wear of steel ropes operating in the luffing systems of basic opencast mining machines*, Mining Machines, vol. 40, iss. 4, s. 229–237. <https://doi.org/10.32056/KOMAG2022.4.5>.
- Peng Y.-x., Zhu Z.-c., Chen G.-a., Cao G.-h., 2007, *Effect of tension on friction coefficient between lining and wire rope with low speed sliding*, Journal of China University of Mining and Technology, vol. 17, iss. 3, s. 409–413. [https://doi.org/10.1016/S1006-1266\(07\)60115-1](https://doi.org/10.1016/S1006-1266(07)60115-1).
- Peterka P., Krešák J., Vojtko M., Halek B., Heinz D., 2020, *The failure analysis of the drilling rig hoisting steel wire rope*, Eksploatacja i Niezawodność – Maintenance and Reliability, vol. 22, iss. 4, s. 667–675. <https://doi.org/10.17531/ein.2020.4.10>.
- Radu S.M., Vilceanu F., Toderas M., Lihoacă A., Dinescu S., 2024, *Determining the level of structural and mechanical degradation of steel in the supporting structure of mining excavation machinery*, Processes, vol. 12, iss. 1, art. 153. <https://doi.org/10.3390/pr12010153>.

- Rozporządzenie Ministra Energii z dnia 23 listopada 2016 r. w sprawie szczegółowych wymagań dotyczących prowadzenia ruchu podziemnych zakładów górniczych (Dz.U. 2017, poz. 1118).
- Teissier J.M., Ridge I.M.L., Evans J.J., Fournier M., 2017, *The effect of wire break distribution on the breaking strength of a wire rope*, [w:] Dohm M.A.R. (ed.), *Proceedings of the OIPEEC Conference 2017. Rope – Present and Future*, OIPEEC, Didcot Oxfordshire, UK, s. 267–293.
- Tytko A., 2021, *Liny stalowe. Budowa, właściwości, eksploatacja, zastosowania*, Wydawnictwo Naukowe PWN, Warszawa.
- Tytko A., Olszyna G., Kocór G., Szot M., 2023, *Some stochastic aspects of safety work of steel wire ropes used in mining-shaft hoists*, *Sustainability*, vol. 15, iss. 9, art. 7590. <https://doi.org/10.3390/su15097590>.
- Wang D., Zhang D., Zhang Z., Ge S., 2012, *Effect of various kinematic parameters of mine hoist on fretting parameters of hoisting rope and a new fretting fatigue test apparatus of steel wires*, *Engineering Failure Analysis*, vol. 22, s. 92–112. <https://doi.org/10.1016/j.engfailanal.2012.01.008>.
- Wang D., Zhang D., Wang S., Ge S., 2013, *Finite element analysis of hoisting rope and fretting wear evolution and fatigue life estimation of steel wires*, *Engineering Failure Analysis*, vol. 27, s. 173–193. <https://doi.org/10.1016/j.engfailanal.2012.08.014>.
- Wang, D., Zhang D., Ge S., 2014, *Effect of terminal mass on fretting and fatigue parameters of a hoisting rope during a lifting cycle in coal mine*, *Engineering Failure Analysis*, vol. 36, s. 407–422. <https://doi.org/10.1016/j.engfailanal.2013.11.006>.
- Zhang J., Wang D., Song D., Zhang D., Zhang Ch., Wang D., Araújo J.A., 2019, *Tribofatigue behaviors of steel wire rope under bending fatigue with the variable tension*, *Wear*, vol. 428–429, s. 154–161. <https://doi.org/10.1016/j.wear.2019.03.004>.
- Zhao D., Liu Y.-X., Ren X.-T., Gao J.-Z., Liu S.-G., Dong L.-Q., Cheng M.-S., 2021, *Fatigue life prediction of wire rope based on grey particle filter method under small sample condition*, *Eksploatacja i Niezawodność – Maintenance and Reliability*, vol. 23, iss. 3, s. 454–467. <https://doi.org/10.17531/ein.2021.3.6>.

Segmentacja aorty z obrazów tomografii komputerowej przy użyciu głębokich sieci neuronowych

Piotr Paniec¹ , Maciej Stanuch² , Andrzej Skalski³ 

¹ Akademia Górniczo-Hutnicza, Wydział Inżynierii Metali i Informatyki Przemysłowej, Kraków

² MedApp S.A., Kraków, Poland

³ Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej, Kraków

Streszczenie: W artykule zaproponowano wykorzystanie architektury nnU-Net do segmentacji aorty z obrazów tomografii komputerowej. Przetwarzanie wstępne obejmowało transformację różnych zakresów wartości do zakresu odpowiadającemu skali jednostek Hounsfielda oraz procedurę okienkowania. Na etapie treningu sieci przeprowadzono pięciokrotną walidację krzyżową oraz wytrenowano pięć modeli, których predykcje na zbiorze testowym zostały połączone. Dla modeli uzyskano średni współczynnik Dice’a na poziomie 0,921 oraz IoU równe 0,858, oznaczające wysoką jakość predykcji. Wartość odległości Hausdorffa ujawniła istnienie elementów odstających, a analiza trójwymiarowych rekonstrukcji wskazała obszary o niższej dokładności segmentacji.

Słowa kluczowe: uczenie głębokie, sieci neuronowe, segmentacja obrazów, obrazowanie medyczne, tomografia komputerowa, informatyka, aorta

AORTA SEGMENTATION FROM CT IMAGES USING DEEP NEURAL NETWORKS

Abstract: The manuscript proposes employing the nnU-Net architecture for the segmentation of the aorta from computed tomography images. Preprocessing involved transforming various value ranges to align with the Hounsfield unit scale, in addition to implementing a windowing procedure. During the network training phase, a five-fold cross-validation protocol was utilized, resulting in the training of five models, whose predictions on the test set were aggregated. The models achieved an average Dice coefficient of 0.921 and an Intersection over Union of 0.858, signifying high prediction quality. The value of the Hausdorff distance indicated the presence of outliers, while the analysis of three-dimensional reconstructions highlighted regions of diminished segmentation accuracy.

Keywords: deep learning, neural networks, image segmentation, medical imaging, computed tomography, computer science, aorta

https://doi.org/10.7494/978-83-68219-80-7_7

1. Wprowadzenie

Tomografia komputerowa (CT, ang. *computed tomography*) jest jedną z najważniejszych technik stosowanych w diagnostyce obrazowej (Hsieh i Flohr 2021). Aorta, największa tętnica organizmu, bywa dotknięta patologiami w budowie i strukturze, które bez diagnozy i odpowiedniego leczenia mogą zagrażać zdrowiu i życiu. W konsekwencji niezwykle istotna jest jej odpowiednia ocena, ustalana na podstawie obrazowania tomografii komputerowej.

Jednym z narzędzi informatycznych, które mogą zostać wykorzystane w tym celu, jest automatyczna segmentacja aorty w obrazach CT, która umożliwia między innymi wyodrębnienie jej światła, trójwymiarową rekonstrukcję geometrii naczynia, a w konsekwencji precyzyjne pomiary geometrii. Celem pracy jest stworzenie narzędzia, bazującego na sieci neuronowej, które umożliwia segmentację w pełni automatycznie, niezależnie od protokołu akwizycji.

Pod względem danych obrazowych wartości powinny być reprezentowane w jednostkach Hounsfielda będących względnym współczynnikiem osłabienia promieniowania rentgenowskiego odniesionego do wody. Woda destylowana reprezentowana jest przez wartość 0 HU (Gücük i Üyetürk 2014, Jhala i Thaker 2015). Ze względu na sposób zapisu wartości w danych, oznaczenia miejsc w obrazach, które były poza obszarem akwizycji, typu tomografu, jednostki czy protokołu akwizycji zakres wartości może być inny.

2. Segmentacja aorty

W pracy zaproponowano na potrzeby segmentacji wykorzystanie narzędzia bazującego na uczeniu głębokim. Jest to podejście często stosowane w analizie obrazów medycznych. Powodem popularności jest automatyczna ekstrakcja cech, oznaczająca, że model samodzielnie, w trakcie procesu uczenia, na podstawie dostarczonych danych, określa istotne cechy, które wpływają na predykcję wyniku.

W niniejszej pracy zastosowano framework nnU-Net (Isensee i in. 2019b, 2021, 2024) w wersji V2. Jest to system stworzony na potrzeby dopasowania rodziny architektur U-Net (Ronneberger i in. 2015) do domeny modalności stosowanych w obrazowaniu medycznym. Framework nnU-Net analizuje statystykę zbioru (rozmiary danych, rozmiary woksela, liczbę klas) i na tej podstawie generuje wariant architektury U-Net oraz odpowiadający mu zestaw parametrów treningu. W konsekwencji dobór topologii sieci czy hiperparametrów jest konfigurowany automatycznie. Jest to narzędzie ułatwiające pracę, przy jednoczesnym zachowaniu najwyższych wyników dla uczenia nadzorowanego (Huang i in. 2022, Isensee i in. 2019b, 2024).

W przypadku nnU-Net możliwe jest uczenie modelu zarówno przekrojami dwuwymiarowymi, jak i danymi 3D, a w obrębie danych 3D występuje kilka domyślnie zdefiniowa-

nych sposobów budowania sieci. Nazywane są one konfiguracjami (ang. *configurations*), z których domyślnie dostępne są cztery: *2d*, *3d_fullres*, *3d_lowres*, *3d_lowres_cascade*. Do treningu wybrano konfigurację *3d_fullres*, która tworzy architekturę opartą na konwolucji trójwymiarowej. Pozwala to sieci analizować ciągłe zależności przestrzenne w obrębie przetwarzanej objętości obrazowania, co zwiększa ilość przestrzennego kontekstu i może podnieść skuteczność predykcji (Isensee i in. 2019b). Zastosowana została pięciokrotna walidacja krzyżowa, a każdy model trenowany był przez 200 epok.

W przypadku współczynnika uczenia (ang. *learning rate*) początkowo używana jest ustalona stała wartość wynosząca 0,01. Domyślnie wartość maleje wraz z postępem treningu, jest zależna od epoki treningu według wzoru (1):

$$lr(e) = lr_0 \left(1 - \frac{e}{E} \right)^{0,9} \quad (1)$$

gdzie:

- lr_0 – wartość początkowa współczynnika uczenia (domyślnie 0,01),
- e – bieżąca epoka,
- E – maksymalna liczba epok (domyślnie 1000),
- $lr(e)$ – wartość współczynnika uczenia dla danej epoki.

Przed rozpoczęciem treningu użyty framework domyślnie projektuje zestaw augmentacji jako element planu eksperymentu, natomiast stosuje je na bieżąco w trakcie uczenia, generując losowo przekształcone wersje obrazów wejściowych. Domyślnie stosowana augmentacja obejmuje losową rotację, losowe skalowanie, szum gaussowski, rozmycie gaussowskie, losową jasność, losowy kontrast, symulowanie niższej rozdzielczości, korekcję gamma, odbicia lustrzane (Isensee i in. 2019a), co ma na celu zwiększenie różnorodności obrazów i poprawę uogólniania. W prezentowanym podejściu zrezygnowano z dodania losowego skalowania, losowej jasności, losowego kontrastu czy korekcji gamma ze względu na fakt, że te transformacje nie odzwierciedlają możliwych zmian w danych. W ostatecznym treningu zastosowano losową rotację, szum gaussowski oraz odbicie lustrzane. Te zabiegi odzwierciedlają możliwość transformacji, które mogą wystąpić w rzeczywistych danych, i jednocześnie poprawiają generalizację modelu.

Do rozwiązania problemu przedstawionego w pracy wykorzystano publicznie dostępny zbiór zanonimizowanych danych tomograficznych. Obszar obrazowania obejmuje odcinek od szyi do miednicy. Zbiór danych zawiera tomografie komputerowe pochodzące z trzech ośrodków (Radl i in. 2022), wraz z referencyjnymi obrysami wykonanymi przez lekarzy. W konsekwencji uzyskano zbiór 56 trójwymiarowych obrazowań tomografii komputerowej i masek, określających położenie aorty i jej odgałęzień.

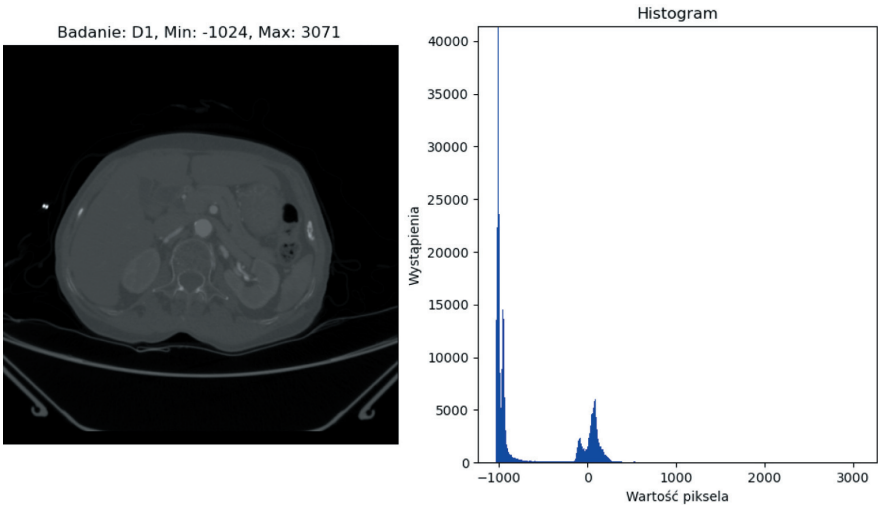
Maski zostały przygotowane w sposób półautomatyczny. Każdy przypadek znajduje się w osobnym folderze. Foldery zawierają plik *.nrrd* z danymi z przeprowadzonego

badania tomografii komputerowej, a także plik *.seg.nrrd* z maską z zaznaczoną aortą, używaną dalej w procesie uczenia. W tabeli 1 przedstawiono podstawowe informacje na temat danych tomograficznych wykorzystanych w pracy.

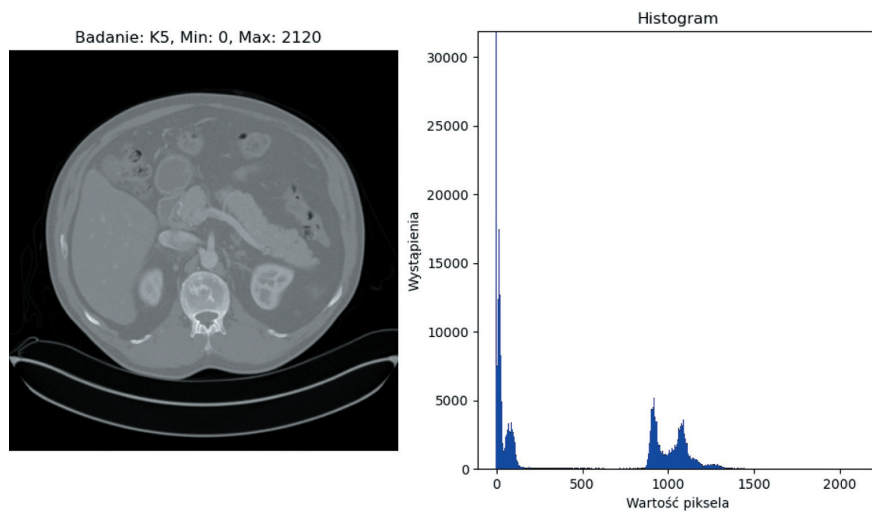
Tabela 1
Zestawienie informacji o zbiorze danych (Radl i in. 2022)

Informacje o obrazach	KiTS	RIDER	Dongyang
Rozdzielczość x/y	512 × 512 pikseli	512 × 512 pikseli	512 × 666 pikseli
Liczba przekrojów względem osi	94/146/1059	260/1008/1140	122/149/251
Grubość przekroju	0,5/5/5 mm	0,625/0,625/2,5 mm	2/3/3 mm
Patologie	brak	rozwarstwienie aorty, tętniaki aorty brzusznej	brak
Liczba przypadków	20	18	18

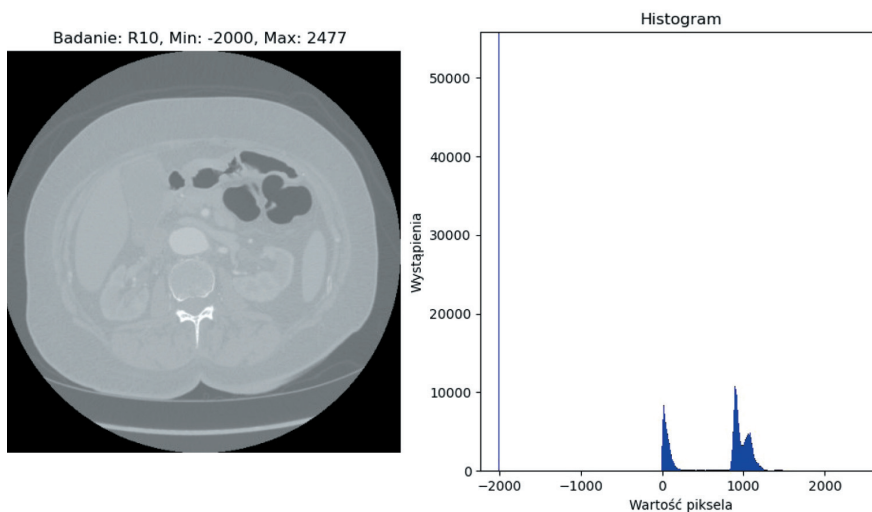
Dane z tomografii komputerowej powinny być zapisane w skali Hounsfielda. Na podstawie histogramów (rys. 1–4) wyznaczonych dla obrazowań zauważono, że przypadki w zbiorze różnią się wartością minimalną, a wartości dla tych samych struktur (np. aorty) różnią się w zależności od pochodzenia danych. Ponadto w przypadku zbioru RIDER obszar zawierający piksele spoza walca obrazowania miał przypisaną wartość -2000 (poza skalą Hounsfielda) jako objętość nieobjętą akwizycją, co przedstawia rysunek 3.



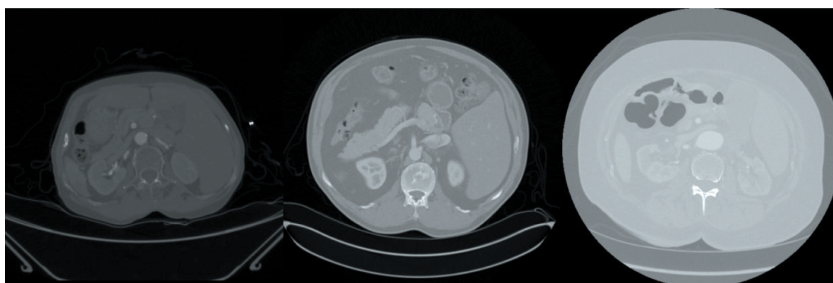
Rys. 1. Przekrój badania D1 oraz jego histogram



Rys. 2. Przekrój badania K5 oraz jego histogram



Rys. 3. Przekrój badania R10 oraz jego histogram



Rys. 4. Przekroje badań D1 (Dongyang), K5 (KiTS) oraz R10 (RIDER) przed przetwarzaniem wstępnym

W konsekwencji dla wartości wokseli obrazowań ze zbioru RIDER, o początkowej wartości -2000 , zaproponowano zastąpienie jej wartością minimalną z obrazowania znalezionej poza tym obszarem.

Rozbieżność sposobu zapisu wartości wokseli wymagała ujednolicenia zakresu względem skali Hounsfielda przed rozpoczęciem uczenia sieci. W związku z tym przyjęto liniową transformację:

$$I_{HU} = \alpha I_{raw} + \beta \quad (2)$$

gdzie:

I_{HU} – wartość woksela w skali Hounsfielda,

I_{raw} – pierwotna, zapisana w pliku wartość woksela,

α – współczynnik skalujący (odpowiada polu *RescaleSlope* w standardzie DICOM),

β – przesunięcie (odpowiada polu *RescaleIntercept* w standardzie DICOM).

W nagłówkach plików *.nrrd* brakowało metadanych *RescaleSlope* i *RescaleIntercept*, dlatego obie wartości oszacowano ręcznie. Współczynniki α oraz β dobrano tak, by charakterystyczne zakresy sygnału odpowiadały typowym poziomom HU dla aorty, płuc i trzonów kręgow. Dzięki temu obrazy ze zbiorów KiTS i RIDER zostały przeskalowane do rzeczywistych jednostek HU (dla zbioru RIDER po zamianie wartości -2000 na wartość minimalną), w przypadku obu zbiorów $\alpha = 1$, $\beta = -1024$, natomiast dane z Dongyang były już poprawnie reprezentowane w skali Hounsfielda.

Otrzymane po tym ujednoliceniu wartości wokseli dla aorty stały się zbliżone, co nie tylko umożliwiło późniejsze okienkowanie, ale także poprawiło jakość procesu uczenia sieci.

W pracy wykorzystano trzy zbiory danych różniące się m.in. rozdzielczością przestrzenną. W konsekwencji konieczne było dostosowanie rozmiarów obrazów do jednej przyjętej wartości. Zdecydowano się ujednolicić rozdzielczość przekrojów do wymiarów 512×512 pikseli. Jest to standardowa wartość stosowana w tomografii komputerowej (Balmer i in. 2022). Zbiory danych RIDER oraz KiTS od początku miały docelową rozdzielczość, zbiór Dongyang należało dopasować.

Obrazy ze zbioru Dongyang mają wymiary 512×666 pikseli. W celu dopasowania rozmiarów danych zdecydowano się wykadrować tomografie z powyższego zbioru tak, aby ustandaryzować rozdzielczość wszystkich obrazowań. Upewniono się, że w usuwanym fragmencie nie znajdują się istotne dane. Jeśli w obszarze, który ma zostać usunięty, występują wartości inne niż 0, wówczas program o tym informuje. W obszarze nie znajdowały się informacje istotne pod względem procesu uczenia, wobec czego można było wykadrować obrazy bez utraty informacji.

Do danych wyrażonych w skali Hounsfielda zastosowano okienkowanie wyrażone wartościami domykającymi przedział. Badanie wykazało istotny wpływ okienkowania

na trening modelu, a także usunięcie artefaktów powstałych w trakcie wykonywania badania. Na podstawie wizualizacji danych, po zastosowaniu różnych wartości okienkowania do ostatecznego rozwiązania użyto przedziału $[-50; 600]$.

Dane używane do trenowania sieci powinny być znormalizowane. W nnU-Net domyślna normalizacja dla tomografii komputerowej polega na progowaniu intensywności do percentyli 0,5 i 99,5 wyznaczonych na podstawie danych ze zbioru treningowego oraz na standaryzacji (z-score) z użyciem globalnej średniej i odchylenia standardowego. W tej pracy wyłączono domyślnie oferowaną normalizację intensywności, a zastosowano normalizację polegającą na transformacji wejściowego przedziału po okienkowaniu do przedziału $[0; 1]$, na podstawie wzoru (3).

$$I'(x) = \frac{I(x) - I_{\min.}}{I_{\max} - I_{\min.}} \quad (3)$$

gdzie:

- $I(x)$ – oryginalna wartość woksela w punkcie x ,
- $I_{\min.}$ – wartość minimalna w całym obrazie po okienkowaniu,
- $I_{\max}$ – wartość maksymalna w całym obrazie po okienkowaniu,
- $I'(x)$ – wartość normalizowana, należąca do przedziału $[0; 1]$.

Normalizacja pozwala na ograniczenie problemu eksplozji gradientu (ang. *gradient explosion*) podczas procesu uczenia, a tym samym na jego stabilizację. Normalizacja danych, zgodnie z oczekiwaniami, przyniosła poprawę wyników treningu (Cabana i Lago-Fernández 2021).

Rysunek 5 przedstawia przykładowe przekroje z każdego ze zbioru danych po zastosowaniu opisanych metod przetwarzania wstępnego.



Rys. 5. Przekroje badań D1, K5 oraz R10 po przetworzeniu wstępnym

Domyślnie stosowanym w nnU-Net optymalizatorem jest SGD (ang. *stochastic gradient descent*). Innym, bardzo często spotykanym, jest Adam (ang. *adaptive moment estimation*). Chociaż obie metody dążą do tego samego celu, jakim jest minimalizacja

funkcji kosztu, ich szybkość zbieżności może być różna, zależnie od zadania (Dogo i in. 2018). Z tego powodu warto przeprowadzić test dla konkretnych danych.

Na podstawie zebranych danych stwierdzono, że wynik dla obu metod optymalizacji jest zbliżony, różnica uzyskanych oboma metodami współczynników Dice'a wyniosła 0,01. W opisywanym przypadku nie zdecydowano się na dalsze testowanie, ponieważ obie metody dają zbliżone wyniki. W konsekwencji wybrano domyślnie proponowany algorytm SGD, wraz z domyślną początkową wartością współczynnika uczenia wynoszącą 0,01.

Obrazy tomograficzne składają się z przekrojów (ang. *slices*), w których wartości najczęściej przypisane są do kwadratowych pikseli. Dodatkowym parametrem jest grubość przekroju (ang. *thickness*), która umożliwia skonstruowanie przestrzennego woksela danych. Wymiary przestrzenne woksela reprezentują rzeczywisty rozmiar obszaru, którego uśredniona wartość jest zapisywana jako wartość woksela (Tino i in. 2021). Taki sposób zapisu danych pozwala przekształcić wielkość woksela na rzeczywiste jednostki odległości, a tym samym dokonywać pomiarów wielkości zadanych struktur. Objętość przestrzenna, którą zajmuje woxel, jest różna w zależności od sposobu przeprowadzania badania. Oznacza to, że obrazy mające tę samą rozdzielczość mogą prezentować zupełnie różne rzeczywiste wymiary, skąd wynika potrzeba ujednolicenia wymiarów woksela przed rozpoczęciem treningu.

Rozdzielczość obrazowania oraz wymiary woksela są używane do obliczenia rzeczywistych rozmiarów, przez co są ze sobą powiązane według wzoru (4).

$$L_k = N_k \cdot \Delta_k \quad (4)$$

gdzie:

- L_k – fizyczna długość obrazu wzdłuż osi k (mm),
- N_k – liczba wokseli (rozdzielczość) wzdłuż osi k ,
- Δ_k – wielkość pojedynczego woksela (ang. *spacing*) wzdłuż tej osi (mm).

Framework nnU-Net zachowuje oryginalną rozdzielczość obrazów, natomiast dla każdego dopasowuje wymiary woksela. Dla danych treningowych trójwymiarowych, niezależnie dla każdej osi, wyszukiwana jest mediana wymiarów woksela. Wartość ta staje się nową wartością docelową dla wszystkich obrazowań. W przypadku treningu dla sieci stosującej dane dwuwymiarowe (przekroje) wybierane są przekroje względem tych osi, które mają najwyższą rozdzielczość. Jeśli wszystkie osie są izotropowe (dla każdej osi rozdzielczość jest taka sama), przekroje wybierane są względem dwóch ostatnich osi (Huang i in. 2022, Isensee i in. 2021). Interpolacja wykonywana jest z wykorzystaniem funkcji sklejanych trzeciego stopnia (Isensee i in. 2019b).

Dobór optymalnej architektury sieci do problemu stanowi duże wyzwanie, dlatego też twórcy frameworku nnU-Net zautomatyzowali to zadanie. Zaimplementowano sys-

tem ekspertowy, który najpierw wyznacza *dataset fingerprint* (kluczowe właściwości zbioru), a następnie na podstawie zestawu reguł automatycznie konfiguruje sposób przetwarzania danych, w tym wstępne przetwarzanie, architekturę sieci U-Net, parametry treningu oraz końcowe przetwarzanie.

Ręcznie podzielono wszystkie 56 obrazowań na zbiór treningowy oraz testowy w stosunku 0,9 względem przypadków, co dało 48 przypadków treningowych (85,7%) oraz 8 testowych (14,3%). Do zbioru treningowego dodano badania R1, R7 oraz R8, pochodzące ze zbioru RIDER, zawierające obrazowania z rozwarstwienia aorty, a także badanie R6, zawierające obrazowanie z tętniakami aorty. Do zbioru testowego przydzielono badania R2 oraz R5, również należące do zbioru RIDER, zawierające obrazowania z rozwarstwienia aorty. Algorytm tworzący podziały służące do realizacji walidacji krzyżowej podzielił zbiór treningowy względem liczby obrazowań w sposób przedstawiony w tabeli 2.

Tabela 2
Podział walidacji krzyżowej względem liczności zbiorów

	Obrazowania treningowe	Obrazowania walidacyjne
Model 1	38	10
Model 2	38	10
Model 3	38	10
Model 4	39	9
Model 5	39	9

Do oceny stworzonego rozwiązania użyto metryk Dice’a Score (5), IoU (6) oraz odległości Hausdorffa (7). Dzięki zastosowaniu różnorodnych metryk oraz oceny jakościowej możliwa jest dokładna i kompleksowa ocena modelu. Każda z metryk uwypukla inne aspekty, co przekłada się na pełniejsze zrozumienie możliwości stworzonego narzędzia oraz identyfikację jego mocnych i słabych stron, a więc umożliwia określenie dalszego kierunku rozwoju.

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (5)$$

gdzie:

A – obraz maski wyliczony jako predykcja modelu,

B – obraz prawidłowej maski, dostarczony wraz ze zbiorem,

$|A \cap B|$ – część wspólna obu masek,

$|A| + |B|$ – suma arytmetyczna pikseli należących do masek A oraz B .

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

gdzie:

$|A \cap B|$ – część wspólna obu masek,

$|A \cup B|$ – suma logiczna pikseli należących do masek A oraz B .

$$H(A, B) = \max\left(\sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(b, a)\right) \quad (7)$$

gdzie:

$d(a, b), d(b, a)$ – odległość euklidesowa pomiędzy punktami a oraz b ,

$\inf_{b \in B} d(a, b)$ – infimum odległości (najmniejsza możliwa odległość) pomiędzy punktem a ze zbioru A a wszystkimi punktami b ze zbioru B (analogicznie dla $\inf_{a \in A} d(b, a)$),

$\sup_{a \in A} \inf_{b \in B} d(a, b)$ – supremum odległości (największa możliwa odległość) dla każdego punktu a ze zbioru A ze zbioru wcześniej obliczonych infimów (analogicznie dla $\sup_{b \in B} \inf_{a \in A} d(b, a)$),

$\max()$ – funkcja zwracająca maksimum z obu obliczonych wartości.

Oprócz klasycznej odległości Hausdorffa raportuje się również metrykę Hausdorff95, czyli 95. percentyl rozkładu odległości między granicami (powierzchniami) dwóch masek segmentacji. Metryka ta ogranicza wpływ pojedynczych wartości odstających.

3. Wyniki

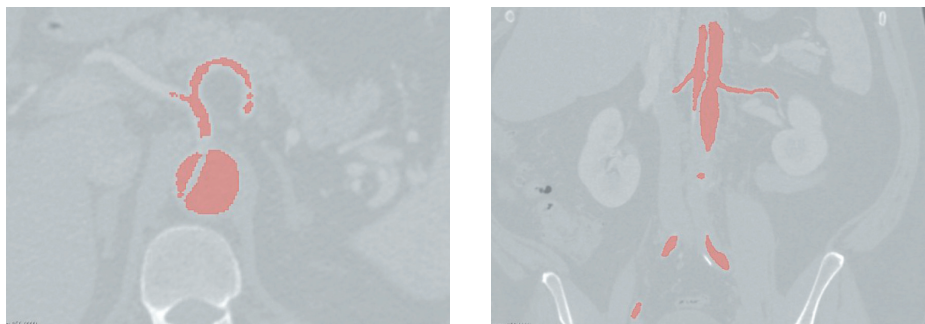
Wyniki przewidywań zbioru testowego z pięciu modeli oraz wyniki uśrednione przez nnU-Net przedstawia tabela 3.

Tabela 3
Metryki dla poszczególnych modeli

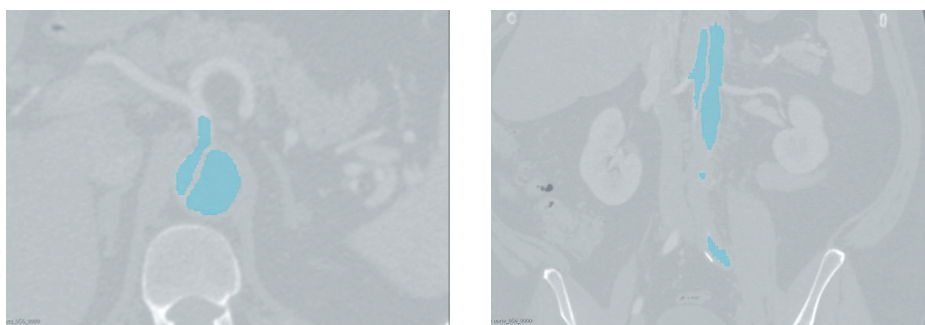
	Model 1	Model 2	Model 3	Model 4	Model 5	Model uśredniony
Dice	0,919	0,895	0,921	0,916	0,908	0,921
IoU	0,855	0,824	0,858	0,849	0,837	0,858
Hausdorff95 [mm]	6,233	12,467	6,013	13,784	12,770	5,870
Hausdorff [mm]	152,876	131,122	133,5	117,431	130,914	114,208

W celu ewaluacji końcowego rozwiązania zwizualizowano powstałe segmentacje oraz porównano je z prawidłowymi maskami, co umożliwiło ocenę problematycznych miejsc.

Jednym ze znanych problemów rodziny architektur U-Net jest trudność z odwzorowaniem szczegółów (Ahmed i in. 2025, Valanarasu i in. 2020). W przypadku używanego zbioru danych sytuacja występuje przy mniejszych odgałęzieniach, co pokazują rysunek 6 oraz rysunek 7.



Rys. 6. Prawidłowa segmentacja odgałęzienia

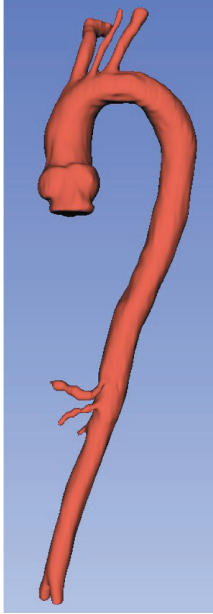
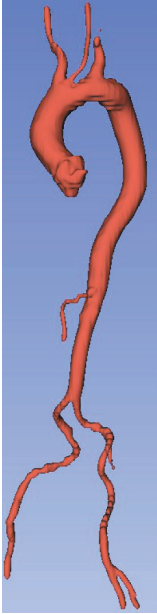
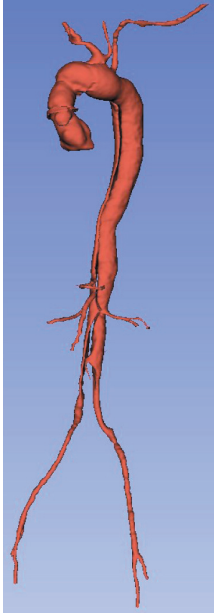
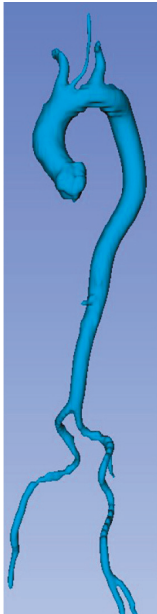
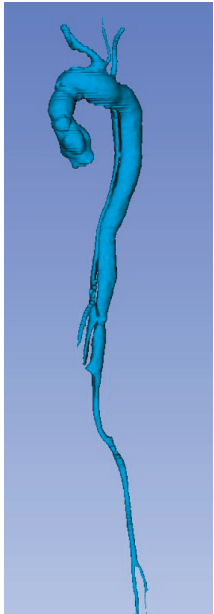


Rys. 7. Segmentacja utworzona przez sieć

Tabela 4 przedstawia trójwymiarowe modele, stworzone na podstawie dostarczonych ze zbiorem danych (czerwone) oraz wygenerowanych przez sieć (niebieskie) masek segmentacji. Na ich podstawie można potwierdzić przypuszczenie, że model dobrze odnajduje duże struktury, natomiast znacznie gorzej przewidyuje odgałęzienia.

Wartość współczynnika Dice'a na poziomie 0,921 oraz IoU wynoszące 0,858 oznaczają stosunkowo wysoką dokładność przewidywań, co zweryfikowano również na podstawie wizualizacji wyników jako trójwymiarowych modeli. Odległość Hausdorffa wskazuje na obecność elementów odstających, natomiast 95. percentyl o wartości 5,870 mm oznacza stosunkowo wysoką jakość przewidywań w dominującej objętości struktury. Stworzone rozwiązanie dobrze realizuje zadanie, jakim jest segmentacja aorty. Jeśli celem predykcji jest również wyodrębnienie odgałęzień oraz mniejszych naczyń, zachodzi potrzeba zweryfikowania zastosowanego podejścia. Ograniczeniem metody jest brak możliwości odwzorowania małych naczyń odchodzących od aorty.

Tabela 4
Porównanie modeli 3D prawidłowych oraz predykcji

	Dongyang	KiTS	RIDER
Model 3D stworzony na podstawie dostarczonych masek			
Model 3D stworzony na podstawie predykcji sieci			

4. Podsumowanie oraz wnioski

Zaprezentowane rozwiązanie miało za zadanie realizować segmentację aorty z obrazów tomografii komputerowej. Otrzymane wyniki potwierdzają skuteczność metody w zakresie segmentacji aorty, nawet w trudniejszych przypadkach, takich jak jej rozwarstwienie. Można stwierdzić, że segmentacja większych, jednolitych struktur działa zadowalająco, w przeciwieństwie do odgałęzień i drobniejszych naczyń.

Framework nnU-Net jest odpowiednim narzędziem, jeżeli celem segmentacji jest jednolita, dużych rozmiarów struktura. Ze względu na uproszczenie wielu etapów tworzenia rozwiązania opartego na sieciach neuronowych, a także automatyczne wykrywanie rodzaju danych treningowych i dopasowywanie do nich architektury, nnU-Net jest przydatnym i wygodnym w użyciu rozwiązaniem. Kluczowym elementem jest prawidłowe przeprowadzenie przetwarzania wstępnego umożliwiającego generalizację modelu dla danych z różnych ośrodków.

Dalszym obszarem badań może być opracowanie rozwiązania, które wychwytyje globalny kontekst i jednocześnie dba o szczegóły. Możliwe jest również wykorzystanie mechanizmu uwagi, zastosowanie augmentacji poprawiających kontrast drobnych struktur czy zastosowanie funkcji kosztu minimalizującej błędy topologiczne.

Literatura

- Ahmed R.M., Iltaf A., Elmann M., Zhao G., Li H., Du Y., Li B., Zhou S., 2025, *MSA-UNet3+: Multi-scale attention UNet3+ with new supervised prototypical contrastive loss for coronary DSA image segmentation*, arXiv:2504.05184v3. <https://arxiv.org/pdf/2504.05184v3>.
- Balmer B.D., Blüthgen C., Bäessler B., Martini K., Huber F.A., Ruby L., Schönenberger A., Frauenfelder T., 2022, *Influence of CT image matrix size and kernel type on the assessment of HRCT in patients with SSC-ILD*, *Diagnostics*, vol. 12(7), 1662. <https://doi.org/10.3390/diagnostics12071662>.
- Cabana A., Lago-Fernández L.F., 2021, *Backward gradient normalization in deep neural networks*, arXiv:2106.09475. <https://doi.org/10.48550/arXiv.2106.09475>.
- Dogo E.M., Afolabi O.J., Nwulu N.I., Twala B., Aigbavboa C.O., 2018, *A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks*, [w:] 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), IEEE, s. 92–99. <https://doi.org/10.1109/CTEMS.2018.8769211>.
- Gücük A., Üyetürk U., 2014, *Usefulness of Hounsfield unit and density in the assessment and treatment of urinary stones*, *World Journal of Nephrology*, vol. 3(4), s. 282–286. <https://doi.org/10.5527/wjn.v3.i4.282>.
- Hsieh J., Flohr T., 2021, *Computed tomography recent history and future perspectives*, *Journal of Medical Imaging*, vol. 8(5), 052109. <https://doi.org/10.1117/1.JMI.8.5.052109>.

- Huang Z., Wang H., Ye J., Niu J., Tu C., Yang Y., Du S., Deng Z., Gu L., He J., 2022, *Revisiting nnU-Net for iterative pseudo labeling and efficient sliding window inference*, [w:] Ma J., Wang B. (Eds.), *Fast and Low-Resource Semi-supervised Abdominal Organ Segmentation*, FLARE 2022. Lecture Notes in Computer Science, vol. 13816, s. 178–189. https://doi.org/10.1007/978-3-031-23911-3_16.
- Isensee F., Jäger P.F., Kohl S.A.A., Petersen J., Maier-Hein K.H., 2019a, *Automated design of deep learning methods for biomedical image segmentation*, arXiv:1904.08128v2. <https://doi.org/10.48550/arXiv.1904.08128>.
- Isensee F., Petersen J., Kohl S.A.A. Jäger P.F., Maier-Hein K.H., 2019b, *nnU-Net: Breaking the spell on successful medical image segmentation*, arXiv:1904.08128v1. <https://arxiv.org/abs/1904.08128v1>.
- Isensee F., Jaeger P.F., Kohl S.A.A., Petersen J., Maier-Hein K.H., 2021, *nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation*, Nature Methods, vol. 18(2), s. 203–211. <https://doi.org/10.1038/s41592-020-01008-z>.
- Isensee F., Wald T., Ulrich C., Baumgartner M., Roy S., Maier-Hein K., Jäger P., 2024, *nnU-Net revisited: A call for rigorous validation in 3D medical image segmentation* arXiv:2404.09556v2. <https://doi.org/10.48550/arXiv.2404.09556>.
- Jhala V.M., Thaker V.S., 2015, *X-ray computed tomography to study rice (Oryza sativa L.) panicle development*, Journal of Experimental Botany, vol. 66(21), s. 6819–6825. <https://doi.org/10.1093/jxb/erv387>.
- Radl L., Jin Y., Pepe A., Li J., Gsaxner C., Zhao F., Egger J., 2022, *AVT: Multicenter aortic vessel tree CTA dataset collection with ground truth segmentation masks*, Data in Brief, vol. 40, 107801. <https://doi.org/10.1016/j.dib.2022.107801>.
- Ronneberger O., Fischer P., Brox T., 2015, [w:] Navab N., Hornegger J., Wells W.M., Frangi A.F. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI. 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18, Lecture Notes in Computer Science, vol. 9351, s. 234–241, arXiv:15.05.04597v1. <https://doi.org/10.48550/arXiv.1505.04597>.
- Tino R., Yeo A., Brandt M., Leary M., Kron T., 2021, *The interlace deposition method of bone equivalent material extrusion 3D printing for imaging in radiotherapy*, Materials & Design, vol. 199, 109439. <https://doi.org/10.1016/j.matdes.2020.109439>.
- Valanarasu J.M.J., Sindagi V.A., Hacihaliloglu I., Patel V.M., 2020, *Kiu-net: Towards accurate segmentation of biomedical images using over-complete representations*, [w:] Martel A.L., Abolmaesumi P., Stoyanov D., Mateus D., Zuluaga M.A., Zhou S.K., Racocanu D., Joskowicz L. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV*, Lecture Notes in Computer Science, vol. 12263, s. 363–373.

Projekt i realizacja tunelu aerodynamicznego o obiegu zamkniętym

Michał Przepiórski , Konrad Staruch 

Akademia Górniczo-Hutnicza, Wydział Energetyki i Paliw, Kraków

Streszczenie: W pracy opisano projekt i realizację tunelu aerodynamicznego o obiegu zamkniętym, przeznaczonego do badań metodą PIV. Tunel posiada przekrój roboczy $0,8\text{ m} \times 0,8\text{ m}$ i umożliwia osiągnięcie maksymalnej prędkości roboczej przepływu do 10 m/s . Całość zaprojektowano w środowisku CAD, a kluczowe komponenty zaprojektowano metodą CFD.

Słowa kluczowe: tunel aerodynamiczny, obieg zamknięty, metoda PIV, aerodynamika, CAD, CFD

DESIGN AND CONSTRUCTION OF A CLOSED-LOOP WIND TUNNEL

Abstract: This paper describes the design and implementation of a closed-circuit wind tunnel for PIV testing. The tunnel has a working cross-section of $0.8\text{ m} \times 0.8\text{ m}$ and allows for flow velocities of up to 10 m/s . The entire system was designed in a CAD environment, and key components were designed using the CFD method.

Keywords: wind tunnel, closed-circuit design, PIV (Particle Image Velocimetry), aerodynamics, CAD, CFD

https://doi.org/10.7494/978-83-68219-80-7_8

1. Wstęp

Od lotnictwa do konstrukcji lądowych aerodynamika decyduje o wydajności maszyn czy planowaniu przestrzennym miast. Do badania właściwości opływanych przez powietrze obiektów służą tunele aerodynamiczne. Zasada działania tunelu aerodynamicznego jest prosta. Powietrze wprowadzone w ruch opływa próbkę, symulując interakcję pomiędzy rzeczywistym obiektem i płynem. Istnieją dwa typy konstrukcji tuneli aerodynamicznych: tunele otwarte i zamknięte (Ahmed 2013). W tunelu otwartym powietrze jest zasysane, a po przejściu przez część roboczą jest wypuszczane z powrotem do otoczenia. W konstrukcjach zamkniętych cały czas używany jest ten sam gaz (brak wymiany płynu z otoczeniem). Główną zaletą tunelu otwartego jest jego prostsza względem tunelu zamkniętego konstrukcja, która umożliwia szybszą i tańszą budowę, a potem obsługę, niż w przypadku tunelu zamkniętego. Przykłady tuneli otwartych przedstawione są w pracach Jewel Barlowa i współautorów (1999) oraz Ryszarda Szwaiby i współautorów (2021). Tunel zamknięty, w związku z bardziej skomplikowaną budową, wymaga większych nakładów finansowych. Wyzwaniami są utrzymanie szczelności w obwodzie i opracowanie systemu, który umożliwia umiejscowienie próbki. Jednocześnie tunel zamknięty oferuje wyższą sprawność energetyczną i nie rozrzuca posiewu laboratoryjnego poza tunel (Ahmed 2013).

Różne sposoby badawcze umożliwiają opisanie zachowania próbki oraz opływającego ją płynu. Układ pomiarowy z dwóch wag umożliwia wyznaczenie współczynnika oporu i nośności próbki (Wood i in. 2025). Zaawansowane metody badawcze, takie jak np. PIV (*Particle Image Velocimetry*), umożliwiają wyznaczenie rozkładu prędkości. Metoda PIV opiera się na śledzeniu trajektorii cząsteczek wpuszczonych do płynu znajdującego się w ruchu. System pomiarowy mierzy przemieszczenie cząsteczek w bardzo krótkich odstępach czasu, dzięki czemu wyznacza pole prędkości płynu (LaVision 2025).

Głównymi parametrami opisującymi ruch płynu w tunelu aerodynamicznym są liczby bezwymiarowe, takie jak liczba Reynoldsa czy liczba Macha (Cattafesta i in. 2010). Wykorzystanie odpowiednio skalowanych modeli, mieszczących się w komorze pomiarowej, jest możliwe dzięki zastosowaniu twierdzenia Buckinghama. Umożliwia ono grupowanie wielkości fizycznych charakteryzujących dany przepływ wokół próbki oraz – przy zachowaniu takich samych wartości liczb bezwymiarowych – odtworzenie warunków rzeczywistych w przeskalowanym modelu.

W niniejszej pracy przedstawiono procedurę projektową oraz wykonawczą tunelu aerodynamicznego niskich prędkości. Opierając się na tym, dla jakich obiektów przeznaczony jest tunel, omówiono jego konstrukcję oraz dobrano moc i wielkość wentylatora.

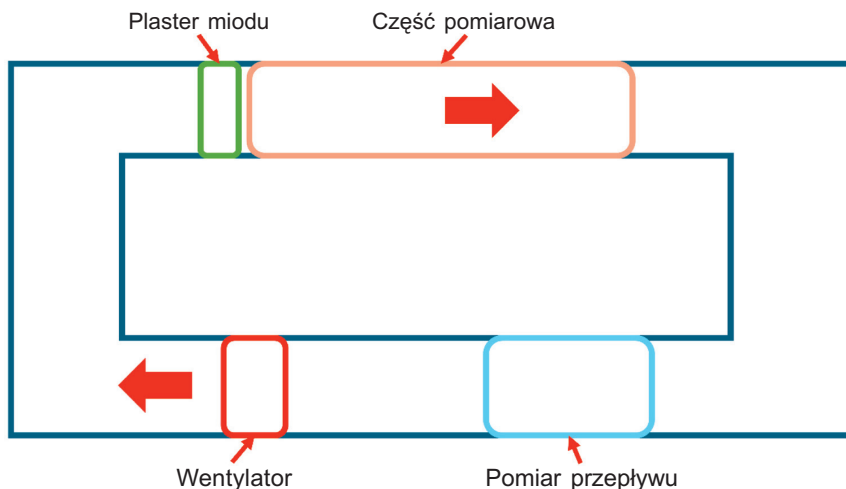
2. Metodologia

2.1. Pomiary

Tunel będzie współpracował z systemem PIV. Kamery PIV wymagają odpowiedniego ustawienia względem próbki pomiarowej (LaVision 2025), a cały tunel musi być przystosowany do współpracy z posiewem. Zebrane doświadczenia laboratoryjne wskazują na wykorzystanie szkła jako preferowanego materiału oddzielającego przepływ tunelu od aparatury PIV. Aby uzyskać miarodajne wyniki pomiarowe, przepływ wokół próbki musi być maksymalnie jednorodny, gdyż każda turbulencja w przepływie jest wizualizowana przez PIV.

2.2. Koncepcja tunelu

Konstrukcja tunelu opiera się na koncepcji zamkniętej (gaz pracujący nie jest wymieniany z otoczeniem). Przyjęta idea tunelu typu zamkniętego pozwala zatrzymać posiew wykorzystywany w metodzie PIV wewnątrz konstrukcji. Brak wymiany powietrza z otoczeniem pozwoli również na podniesienie sprawności energetycznej tunelu względem projektów typu otwartego (Ahmed 2013). Rysunek 1 prezentuje uproszczony rzut z góry omawianego obiektu. W segmencie pomiarowym tunelu lokalizuje się badaną próbkę. Znajdująca się pomiędzy wentylatorem a segmentem pomiarowym prostownica przepływu (plaster miodu) służy do ujednolicenia i nadania odpowiednich właściwości przepływowi turbulentnemu. Ostatnim elementem tunelu jest sekcja pomiaru właściwości przepływu. Sekcja odpowiada za pomiar prędkości w określonych miejscach przekroju kanału, tak aby uzyskać reprezentatywną wartość średniej prędkości powietrza w tunelu.

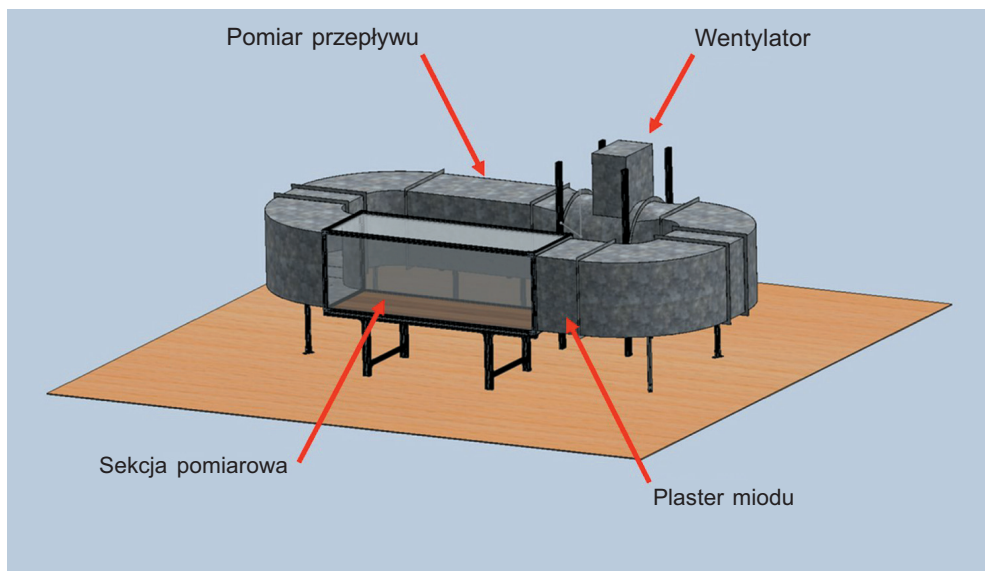


Rys. 1. Schemat przedstawiający rzut od góry tunelu aerodynamicznego typu zamkniętego. Czerwona strzałka oznacza kierunek przepływu powietrza

Wejściowe wielkości opisujące tunel zostały przedstawione w tabeli 1. Poniższe założenia dotyczące wymiarów oraz prędkości w tunelu zostały przyjęte na podstawie doświadczeń zebranych podczas badań aerodynamicznych zrealizowanych w przeszłości oraz ogólnej charakterystyki aparatury pomiarowej omówionej w podrozdziale 2.1. Krytycznym miejscem konstrukcji jest prześwit pomiędzy komorą pomiarową a wentylatorem. Przy założonych maksymalnych wymiarach zewnętrznych prześwit ten wynosi 1,4 m, co umożliwi odpowiednie umiejscowienie kamery PIV. Wszystkie elementy i podzespoły tunelu zostały zaprojektowane w środowisku CAD (SolidWorks). Wizualizacja projektu tunelu została przedstawiona na rysunku 2.

Tabela 1
Wielkości opisujące tunel

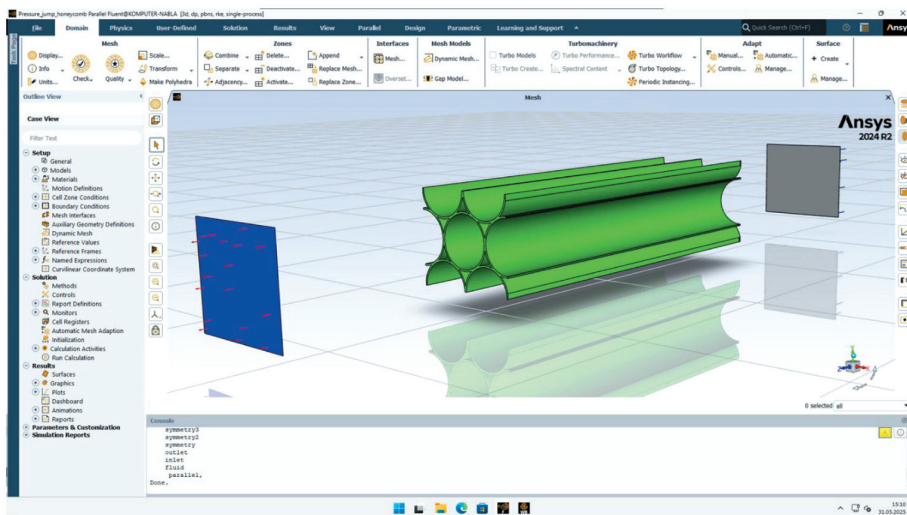
Wielkość opisująca tunel	Wartość liczbową
Maksymalna szerokość tunelu	3 m
Maksymalna długość tunelu	6 m
Maksymalna wysokość tunelu	2,5 m
Długość sekcji pomiarowej	2,5 m
Przekrój roboczy tunelu	0,8 m × 0,8 m
Prędkość maksymalna czynnika	10 m/s



Rys. 2. Wizualizacja tunelu aerodynamicznego w programie SolidWorks

2.3. Prostownica przepływu

Prostownice przepływu używane są w tunelach aerodynamicznych w celu poprawy jednolitości przepływu w sekcji pomiarowej. Podczas konstruowania tunelu założono przypadek prostownicy przepływu, w którym porowatość wynosi 0,9, grubość ścianek komórek wynosi 1,5 mm, kształt komórek został założony jako kołowy, a średnica wewnętrzna pojedynczej komórki wynosi 47 mm. Dobrana długość prostownicy przepływu wynosi 350 mm. Wszystkie wymiary prostownicy przepływu zostały dobrane na podstawie książki Noora Ahmeda (2013). Wizualizacja wycinka plastra miodu jest przedstawiona na rysunku 3.



Rys. 3. Zrzut ekranu przedstawiający wizualizację wycinka prostownicy przepływu w programie Ansys Fluent
Źródło: Ansys (2025)

2.4. Obliczenia wstępne

Sercem każdego tunelu aerodynamicznego jest maszyna wprawiająca płyn w ruch. W przypadku tunelu opisywanego w niniejszej pracy zdecydowano o wyborze wentylatora jako maszyny pędnej. Aby dobrać odpowiedni wentylator, wyznaczono spadek ciśnienia płynu przemieszczającego się w pętli tunelu przy założonym wydatku masowym. Do obliczeń wykorzystano przyjęte parametry wejściowe (tab. 1). Wydatek masowy można obliczyć z równania zachowania masy dla płynów zapisanego w postaci równania (1):

$$v \cdot A = Q \quad (1)$$

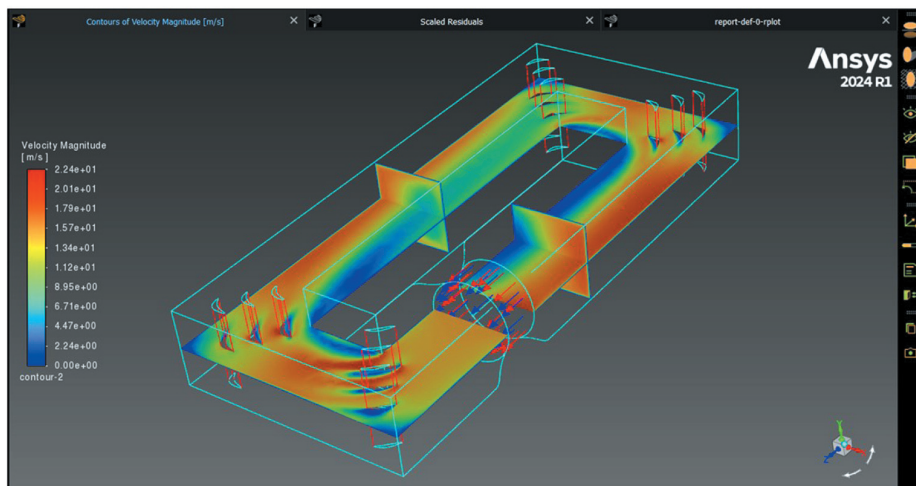
gdzie:

v – prędkość [m/s],

A – pole powierzchni [m²],

Q – wydatek objętościowy powietrza [m³/h].

Zgodnie z równaniem (1) wydatek objętościowy przepływu powietrza przy założonej prędkości czynnika oraz przekroju roboczym wynosi $23\,040\text{ m}^3/\text{h}$. Spadek ciśnienia może być obliczony za pomocą metod analitycznych (Ahmed 2013) lub metod symulacji komputerowych CFD (*Computational Fluid Dynamics*) (Yong i Dol 2015). Dzięki dostępowi do zasobów komputerowych oraz programu Ansys Fluent 2024 R1 (Ansys 2025) wybrano metodykę CFD. W symulacji uwzględniono rozmiary tunelu oraz obecność kierownic w kolankach. Wyznaczone pole prędkości (rys. 4) umożliwiło obliczenie spadku ciśnienia pomiędzy wlotem a wylotem wentylatora. Wyniki symulacji porównano z przybliżonymi wynikami teoretycznymi, wykorzystując zależności analityczne (prawo zachowania energii w przepływie) i tabele współczynników spadku ciśnienia.



Rys. 4. Zrzut ekranu przedstawiający wizualizację pola prędkości w tunelu aerodynamicznym. Wizualizacja powstała w programie Ansys Fluent
Źródło: Ansys (2025)

Kolejnym elementem potrzebnym do oszacowania spiętrzenia wentylatora jest wyznaczenie spadku ciśnienia w prostownicy przepływu znajdującej się przed częścią testową tunelu (rys. 2). Spadek ciśnienia w prostownicy przepływu również wyznaczono za pomocą symulacji komputerowych CFD. Plaster miodu potraktowano jako geometrię składającą się ze skończonej liczby identycznych wycinków. Zostały one wyznaczone przez przecięcie geometrii reprezentującej prostownicę przepływu przez skończoną liczbę płaszczyzn symetrii. Plaster miodu po „przecięciu” domeny obliczeniowej jest przedstawiony na rysunku 3. Przy założonym warunku symetrii przepływu na płaszczyznach symetrii plastra miodu odwzorowano spadek ciśnienia między wlotem a wylotem prostownicy przepływu. Wyniki symulacji porównano z przybliżonymi wynikami teoretycznymi, wykorzystując analityczne zależności (prawo Darcy’ego). Przyjęto założenie, że spadki ciśnienia w tunelu oraz w prostownicy przepływu można zsumować i dzięki temu uzyskać finalne spiętrzenie wentylatora przy zadanym wydatku

objętościowym. Obliczone metodą CFD spiętrzenie wentylatora wynosi 200 Pa. Mając na uwadze przyszły rozwój konstrukcji oraz niedoszacowanie spadków ciśnienia metodą CFD, przemnożono uzyskane spiętrzenie wentylatora przez współczynnik przewymiarowania wynoszący 1,5 i uzyskano wartość 300 Pa. Współczynnik przewymiarowania został dobrany arbitralnie przy uwzględnieniu błędów obliczeń CFD oraz przyszłej przebudowy tunelu w celu uzyskania wyższych prędkości przepływu.

3. Wyniki

3.1. Konstrukcja tunelu

Tunel aerodynamiczny składa się z czterech kanałów prostych oraz z czterech kolanek. Kanał, w którym znajduje się prostownica przepływu, ma długość 0,5 m. Długość pozostałych przewodów oraz kolanek wentylacyjnych została dobrana jako zmienna wynikowa w celu zachowania parametrów opisanych w tabeli 1. Sekcja pomiarowa tunelu została złożona z profili aluminiowych konstrukcyjnych o długości boku 40 mm. Szyby hartowane 8 mm zostały wykorzystane do wypełnienia ścian bocznych i górnej w konstrukcji sekcji pomiarowej. Zastosowano szkło z uwagi na dobrą charakterystykę współpracy z laserem (podrozdz. 2.1). Sekcja testowa, wkomponowana w pozostałe podzespoły tunelu, została zaprezentowana na rysunku 5, który przedstawia sekcję bez włożonych wypełnień ścian bocznych, dolnej i górnej.



Rys. 5. Sekcja testowa tunelu aerodynamicznego ze zdemontowanymi wypełnieniami ścian bocznych

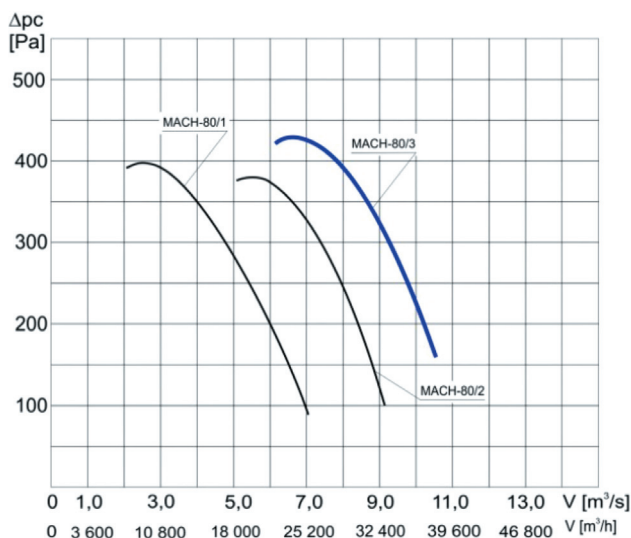
Wszystkie kanały zostały skręcone za pomocą śrub, a pomiędzy każdym kanałem umieszczono uszczelkę gąbkową samoprzylepną o grubości 16 mm w celu utrzymania szczelności w obwodzie. Każdy kanał ma swoją podporę zbudowaną z profili aluminiowych. Rysunek 6 przedstawia cały tunel po złożeniu, lecz bez zainstalowanego wentylatora.



Rys. 6. Zdjęcie zbudowanego tunelu aerodynamicznego.
Na zdjęciu tunel pozbawiony jest wentylatora oraz ścian bocznych w sekcji pomiarowej

3.2. Wentylator

Wentylator dobrano na podstawie obliczeń wykonanych w podrozdziale 2.4 Przyjęty spręż maksymalny wentylatora to 300 Pa przy wydatku objętościowym równym 23 040 m³/h. Zgodność z wyliczonymi parametrami przepływu oferuje wentylator firmy Tywent model MACH-80/3 3F (Tywent 2025). Na rysunku 7 przedstawiona została charakterystyka spiętrzeniowa wentylatora. Porównanie teoretycznych obliczeń z charakterystyką wentylatora wskazuje na przewymiarowanie maszyny względem założonych nominalnych parametrów pracy czynnika roboczego. W celu przystosowania wentylatora do obliczonych parametrów pracy zastosowano inwerter.



Rys. 7. Charakterystyka spiętrzeniowa wentylatora Tywent MACH-80/3 3F (zaznaczona niebieską linią)
Źródło: Tywent (2025)

4. Wnioski i podsumowanie

W pracy przedstawiono procedurę projektową oraz wykonawczą tunelu aerodynamicznego niskich prędkości. Wykorzystując oprogramowania CAD oraz CFD, skonstruowano tunel oraz dobrano moc i wielkość wentylatora. Następnie przedstawiono rzeczywistą konstrukcję złożoną w laboratorium. Zbudowana konstrukcja mieści się w założonych maksymalnych wymiarach tunelu. Rzeczywista długość tunelu to 5,3 m, szerokość to 2,7 m, a prześwit pomiędzy kanałami to 1,039 m. Tunel ma wymiary mieszczące się w tych maksymalnych opisanych w tabeli 1. Tunel ma kwadratowy przekrój roboczy $0,8\text{ m} \times 0,8\text{ m}$, a uzyskana prędkość robocza czynnika wynosi 10 m/s. Koszt budowy instalacji wyniósł 30 000 zł (podana kwota nie zawiera ceny aparatury pomiarowej PIV). Budowa tunelu, bez czasu potrzebnego na dostawę zamówionych części, zajęła dwa dni.

Przewymiarowany wentylator oraz modułowa, skręcana konstrukcja umożliwiającą zmianę parametrów pracy tunelu. Zmniejszenie powierzchni poprzecznej przepływu pozwoli na zwiększenie jego prędkości (prędkość powietrza rośnie odwrotnie proporcjonalnie do powierzchni poprzecznej przepływu). Po przebudowie tunel będzie mógł służyć nie tylko do badań urbanistycznych, ale również do badań z dziedziny motoryzacji. Tunel dodatkowo wymaga opracowania modułu pomiarowego próbek. Moduł ten będzie uzupełnieniem systemu PIV, pozwalającym wykonywać proste badania współczynników nośności i oporu dzięki czujnikom tensometrycznym. W przyszłości konstrukcja tunelu powinna zostać wyposażona w chłodnicę w celu zredukowania efektu nagrzewania się powietrza przez sprężający je wentylator.

Literatura

- Ahmed N., 2013, *Wind tunnel designs and their diverse engineering applications*, InTech. <https://doi.org/10.5772/3403>.
- Ansys 2025, Ansys Fluent. <https://www.ansys.com/products/fluids/ansys-fluent> [dostęp: 12.06.2025].
- Barlow J., Rae W.H. Jr., Pope A., 1999, *Low-speed Wind Tunnel Testing*, Third Edition, John Wiley & Sons, New York – Chichester – Weinheim – Brisbane – Singapore – Toronto.
- Cattafesta L., Bahr C. Mathew J., 2010, *Fundamentals of wind-tunnel design*, [w:] Blockley R., Shyy W. (Eds.), *Encyclopedia of Aerospace Engineering*, John Wiley & Sons, s. 1–10. https://www.researchgate.net/profile/Louis-Cattafesta/publication/230271809_Fundamentals_of_Wind-Tunnel_Design/links/5a0d7a51a6fdcc39e9bfe-324/Fundamentals-of-Wind-Tunnel-Design.pdf.
- LaVision, *Particle Image Velocimetry*. <https://www.lavision.de/en/techniques/piv-ptv/index.php> [dostęp: 12.06.2025].

- Symkaton, Symkaton. <https://symkom.pl/symkaton/> [dostęp: 12.06.2025].
- Szwaba R., Hinc K., Ochrymiuk Z., Krzemianowski Z., Doerffer P., Kurowski M., 2021, *Open low speed wind tunnel – design and testing*, Archives of Thermodynamics, vol. 42(1), s. 57–70. <https://doi.org/10.24425/ather.2021.136947>.
- Tywent, Tywent MACH-80/3 3F. <https://tywent.pl/mach/476-wentylator-kana-owy-mach-80-3-3f-16342.html> [dostęp: 12.06.2025].
- Wood M., Detlefs J., Arnold N., 2025, *2C20.50 – Airfoil Lift and Drag*. <https://www.youtube.com/watch?v=EbOpgUrwK7M> [dostęp: 12.06.2025].
- Yong T.H., Dol S.S. 2015, *Design and development of low-cost wind tunnel for educational purpose*, 78, IOP Conference Series: Materials Science and Engineering, vol. 78, 012039. 10.1088/1757-899X/78/1/012039.

Szczelność naftowych otworów wiertniczych w świetle powierzchniowych badań geochemicznych – studium przypadku z monokliny przedsudeckiej

Henryk Sechman¹ , Sandra Płochocka², Jan Adam Sobon¹ 

¹ Akademia Górniczo-Hutnicza, Wydział Geologii, Geofizyki i Ochrony Środowiska, Kraków

² Shell Polska Sp. z o.o. Oddział w Krakowie, ul. Czerwone Maki 85, 30-392 Kraków

Streszczenie: Celem pracy jest ocena integralności naftowego otworu wiertniczego na monoklinie przedsudeckiej, oparta na wynikach powierzchniowych badań geochemicznych gazów glebowych zrealizowanych metodą „gazu wolnego”. W 16 punktach rozmieszczonych wokół głowicy otworu pobrano próbki z głębokości około 1,2 m i poddano je analizie metodą chromatografii gazowej (FID/TCD) w zakresie oznaczenia: CH₄, alkanów C₂–C₅, gazowych alkenów, H₂, O₂, N₂ i CO₂. Zbiory stężeń metanu (0,9–30 ppm) i etanu (0,012–3,657 ppm) wykazują rozkłady asymetryczne, a wartości sumy alkanów C₂–C₅ powyżej 0,2 ppm określono jako anomalne. Korelacje CH₄–C₂H₆–C₃H₈ ($r \approx 0,93$ – $0,96$) wskazują na wspólne, głębokie źródło emanacji węglowodorów. Strefy anomalne alkanów zostały zarejestrowane w bezpośrednim sąsiedztwie głowicy otworu. Analiza wskaźników geochemicznych wykazała podobieństwo składu gazu glebowego do gazu ziemnego pobranego z otworu, co potwierdza obecność jego śladowych ilości w próbkach. Maksymalne stężenia: 30 ppm CH₄ oraz 4,57 ppm sumy alkanów C₂–C₅ nie są efektem znaczącej nieszczelności otworu, aczkolwiek wskazują na potrzebę dalszego monitoringu. Przeprowadzone badania podkreślają znaczenie doboru odpowiedniej metodyki opróbowania strefy przyotworowej oraz zasadności wykonywania pełnej analizy składu cząsteczkowego dla wczesnego wykrywania nieszczelności otworów naftowych.

Słowa kluczowe: otwór naftowy, gaz glebowy, badania geochemiczne, metan, alkan

WELL INTEGRITY OF PETROLEUM BOREHOLES ASSESSED THROUGH SURFACE GEOCHEMICAL SURVEYS: A CASE STUDY FROM THE FORE-SUDETIC MONOCLINE

Abstract: The aim of this study is to assess the integrity of an oil well located on the Fore-Sudetic Monocline based on near-surface geochemical investigations of soil gases using the “free gas” method. Soil gas samples were collected from a depth of ~1.2 m at 16 points arranged around the wellhead and analyzed by gas chromatography (FID/TCD) for CH₄, total C₂–C₅ alkanes, alkenes, H₂, O₂, N₂ and CO₂. Methane concentrations (0.9–30 ppm) and ethane concentrations (0.012–3.657 ppm) exhibit asymmetric distributions, and the values of total C₂–C₅ alkanes above 0.2 ppm were defined as anomalous. Strong correlations between CH₄–C₂H₆–C₃H₈ ($r \approx 0.93$ – 0.96) indicate a common, deep-seated source of hydrocarbon emanations. Anomalous alkane zones were recorded in the immediate vicinity of the wellhead. Geochemical ratio analysis revealed similarity between soil gas composition and the well gas, confirming the presence of trace amounts of natural gas in the samples. Maximum concentrations (30 ppm CH₄ and 4.57 ppm total C₂–C₅ alkanes) do not indicate significant leakage of the well but point to the necessity of continued monitoring. The study highlights the importance of proper sampling strategy in the near-well zone and the need for comprehensive molecular composition analysis to enable early detection of potential well integrity failures.

Keywords: petroleum well, soil gas, geochemical research, methane, alkanes

https://doi.org/10.7494/978-83-68219-80-7_9

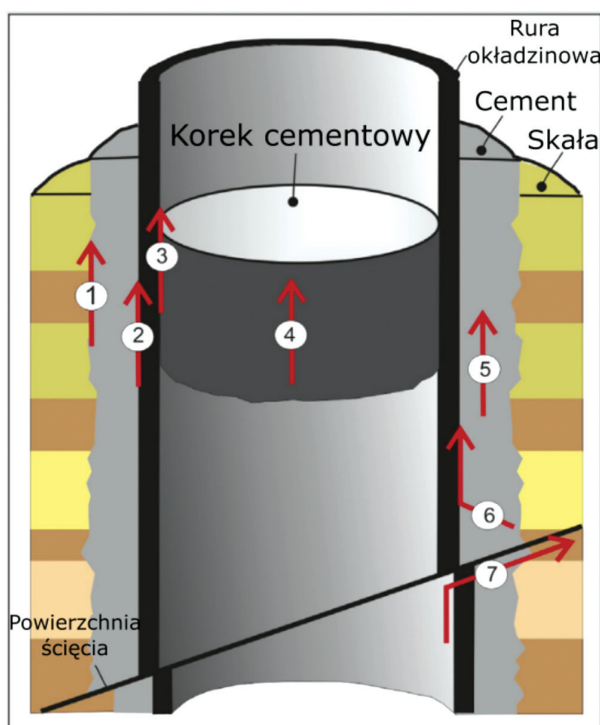
1. Wprowadzenie

Metan jest jednym z najgroźniejszych gazów dla atmosfery (Miller i in. 2013, IPCC 2021, EPA 2021). Według Międzyrządowego Zespołu ds. Zmian Klimatu (ang. IPCC) przez ostatnie 100 lat potencjał tworzenia efektu cieplarnianego spowodowanego emisją metanu jest 34 razy większy niż dwutlenku węgla. Oznacza to, że ilość ciepła zaabsorbowanego przez określoną masę metanu jest 34 razy większa niż ilość ciepła zaabsorbowana przez podobną masę dwutlenku węgla. Tak wysoka wartość tego współczynnika oznacza, że ograniczenie emisji metanu będzie miało kluczowe znaczenie dla zmniejszenia efektu cieplarnianego w najbliższym czasie (Caulton i in. 2014, Kang i in. 2019).

Światowy przemysł paliw kopalnych (produkcja i wykorzystanie gazu ziemnego, ropy i węgla) stanowi ok. 15–22% całkowitej emisji metanu (wg Schwietzke i in. 2016). Wśród znaczących źródeł metanu wymienia się nieszczelne otwory wiertnicze, zarówno te produkcyjne, jak i zlikwidowane. Od kilkunastu lat wycieki gazu ziemnego z otworów naftowych i gazowych są uznane za jedno z największych zagrożeń dla środowiska (Boothroyd i in. 2016, Lyman i in. 2017, Macuda i in. 2017, McMahon i in. 2018, Schout i in. 2019, EPA 2021, Sechman 2022). Zjawiska takie związane są z eksploatacją otworową zarówno konwencjonalnych, jak i niekonwencjonalnych złóż węglowodorów (np. Erno i Schmitz 1996, Darrah i in. 2014, Davies i in. 2014, Hammond 2016, Howarth 2019, Sechman 2022). Pomimo że stosowane obecnie techniki wiercenia są coraz lepsze, wykonanie otworu zawsze powoduje naruszenie ciągłości struktur geologicznych (Zawisza i Macuda 2006, Davies i in. 2014, Kremieniewski 2016). W konsekwencji prowadzi to do tworzenia się szczelin w strefach przyotworowych, które umożliwiają migrację gazu wzdłuż orurowania otworu ku powierzchni terenu (Yudhowijoyo i in. 2018). Davies wymienia siedem możliwych ścieżek migracji gazu ziemnego w strefie przyotworowej (Davies i in. 2014). Gaz może migrować (rys. 1): między cementem a otaczającymi go formacjami skalnymi (1); między rurami okładzinowymi a cementem (2); między korkiem cementowym a obudową (3); przez korek cementowy (4); przez mikrospeknięcia w cemencie wypełniającym przestrzeń pomiędzy rurami okładzinowymi a formacjami skalnymi (5); przez cement między obudową a formacją skalną (6); wzdłuż powierzchni ścięcia powstałego na skutek przemieszczenia się mas skalnych (7).

Ilość przeciekającego gazu jest uzależniona od wielu czynników, np. od ciśnienia gazu ziemnego w złożu oraz stopnia uszkodzenia strefy przyotworowej. Na temat przyczyn pojawiania się nieszczelności otworów pisało wielu badaczy (np. Davies i in. 2014, Kang i in. 2016, Forde i in. 2019). Relatywnie niewielkie przecieki gazu ziemnego wzdłuż kolumny rur otworu wiertniczego objawiają się w postaci podwyższonych stężeń głównie metanu w gazie glebowym. Jest to z reguły początkowy etap mogącej pojawić się w późniejszym okresie ekshalacji gazu. W przypadku stosunkowo dużej ilości przecieka-

jącego gazu na powierzchni terenu można makroskopowo zaobserwować efekty jego migracji (np. bąble gazu w kałużach). Poza tym gaz może również migrować lateralnie, nasycając przypowierzchniowe środowisko gruntowo-wodne (Sechman i in. 2013, Uliasz-Misiak i in. 2019, Płochocka 2021, Sechman 2022). W konsekwencji powoduje to pojawienie się anomalnych stężeń metanu i jego homologów nie tylko w bliskim sąsiedztwie nieszczelnego otworu. W przypadku przedostania się metanu do obiektów kubaturowych tworzy on z powietrzem mieszaninę, która w odpowiednich proporcjach może stanowić bezpośrednie zagrożenie wybuchem lub pożarem. W związku z tym niezwykle ważne jest określenie zasięgu anomalnych stężeń metanu w glebie pojawiających się w sąsiedztwie nieszczelnych otworów wiertniczych (Sechman 2022).



Rys. 1. Możliwe ścieżki migracji węglowodorów w obrębie otworu wiertniczego (Davies i in. 2014, zmodyfikowane – Płochocka 2021)

Przepisy dotyczące ochrony środowiska podczas prowadzenia robót wiertniczych w Polsce są regulowane przez *Rozporządzenie Ministra Gospodarki z dnia 25 kwietnia 2014 r. w sprawie szczegółowych wymagań dotyczących prowadzenia ruchu zakładów górniczych wydobywających kopaliny otworami wiertniczymi* (Rozporządzenie... 2014). Rozporządzenie nakazuje, aby przed rozpoczęciem prac poszukiwawczych i eksploatacyjnych złóż węglowodorów przeprowadzić między innymi pomiary gazowego tła w glebie.

Za minimalizowanie wpływu działalności górniczej na środowisko odpowiada kierownik ruchu zakładu otworowego. Wymagane są stałe obserwacje dotyczące zmian w tle gazowym. Pomiary pozwalają ocenić zagrożenia, ich zasięg i skalę, a wyniki nanosi się na mapy górnicze. W przypadku wycieków substancji szkodliwych, np. metanu, należy je natychmiast usuwać, a teren przywracać do stanu użyteczności (*Rozporządzenie...* 2014). Przepisy regulują również kwestie likwidacji otworu oraz późniejszego monitorowania stanu środowiska. Otwór musi być zlikwidowany w taki sposób, aby zapewnić szczelną izolację poziomów wodonośnych, roponośnych i gazonośnych. Sposób, w jaki otwór zostanie zlikwidowany, oraz użyte do tego materiały muszą być dostosowane do istniejących warunków geologicznych. Po procesie likwidacji należy zbadać tło gazowe i wyznaczyć strefę ochronną wokół otworu.

Z przedstawionych informacji wynika, że kluczowe jest przeprowadzenie badań pierwotnego tła gazowego w glebie oraz wykonanie analogicznych badań po zakończeniu wiercenia, a w przypadku likwidacji otworu wykonanie takich badań po zakończeniu prac likwidacyjnych i przeprowadzeniu rekultywacji terenu. Ustawodawca nie precyzuje jednak, jaką strefę należy objąć badaniami, jaka powinna być liczba punktów pomiarowych, a także nie wskazuje, jaką metodykę należy zastosować, aby uzyskać wiarygodne wyniki badań. O zakresie i sposobie wykonania takich badań decyduje kierownik ruchu zakładu górniczego, czyli kierownik wiertni lub kopalni gazu ziemnego/ropy naftowej.

Uwzględniając kilkudziesięcioletnie doświadczenie zespołu Katedry Surowców Energetycznych Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie (KSE AGH) w zakresie realizacji powierzchniowych badań geochemicznych, należy stwierdzić, że najlepsze efekty badania tła gazowego i badań szczelności otworów wiertniczych uzyskuje się, stosując tzw. wariant gazu wolnego. Generalnie badania te polegają na pobraniu odpowiedniej liczby próbek gazu glebowego usytuowanych w strefie badanego otworu, wykonaniu szczegółowej analizy ich składu cząsteczkowego, a następnie interpretacji uzyskanych wyników badań.

Celem publikacji jest scharakteryzowanie problematyki dotyczącej integralności naftowych otworów wiertniczych oraz przedstawienie przykładowych wyników badań zrealizowanych przez Zespół KSE AGH w zakresie oceny szczelności otworu zlokalizowanego na monoklinie przedsudeckiej.

2. Charakterystyka przyczyn niekontrolowanej migracji gazu ziemnego w strefach otworów wiertniczych

Migracja węglowodorów w przestrzeniach pomiędzyrurowych i pozarurowych w otworach wiertniczych oraz emisja metanu do atmosfery jest w dzisiejszych czasach coraz bardziej poważnym problemem. Ekshalacje gazu objawiają się na dwa sposoby. Gaz może bezpośrednio wypływać z przestrzeni międzyrurowych lub objawiać się

w postaci podwyższonych stężeń metanu w powietrzu glebowym w sąsiedztwie otworu (np. Herman 2008, Kremieniewski 2016). Węglowodory migrują pionowo ku powierzchni w wyniku procesów adwekcyjnych lub dyfuzyjnych. Część wyciekających lekkich węglowodorów może zostać utleniona wskutek procesów mikrobialnych zachodzących w glebie (np. Shennan 2006). Przyczyny wycieków węglowodorów z otworów wiertniczych i emisji metanu do atmosfery dzielą się na geologiczne i techniczno-technologiczne. Zazwyczaj nie da się jednoznacznie określić jednej głównej przyczyny migracji gazu ziemnego występującej w strefach otworów wiertniczych. Jest to bowiem wynik połączenia wielu czynników geologicznych i techniczno-technologicznych (Davies i in. 2014, Kremieniewski 2016, Yudhowijoyo i in. 2018).

Jednym z najważniejszych czynników decydujących o migracji gazu z otworów wiertniczych jest budowa geologiczna rozpatrywanego regionu. W tym zakresie dużą rolę odgrywają właściwości fizykochemiczne przewiercanych skał oraz ich cechy petrofizyczne, takie jak: porowatość, przepuszczalność, kawernistość i szczelinowatość. Inne czynniki geologiczne to między innymi: ciśnienie górotworu, rozkład gradientów ciśnień złożowych, porowych i szczelinowania, występowanie warstw przepuszczalnych i ich miąższość, rodzaj płynów złożowych, temperatura, warunki hydrogeologiczne rozpatrywanego obszaru, zaburzenia tektoniczne (Herman 2008, Kołodziejak 2010, Kremieniewski 2016).

W przypadku czynników techniczno-technologicznych do określenia uszkodzeń w otworze wiertniczym używane są dwa terminy: awaria integralności odwiertu (ang. *well integrity failure*) oraz uszkodzenie ścian odwiertu (ang. *well barrier failure*). Mogą one wystąpić podczas wiercenia, eksploatacji lub po zlikwidowaniu otworu wiertniczego. Termin „uszkodzenie bariery odwiertu” jest stosowany, kiedy awarii ulega jeden z elementów otworu, np. rura okładzinowa, lecz nie powoduje to wykrywalnego w środowisku wycieku gazu. Termin „awaria integralności odwiertu” jest używany w przypadku, gdy zawodzą wszystkie bariery cementowe i okładzinowe, tworząc węglowodorem ścieżkę migracji do otaczającego środowiska, między innymi do atmosfery (Davies i in. 2014, Płochocka 2021).

Istnieje wiele przyczyn utraty integralności odwiertu. W większości przypadków jest to nieprawidłowo przeprowadzony proces cementowania i rurowania otworu (Kremieniewski 2016). Kolumny okładzinowe są zazwyczaj nieszczelne w miejscu połączeń gwintowych lub w wyniku korozji kwasowej. Kolejnym czynnikiem jest kurczenie się cementu, przez co tworzą się w nim pęknięcia i kanały. Inną przyczyną jest słabe wiązanie cementu między ścianą odwiertu a kolumną okładzinową (Kremieniewski 2016, Boothroyd i in. 2016).

Wycieki węglowodorów przez nieszczelności otworów wiertniczych mogą być katastrofalne w skutkach, o czym świadczą eksplozje gazu ziemnego. Mogą być również tak niskie, że nie da się ich wykryć. Źródłami migrujących płynów złożowych mogą być

akumulacje węglowodorów, nieprodukcyjne formacje przepuszczalne, pokłady węgla oraz biogenne lub termogeniczne gazy z płytkich formacji skalnych. Awaryjne mogą wystąpić w każdym rodzaju naftowych otworów wiertniczych, niezależnie, czy jest on aktualnie wiercony, czy są już nim wydobywane węglowodory, czy też został on zlikwidowany (Davies i in. 2014).

3. Metodyka badań

3.1. Badania terenowe – pobór próbek gazu glebowego

W badaniach szczelności naftowych otworów wiertniczych najlepsze efekty uzyskujemy, stosując badania składu gazów glebowych z wykorzystaniem metody gazu wolnego (Sechman i in. 2013, Sechman 2022). W strefach zlikwidowanych i czynnych otworów wiertniczych korzystnym rozwiązaniem jest takie rozmieszczenie punktów pomiarowych, aby szczegółowo zbadać strefę przyotworową, a także objąć badaniami obszar w promieniu do kilkudziesięciu metrów od otworu (Sechman 2022). Zaleca się, aby badania wykonywać, stosując tzw. optymalne rozmieszczenie punktów pomiarowych przedstawione na rysunku 2a, szczególnie w przypadku lokalizacji otworów w obszarach o skomplikowanej budowie geologicznej (np. Karpaty zewnętrzne). Takie rozmieszczenie punktów zapewnia szczegółowe zbadanie zarówno strefy przyotworowej, jak i terenu przyległego w promieniu ok. 60 m od otworu (Sechman 2022). Jednak w wyjątkowych przypadkach liczba punktów pomiarowych może być mniejsza. Według autorów niniejszej publikacji minimalna liczba punktów pomiarowych wynosząca 16 gwarantuje uzyskanie wiarygodnych informacji w zakresie określenia tzw. tła gazowego. Dotyczy to głównie badanych otworów wiertniczych zlokalizowanych na obszarach platformowych. Zaleca się, aby punkty te były rozmieszczone promieniście wokół otworu – najlepiej tak jak na przedstawionym schemacie (rys. 2b).

W strefie badanego otworu – opisywanego w ramach niniejszej publikacji – zastosowano wariant z minimalnym rozmieszczeniem punktów pomiarowych (rys. 2b). Wynika to z faktu, że badany otwór znajduje się w zachodniej części Polski, na obszarze monokliny przedsudeckiej, charakteryzującej się obecnością utworów permu i triasu zapadających pod małym kątem ku NE, stanowiących platformową pokrywę utworów karbońskiego podłoża (Żelaźniewicz i Aleksandrowski 2008).

Próbki gazu glebowego pobrano za pomocą specjalnego zestawu i metodyki opracowanej przez pracowników KSE AGH (Dzieniewicz i Sechman 2002, Sechman 2022). Opatentowana metodologia polega na pobraniu gazu glebowego za pomocą specjalnej sondy geochemicznej i gazoszczelnej strzykawki. Próbki pobierano do specjalnych szklanych pojemników, które uprzednio były napełnione nasyconym roztworem czystego chemicznie chlorku sodu. W świetle badań eksperymentalnych i danych literaturo-

Rys. 2. Sposoby rozmieszczania geochemicznych punktów pomiarowych wokół badanego otworu:
a) optymalna liczba punktów; b) najmniejsza dopuszczalna liczba punktów

3.2. Badania laboratoryjne – analizy składu cząsteczkowego próbek gazów glebowego

Próbki gazu glebowego analizowano pod kątem obecności metanu i lekkich alkanów (etanu, propanu, i-butanu, n-butanu, neopentanu, i-pentanu, n-pentanu), alkenów gazowych (etylenu, propylenu, 1-butenu), wodoru, tlenu, azotu i dwutlenku węgla. Analizy przeprowadzono metodą chromatografii gazowej z wykorzystaniem detekcji płomieniowo-jonizacyjnej (FID) i cieplno-przewodnościowej (TCD).

Oznaczeń węglowodorów dokonano na aparacie GC 8160 firmy Fisons Instruments z dokładnością do 0,01 ppm, stosując następujące parametry pracy:

- kolumna metalowa o długości 1,3 m wypełniona aktywowanym tlenkiem aluminium (mesh 100/120),
- przepływ gazu nośnego – hel (60 ml/min),
- przepływy gazów roboczych: wodór (26 ml/min), powietrze (300 ml/min),
- programowana temperatura termostatu: 80°C – 3 min, narost: 80°C do 200°C (20°/min), 200°C – 3 min,
- temperatura detektora FID – 270°C,
- temperatura komory iniekcyjnej – 100°C,
- objętość analizowanej próbki – 2 ml.

Do oznaczenia helu, wodoru, tlenu, azotu i dwutlenku węgla wykorzystano aparat firmy Carlo Erba Instruments, stosując następujące parametry pracy:

- kolumna metalowa o długości 2,5 m wypełniona sitami molekularnymi 5A (dla helu i wodoru) i porowatym polimerem (dla dwutlenku węgla),
- przepływ gazu nośnego – argon (30 ml/min),
- temperatura termostatu – 65°C – stała,
- temperatura detektora TCD (baza – 60°C, włókna – 180°C),
- objętość analizowanej próbki – 2 ml (przy użyciu dziesięciodrożnego, automatycznego zaworu dozującego firmy VALCO).

Dozowanie prób zostało przeprowadzone ręcznie za pomocą gazoszczelnej strzykawki, przy czym przed każdą serią i po każdej serii prób (10–15) wykonane zostały nastroiki mieszanek wzorcowych w celu kalibracji oraz ewentualnej rekalkulacji analizatora.

Próg oznaczalności detektora FID dla węglowodorów wynosi średnio 0,01 ppm. Próg oznaczalności detektora TCD: dla helu – 10 ppm, dla wodoru – 10 ppm, dla tlenu, azotu i dwutlenku węgla – 0,01% obj.

4. Wyniki badań i ich analiza

4.1. Charakterystyka statystyczna zbiorów pomierzonych stężeń

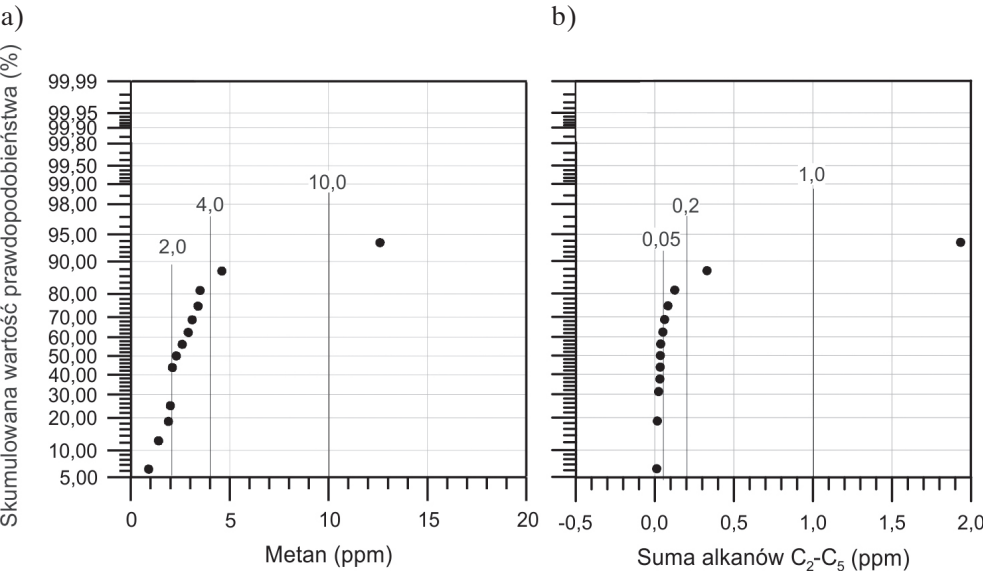
Spośród składników węglowodorowych oznaczonych w próbkach gazu glebowego dominuje metan, który został zarejestrowany w każdej z badanych próbek (tab. 1). Zbiór ten charakteryzuje się wartością średnią prawie dwukrotnie wyższą od mediany (tab. 2), co wskazuje na obecność wartości anomalnych. Potwierdza to skonstruowany wykres prawdopodobieństwa dla metanu (rys. 3a), na którym wartości stężenia powyżej 4 ppm należy uznać za anomalne. Nie są to oczywiście stężenia wysokie, a użycie terminu „anomalne” ma podkreślić ich wyraźnie odstający charakter w relacji do pozostałych wartości.

W każdej próbce gazu glebowego stwierdzono również obecność etanu powyżej progu oznaczalności detektora FID. Stężenia etanu wahają się w zakresie dwóch rzędów wielkości (tab. 1). Podobnie jak w przypadku metanu, zbiór wartości stężenia etanu charakteryzuje się wyraźnie wyższą średnią od mediany, a także relatywnie wysoką wartością odchylenia standardowego (tab. 2). Wskazuje to również na obecność wartości anomalnych. Stężenia pozostałych składników alkanowych (propan, i-butan i n-butan) powyżej progu oznaczalności detektora FID zarejestrowano odpowiednio w ponad 56%, prawie 19% i 25% pobranych próbek (tab. 1 i 2). W strefie badanego otworu nie stwierdzono obecności pentanów, a spośród składników alkenowych oznaczono śladową zawartość propylenu tylko w jednej próbce (tab. 1).

Tabela 1
Skład cząsteczkowy próbek gazu glebowego pobranych w strefie badanego otworu

Nr próbki	CH ₄ [ppm]	C ₂ H ₆ [ppm]	C ₃ H ₈ [ppm]	iC ₄ H ₁₀ [ppm]	nC ₄ H ₁₀ [ppm]	meoC ₃ H ₁₂ [ppm]	iC ₅ H ₁₂ [ppm]	nC ₅ H ₁₂ [ppm]	C ₂ H ₄ [ppm]	C ₃ H ₆ [ppm]	1-C ₄ H ₈ [ppm]	He [% obj.]	H ₂ [% obj.]	O ₂ [% obj.]	N ₂ [% obj.]	CO ₂ [% obj.]
1	30,0	3,66	0,73	0,07	0,11	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	17,6	77,0	1,82
2	3,4	0,25	0,05	0,01	0,02	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,04	p.g.o.	p.g.o.	p.g.o.	18,2	77,3	1,50
3	12,6	1,52	0,32	0,03	0,06	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	18,2	77,5	1,42
4	2,3	0,03	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	18,2	77,5	1,42
5	2,1	0,02	0,01	p.g.o.	0,01	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	18,1	77,7	1,55
6	4,6	0,11	0,01	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,004	17,9	77,8	1,48
7	3,1	0,04	0,01	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	18,1	77,7	1,55
8	0,9	0,02	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,027	18,1	78,9	0,54
9	1,4	0,02	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,002	17,8	78,1	1,68
10	2,1	0,02	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	17,2	78,2	2,60
11	2,1	0,02	0,01	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,005	19,4	77,9	0,61
12	3,5	0,07	0,02	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	18,3	77,6	1,85
13	1,9	0,02	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,050	18,0	78,2	1,12
14	2,6	0,03	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	16,7	78,2	2,88
15	2,0	0,01	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,007	18,1	77,9	1,54
16	2,9	0,05	0,01	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	p.g.o.	0,010	19,6	77,3	0,41

p.g.o. – poniżej granicy oznaczalności



Rys. 3. Wykresy prawdopodobieństwa dla zbioru stężeń metanu (a) i sumy alkanów C₂-C₅ (b)

Tabela 2
Podstawowe parametry statystyczne składników alkanowych pomierzonych w próbkach gazu glebowego

Parametr statystyczny	Jednostka	CH ₄	C ₂ H ₆	C ₃ H ₈	<i>i</i> -C ₄ H ₁₀	<i>n</i> -C ₄ H ₁₀	Suma alkanów C ₂ -C ₅
Wartość minimalna	[ppm]	0,9	0,012	<i>p.g.o.</i>	<i>p.g.o.</i>	<i>p.g.o.</i>	0,012
Wartość maksymalna	[ppm]	30,0	3,657	0,733	0,065	0,111	4,566
Średnia	[ppm]	4,8	0,369	0,073	0,006	0,013	0,461
Mediana	[ppm]	2,5	0,034	0,009	0	0	0,036
Odch. standardowe	[ppm]	7,2	0,953	0,193	0,017	0,031	1,192
Udział proc. próbek*	[%]	100,0	100,0	56,3	18,8	25,0	100,0
* – udział procentowy próbek z wartością stężenia powyżej progu oznaczalności detektora FID, <i>p.g.o.</i> – poniżej granicy oznaczalności							

Metan może mieć pochodzenie poligenetyczne, dlatego w interpretacji wyników powierzchniowych badań geochemicznych jako wiodący wykorzystuje się wskaźnik sumy alkanów C₂-C₅. Wynika to z faktu, że alkany cięższe od metanu mają pochodzenie wgłębne i nie generują się w znaczących ilościach w procesach mikrobialnych (Tedesco 1995, Jones i in. 2000, Sechman 2012). Zatem podwyższona koncentracja

sumy alkanów C_2-C_5 w gazie glebowym może wskazywać na ich mikroprzeziąkanie i migrację z głębi lub z nieuszczelnego otworu. Parametry statystyczne charakteryzujące zbiór wartości stężenia sumy alkanów C_2-C_5 wskazują na obecność wartości anomalnych w jego obrębie (tab. 2), co znajduje potwierdzenie w skonstruowanym wykresie prawdopodobieństwa (rys. 3b). Stężenie sumy alkanów C_2-C_5 przekraczające 0,2 ppm można uznać za anomalne.

Spośród składników niewęglowodorowych uwagę zwracają podobne wartości średniej i mediany dla zbioru stężenia dwutlenku węgla oraz identyczne w przypadku zbioru stężenia tlenu i azotu (tab. 3). Oznacza to, że zbiory te cechują się rozkładem zbliżonym do normalnego, a w ich obrębie nie występują wartości odstające, które można uznać za anomalne.

Tabela 3

Podstawowe parametry statystyczne zbiorów stężeń dwutlenku węgla, wodoru, tlenu i azotu pomierzonych w próbkach gazu glebowego

Parametr statystyczny	CO ₂ [% obj.]	H ₂ [% obj.]	O ₂ [% obj.]	N ₂ [% obj.]
Wartość minimalna	0,41	<i>p.g.o.</i>	16,7	77,0
Wartość maksymalna	2,88	0,0497	19,6	78,9
Średnia	1,50	0,0066	18,1	77,8
Mediana	1,52	0	18,1	77,8
Odch. standardowe	0,66	0,0135	0,7	0,5
Udział proc. próbek*	100,0	43,8	100,0	100,0
* – udział procentowy próbek z wartością stężenia powyżej progu oznaczalności detektora FID, <i>p.g.o.</i> – poniżej granicy oznaczalności				

4.2. Ocena relacji i zależności pomiędzy analizowanymi zbiorami stężeń

Szczegółowy skład cząsteczkowy próbek gazu glebowego pobranych w strefie badanego otworu pozwala na zbadanie relacji i zależności pomiędzy uzyskanymi zbiorami stężeń. Analiza taka umożliwia wstępną ocenę genetyczną zarejestrowanych stężeń metanu, a także interpretację w zakresie określenia procesów biochemicznych zachodzących w strefie badanego otworu, które mogą być związane z obecnością podwyższonych stężeń węglowodorów i innych składników.

W przypadku badanego otworu bardzo wysokie wartości współczynnika korelacji liniowej (r) występują tylko pomiędzy wskaźnikami węglowodorowymi (tab. 4, rys. 4a, b). Należy zaznaczyć, że metan koreluje się bardzo dobrze z etanem i propanem. Wyższe od metanu alkany rejestrowane w strefie przypowierzchniowej mają głównie pochodzenie

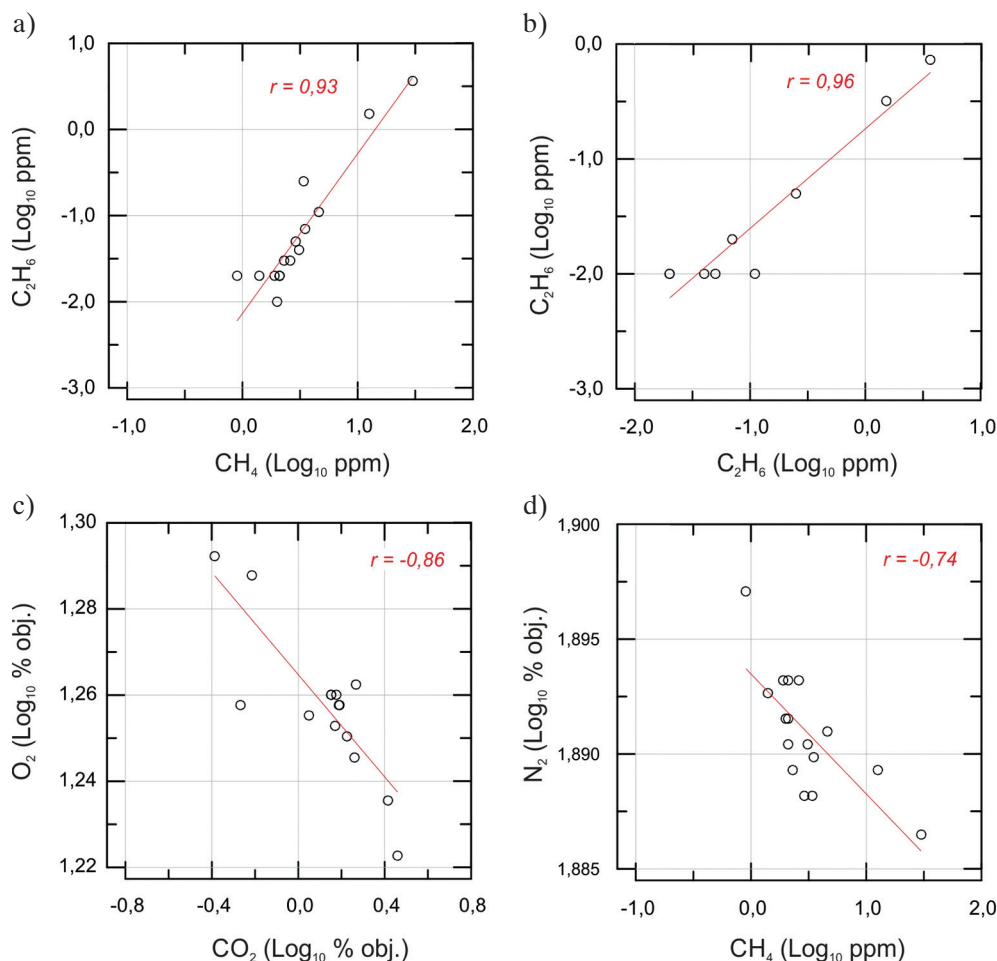
wgłębne (np. Tedesco 1995, Saunders i in. 1999, Sechman 2012), co oznacza wgłębne pochodzenie metanu. Z kolei bardzo wysoka wartość współczynnika korelacji pomiędzy etanem a propanem wskazuje na to, że alkany wyższe od metanu zarejestrowane w strefie przypowierzchniowej pochodzą generalnie z jednego pod względem charakteru źródła rozpraszania. Tego typu relacje wykorzystywano w interpretacji wyników badań geochemicznych wykonanych w wielu obszarach ropogazonośnych (Jones i Drozd 1983, Jones i in. 2000, Harbert i in. 2006, Sechman 2012, Sechman i in. 2016, 2025, Twarog i in. 2024).

Tabela 4

Współczynniki korelacji liniowej Pearsona pomiędzy zlogarytmowanymi wartościami zbiorów analizowanych stężeń

	Log(CH ₄)	Log(C ₂ H ₆)	Log(C ₃ H ₈)	Log(H ₂)	Log(O ₂)	Log(N ₂)	Log(CO ₂)
Log(CH ₄)	1						
Log(C ₂ H ₆)	0,93	1					
Log(C ₃ H ₈)	0,93	0,96	1				
Log(H ₂)	−0,30	−0,19	brak	1			
Log(O ₂)	−0,09	−0,06	−0,48	0,01	1		
Log(N ₂)	−0,74	−0,66	−0,71	0,40	−0,30	1	
Log(CO ₂)	0,26	0,15	0,40	−0,41	−0,83	−0,09	1

Zbiory stężenia tlenu i dwutlenku węgla posiadają przeciwną korelację wyrażoną wartością współczynnika korelacji liniowej równą −0,83 (tab. 4, rys. 4c). Oznacza to, że tlen jest zużywany, a dwutlenek węgla wytwarzany w procesie oddychania glebowego. Tlen jest niezbędny dla korzeni roślin i organizmów glebowych, natomiast CO₂ jest wytwarzany w procesie ich metabolizmu (oddychania) (Waleńczak 1987). Względne stężenia O₂ i CO₂ są odwrotnie proporcjonalne – w glebach słabo napowietrzonych występuje niższy poziom tlenu, a wyższy poziom dwutlenku węgla (Le Mer i Roger 2001, Shahzad i in. 2019). Takie relacje pomiędzy tlenem a dwutlenkiem węgla mogą również wskazywać na procesy mikrobialnego utleniania materii organicznej i/lub węglowodorów migrujących z głębi do powierzchni terenu (Le Mer i Roger 2001, Shennan 2006). Jednak obecność tego ostatniego procesu byłaby potwierdzona odwrotną zależnością pomiędzy metanem a dwutlenkiem węgla, a wyniki naszych badań nie wykazały takich relacji (tab. 4). Niewielki stopień utleniania metanu i wyższych alkanów potwierdza również odwrotna zależność pomiędzy azotem a metanem i wyższymi alkanami. Wykazano bowiem, że azot glebowy wpływa na wielkość stężenia metanu przez potencjalne hamowanie jego bakteryjnego utleniania (Bodelier i Laanbroek 2004).



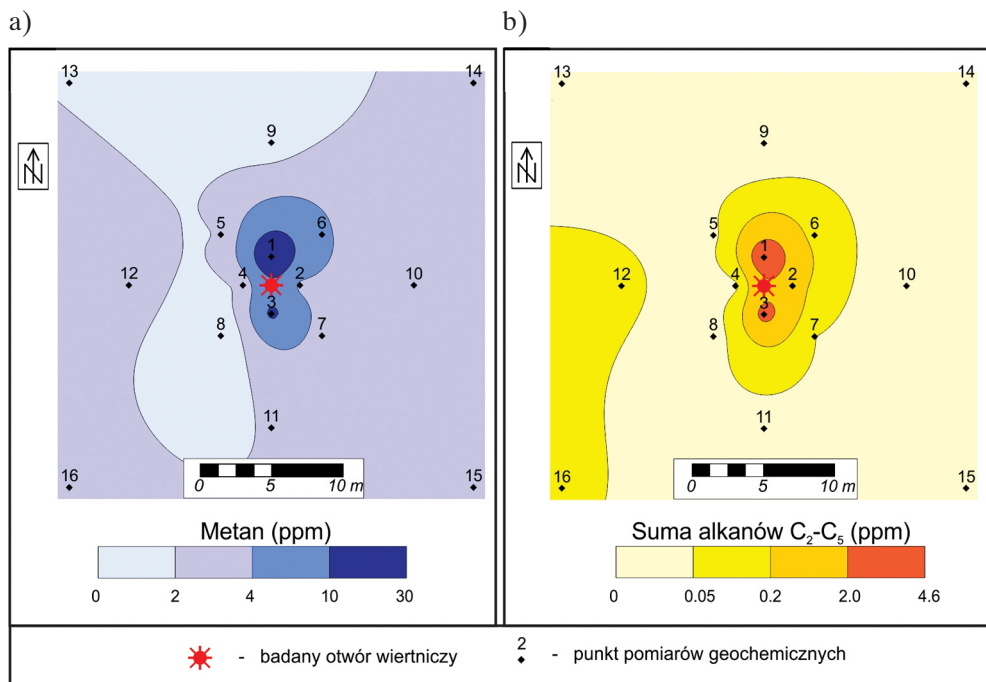
Rys. 4. Wykresy korelacyjne dla zlogarytmowanych zbiorów stężeń:
a) metan – etan; b) etan – propan; c) dwutlenek węgla – tlen; d) metan – azot

4.3. Charakterystyka powierzchniowych rozkładów stężeń metanu i wyższych alkanów w strefie badanego otworu

Stężenia metanu i sumy alkanów C_2 – C_5 pomierzone w próbkach gazu glebowego pobranych wokół otworu O-1 przedstawiono graficznie w formie rozkładów powierzchniowych (rys. 5a, b). Cięcia warstwowe przyjęto na podstawie skonstruowanych rozkładów prawdopodobieństwa (rys. 3a, b). W celu uzyskania map rozkładów stężeń wykonano interpolację danych metodą krigingu z wykorzystaniem programu Surfer.

Prezentowane rozkłady stężeń metanu i sumy alkanów C_2 – C_5 są podobne pod względem zasięgu wartości anomalnych (rys. 5a, b). Występują one w próbkach pobranych w bezpośrednim sąsiedztwie głowicy badanego otworu. Relatywnie znaczący

udział cięższych składników alkanowych w stosunku do metanu we wspomnianych próbkach wynika z łatwej i szybkiej migracji węglowodorów do powierzchni. Aktywną migrację alkanów potwierdza znikomy udział składników alkenowych (Saunders i in. 1999, Jones i in. 2000, Harbert i in. 2006, Sechman 2012, Sechman i in. 2015, 2016, Twaróg i in. 2024).

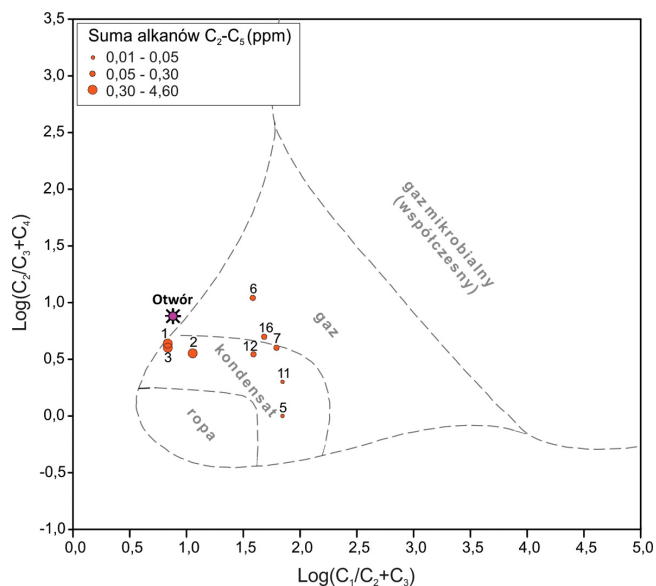


Rys. 5. Rozkłady powierzchniowe pomierzonych stężeń metanu (a) i sumy alkanów C_2-C_5 (b)

Współwystępowanie anomalii metanowych i sumy wyższych alkanów na mapach potwierdza głębne pochodzenie metanu (Jones i Drozd 1983, Jones i in. 2000, Sechman 2012). Należy jednak zaznaczyć, że poziom zarejestrowanych stężeń metanu jest stosunkowo niski (tab. 1 i 2) i nie świadczy o nieszczelności badanego otworu. Obecność wspomnianych anomalii węglowodorowych w sąsiedztwie otworu jest konsekwencją naruszenia ciągłości warstw geologicznych. Śladowe ilości węglowodorów mogą przemieszczać się, wykorzystując spękania i mikrospeknięcia powstałe wzdłuż trajektorii otworu (rys. 1), opisane szczegółowo w literaturze (Davies i in. 2014).

Ocena genetyczna zarejestrowanych stężeń metanu opiera się zwykle na wynikach badań izotopowych. Są to badania pozwalające na stwierdzenie, czy w próbkach gazu głębokiego jest obecny metan pochodzący z otworu. Jednak w opisywanym przypadku relatywnie niski poziom zarejestrowanych stężeń tego gazu – maksymalnie osiągający wartość 30 ppm (tab. 1) powoduje, że błąd oznaczenia składu izotopowego mógłby być

dość znaczny. Zatem identyfikacja genetyczna metanu w próbkach gazu glebowego byłaby niejednoznaczna. W związku z tym posłużono się analizą oceny relacji pomiędzy wartościami współczynników węglowodorowych przy wykorzystaniu tzw. wykresu Jonesa (Jones i in. 2000, Sechman i in. 2012). Wykres przedstawia zależność między zlogarytmowanymi wartościami współczynników $C_1/(C_2 + C_3)$ oraz $C_2/(C_3 + C_4)$, obliczonymi na podstawie zmierzonych stężeń składników alkanowych (rys. 6). Interpretując położenie punktów na wykresie, można oszacować, z jakiego typu akumulacji – gazu, ropy lub kondensatu – pochodzą śladowe stężenia alkanów zarejestrowane w próbkach gazu glebowego. Umieszczenie na tym samym wykresie punktu wyznaczonego na podstawie składu cząsteczkowego gazu z otworu (tab. 5) pozwala, przez analizę wzajemnego położenia punktów, ocenić, czy składniki zarejestrowane w gazie glebowym pochodzą z tego otworu.



Rys. 6. Wykres Jonesa dla próbek gazu glebowego oraz punktu reprezentującego skład cząsteczkowy gazu ziemnego pochodzącego z badanego otworu

Wyniki badań wykazały, że punkty 1, 2, 3 o najwyższych stężeniach składników alkanowych (tab. 1) położone są na wykresie Jonesa (rys. 6) w sąsiedztwie punktu reprezentującego skład cząsteczkowy gazu pochodzącego z badanego otworu (tab. 5). Wartości współczynnika $C_1/(C_2 + C_3)$ dla punktów pomiarowych 1, 2 i 3 są bardzo podobne do wartości tego współczynnika obliczonego na podstawie stężeń węglowodorów pomierzone w próbce gazu ziemnego pochodzącego z otworu (tab. 5). Z kolei nieco niższe wartości współczynnika $C_2/(C_3 + C_4)$ dla punktów reprezentujących anomalne stężenia alkanów w gazie glebowym w relacji do tego współczynnika obliczonego

dla gazu z otworu może być konsekwencją procesu aktywniejszego uwalniania się z gleby pewnej części lżejszego etanu w porównaniu z wyższymi alkanami (propan, i-butan i n-butan). Zatem można stwierdzić, że w punktach 1, 2 i 3 (tab. 1, rys. 6) zarejestrowano obecność gazu ziemnego pochodzącego z badanego otworu wiertniczego. Pozostałe punkty na wykresie zależności współczynników węglowodorowych (rys. 6) wskazują na naturalne mikropenetracje składników węglowodorowych z akumulacji wgłębnej przez nadkład skalny do powierzchni terenu.

Tabela 5

Skład cząsteczkowy próbki gazu ziemnego pobranego z badanego otworu O-1

Składnik	Stężenie [% obj.]
CH ₄	49,8
C ₂ H ₆	6,02
C ₃ H ₈	0,54
<i>i</i> C ₄ H ₁₀	0,22
<i>n</i> C ₄ H ₁₀	0,32
<i>neo</i> C ₅ H ₁₂	0,003
<i>i</i> C ₅ H ₁₂	0,163
<i>n</i> C ₅ H ₁₂	0,124
C ₂ H ₄	0,0550
C ₃ H ₆	0,0003
1-C ₄ H ₈	0,0001
He	0,09
H ₂	3,06
O ₂	0,7
N ₂	37,6
CO ₂	1,29

5. Podsumowanie i wnioski

W artykule przedstawiono zagadnienia dotyczące problematyki związanej ze szczelnością naftowych otworów wiertniczych. Szczegółowo przeanalizowano wyniki powierzchniowych badań geochemicznych uzyskane w strefie jednego z otworów naftowych położonego na obszarze monokliny przedsudeckiej.

Wykazano, że zastosowanie w badaniach terenowych metody „gazu wolnego” z wykorzystaniem opatentowanego zestawu pomiarowego oraz szczegółowa analiza składu cząsteczkowego pobranych próbek gazu glebowego umożliwiają uzyskanie wyników pozwalających na skuteczną ocenę szczelności otworu już na etapie pojawienia się śladowych ilości gazu ziemnego w strefie przyotworowej.

Liczba punktów pomiarowych i sposób ich rozmieszczenia mają duże znaczenie dla wiarygodnej oceny szczelności naftowych otworów wiertniczych. W tym zakresie autorzy preferują rozmieszczenie 40 punktów pomiarowych – dotyczy to szczególnie obszarów o skomplikowanej budowie geologicznej. W przypadku otworów zlokalizowanych w obszarach platformowych zastosowanie 16 punktów pomiarowych rozmieszczonych w promieniu około 20 m zapewnia uzyskanie wiarygodnych wyników, co potwierdziły wyniki badań zaprezentowane w niniejszej publikacji.

Maksymalne wartości stężenia metanu i sumy alkanów C_2 – C_5 zanotowane w strefie badanego otworu wynoszące odpowiednio: 30 ppm i 4,57 ppm nie wskazują na ewidentną nieszczelność badanego otworu, aczkolwiek w jego sąsiedztwie występują procesy aktywnego mikrop przenikania gazu ziemnego spowodowane najprawdopodobniej mikrospekaniem strefy przyotworowej. Bardzo wysoka korelacja metanu z wyższymi alkanami wskazuje na jego wgłębne pochodzenie, a śladowe ilości alkenów potwierdzają aktywny proces mikrop przenikania alkanów do powierzchni terenu.

Analiza wykresu krzyżowego dla współczynników $C_1/(C_2 + C_3)$ vs. $C_2/(C_3 + C_4)$ obliczonych w oparciu o stężenia pomierzone w próbkach gazu glebowego i próbki gazu ziemnego pobranego z otworu pozwoliła na wstępną ocenę genetyczną zarejestrowanych stężeń metanu w próbkach gazu glebowego. Zidentyfikowano w nich obecność gazu ziemnego z otworu. Tego typu analiza jest szczególnie istotna w przypadku niskich stężeń metanu (rzędu ppm), gdy wyniki ewentualnych badań izotopowych mogą być obciążone relatywnie dość dużym błędem.

Uzyskane wyniki badań mają szczególne znaczenie dla ukierunkowania dalszych działań badawczych i technicznych podejmowanych przez operatora. W czasie prowadzenia pomiarów nie stwierdzono przesłanek świadczących o istotnej nieszczelności badanego otworu wiertniczego. Jednak wykryta śladowa obecność gazu ziemnego pochodzącego z otworu w kilku próbkach gazu glebowego w bezpośrednim jego sąsiedztwie wskazuje na symptomy mogącej się pojawić w przyszłości utraty integralności badanego otworu. Zatem zaleca się dalszy monitoring geochemiczny otworu w celu wczesnego wykrycia ewentualnych nieszczelności.

Badania zostały sfinansowane z funduszu statutowego (grant nr 16.16.140.315) Wydziału Geologii, Geofizyki i Ochrony Środowiska, Akademii Górniczo-Hutniczej w Krakowie.

Literatura

- Bodelier P.L.E., Laanbroek H.J., 2004, *Nitrogen as a regulatory factor of methane oxidation in soils and sediments*, FEMS Microbiology Ecology, vol. 47, s. 265–277.
- Boothroyd I.M., Almond S., Qassim S.M., Worrall F., Davies R.J., 2016, *Fugitive emissions of methane from abandoned, decommissioned oil and gas wells*, Science of the Total Environment, vol. 547, s. 461–469.
- Caulton D.R., Shepson P.B., Santoro R.L., Sparks J.P., Howarth R.W., Ingraffea A.R., Cambaliza M.O.L., Sweeney C., Karion A., Davis K.J., 2014, *Experimental investigation of wellbore integrity and CO₂-brine flow along the casing-cement microannulus*, Proceedings of the National Academy of Sciences, art. 201316546.
- Darrah T.H., Vengosh A., Jackson R.B., Warner N.R., Poreda R.J., 2014, *Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales*, Proceedings of the National Academy of Sciences of the United States of America, vol. 111, s. 14076–14081.
- Davies R.J., Almond S., Ward R.S., Jackson R.B., Adams C., Worrall F., Herringshaw L.G., Gluyas J.G., Whitehead M.A., 2014, *Oil and gas wells and their integrity: implications for shale and unconventional resource exploitation*, Marine and Petroleum Geology, vol. 56, s. 239–254.
- Rozporządzenie Ministra Gospodarki z dnia 25 kwietnia 2014 r. w sprawie szczegółowych wymagań dotyczących prowadzenia ruchu zakładów górniczych wydobywających kopaliny otworami wiertniczymi, Dz.U. 2014 poz. 812.
- Dzieniewicz M., Kuśmierk J., Rusta T., 1985, *Optimization of soil-gas field collection techniques based on experimental studies*, Association Petroleum Geochemical Explorationists Bulletin, vol. 1, s. 43–56.
- Dzieniewicz M., Sechman H., 2002, *Zestaw do ręcznego pobierania próbek gazowych z warstw przypowierzchniowych*, Patent RP nr PL 184080, WUP, 08/2002.
- EPA, 2021, U.S. Environmental Protection Agency. *Climate Change Indicators: Atmospheric Concentrations of Greenhouse Gases*. <https://www.epa.gov/climate-indicators/climate-change-indicators-atmospheric-concentrations-greenhouse-gases> [dostęp: 4.02.2025].
- Erno B., Schmitz R., 1996, *Measurements of soil gas migration around oil and gas wells in the Lloydminster area*, The Journal of Canadian Petroleum Technology, vol. 35, s. 37–46. <https://doi.org/10.3997/2214-4609.201408077>.
- Forde O.N., Mayer K.U., Hunkeler D., 2019, *Identification, spatial extent and distribution of fugitive gas migration on the well pad scale*, Science of the Total Environment, vol. 652, s. 356–366.
- Hammond P.A., 2016, *The relationship between methane migration and shale-gas well operations near Dimock. Pennsylvania, USA*. Hydrogeology Journal, vol. 24, s. 503–519.

- Harbert W., Jones V.T., Izzo J., Anderson T.H., 2006, *Analysis of light hydrocarbons in soil gases, Lost River region, West Virginia: relation to stratigraphy and geological structures*. Association Petroleum Geochemical Explorationists Bulletin, vol. 90, nr 5, s. 715–734.
- Herman Z., 2008, *Migracje i ekshalacje gazu ziemnego z przestrzeni międzyrurowych i pozarurowych odwiertów eksploatacyjnych na obszarze Przedgórza Karpat*, *Warsztaty górnicze 2008 z cyklu: Zagrożenia naturalne w górnictwie*, Baranów Sandomierski, 18–20 czerwiec 2008, Instytut Gospodarki Surowcami Mineralnymi i Energią PAN, 2008, s. 63–71.
- Howarth R.W., 2019, *Ideas and perspectives: is shale gas a major driver of recent increase in global atmospheric methane?*, *Biogeosciences*, vol. 16, s. 3033–3046.
- IPCC, 2021, *Intergovernmental Panel on Climate Change*. <https://www.ipcc.ch/> [dostęp: 20.05.2025].
- Jones V.T., Drozd R.J., 1983, *Prediction of oil or gas potential by near-surface geochemistry*, *American Association of Petroleum Geologist Bulletin*, vol. 67, s. 932–952.
- Jones V.T., Matthews M.D., Richers D.M., 2000, *Light hydrocarbons for petroleum and gas prospecting*. W: Hale M. (Ed.), *Handbook of Exploration Geochemistry*, vol. 7, Elsevier Science Publishers, s. 133–212.
- Kang M., Christian S., Celia M. A., Mauzerall D. L., Bill M., Miller A. R., Chen Y., Conrad M.E., Darrah T.H., Jackson R.B., 2016, *Identification and characterization of high methane-emitting abandoned oil and gas wells*, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113 (48), s. 13636–13641.
- Kołodziejak G., 2010, *Ocena możliwości minimalizacji zagrożeń powodowanych ekshalacjami gazu ziemnego na terenie miasta Przeworsk*, *Nafta – Gaz*, vol. 7, s. 591–596.
- Kremieniewski M., 2016, *Ograniczenie ekshalacji gazu w otworach wiertniczych poprzez modyfikację receptur oraz kształtowanie się struktury stwardniałych zaczynów uszczelniających*, Instytut Nafty i Gazu – Państwowy Instytut Badawczy, Kraków. https://inig.pl/magazyn/prace_naukowe/199.pdf [dostęp: 6.10.2025].
- Le Mer J., Roger P., 2001, *Production, oxidation, emission and consumption of methane by soils: A review*, *European Journal of Soil Biology*, vol. 37, s. 25–50.
- Lyman S.N., Watkins C., Jones C.P., Mansfield M.L., McKinley M., Kenney D., Evans J., 2017, *Hydrocarbon and Carbon Dioxide Fluxes from Natural Gas Well Pad Soils and Surrounding Soils in Eastern Utah*. *Environmental Science & Technology*, vol. 51, s. 11625–11633.
- Macuda J., Bogacki M., Siemek J., 2017, *Effect of drilling for shale gas on the quality of atmospheric air*. *Problemy Ekorozwoju. Polska Akademia Nauk. Komitet Człowiek i Środowisko*, vol. 12, no. 1, s. 91–100.

- McMahon P.B., Thomas J.C., Crawford J.T., Dornblaser M.M., Hunt A.G., 2018, *Methane in groundwater from a leaking gas well, Piceance Basin, Colorado, USA*, Science of the Total Environment, vol. 634, s. 791–801.
- Miller S.M., Wofsy S.C., Michalak A.M., Kort E.A., Andrews A.E., Biraud S.C., Dlugokencky E.J., Eluszkiewicz J., Fischer M.L., Janssens-Maenhout G., Miller B.R., Miller J.B., Montzka S.A., Nehrkorn T., Sweeney C., 2013, *Anthropogenic emissions of methane in the United States*, Earth, Atmospheric, and Planetary Sciences, vol. 110, s. 20018–20022. <https://doi.org/10.1073/pnas.1314392110>.
- Płochocka S., 2021, *Naftowe otwory wiertnicze jako źródło emisji metanu do atmosfery*. Projekt inżynierski, AGH, Archiwum KSE AGH [praca niepublikowana].
- Saunders D.F., Burson K.R., Thompson C.K., 1999, *Model for hydrocarbon microseepage and related near-surface alterations*, AAPG Bulletin Am. Assoc. Petrol. Geolog. Bull., vol. 83, s. 170–184.
- Shahzad K., Bary A.I., Collins D.P., Chalker-Scott L., Abid M., Sintim H.Y., Flury M., 2019, *Carbon dioxide and oxygen exchange at the soil-atmosphere boundary as affected by various mulch materials*, Soil and Tillage Research, vol. 194, no. 11, art. no. 104335.
- Shennan J.L., 2006. *Review. Utilisation of C₂–C₄ gaseous hydrocarbons and isoprene by microorganisms*, Journal of Chemical Technology and Biotechnology, vol. 81, s. 237–256.
- Schout G., Griffioen J., Hassanizadeh S.M., Cardon de Lichtbuer G., Hartog N., 2019, *Occurrence and fate of methane leakage from cut and buried abandoned gas wells in the Netherlands*. Science of the Total Environment, vol. 659, s. 773–782.
- Schwietzke S., Sherwood O.A., Bruhwiler L.M.P., Miller J.B., Etiope G., Dlugokencky E. J., Michel, S.E., Arling V.A., Vaughn B. H., White J.W.C., and Tans P.P., 2016, *Upward revision of global fossil fuel methane emissions based on isotope database*, Nature, vol. 538, s. 88–91. <https://doi.org/10.1038/nature19797>.
- Sechman H., 2006, *Głębokość poboru próbki gazu podglebowego w powierzchniowych badaniach geochemicznych – próba optymalizacji*, Geologia, vol. 32, nr 2, s. 117–139.
- Sechman H., Mościcki J. W., Dzieńiewicz M., 2013, *Pollution of near-surface zone in the vicinity of gas wells*, Geoderma, vol. 197–198, s. 193–204.
- Sechman H., Izydor G., Guzy P., Dzieńiewicz M., 2015. *Surface geochemical exploration for hydrocarbons in the area of prospective structures of the Lublin Trough (Eastern Poland)*, Marine and Petroleum Geology, vol. 61, s. 22–38.
- Sechman H., Kuśmierk J., Machowski G., Guzy P., Dzieńiewicz M., 2016, *Surface geochemical anomalies in the vicinity of the Wańkowa oil field (SE Polish Carpathians)*, Annales Societatis Geologorum Poloniae, vol. 86, iss. 2, s. 219–235.
- Sechman H., 2022, *Detailed analysis of gaseous components in soil gases around petroleum wells – an effective tool for evaluation of their integrity*, Applied Geochemistry, vol. 142, art. no. 105346, s. 1–13.

- Tedesco S.A., 1995, *Surface Geochemistry in Petroleum Exploration*, Chapman & Hall Int. Thomson Publ. Co., New York.
- Twaróg A., Starzec K., Sechman H., Schnabel W. 2024, *Surface geochemistry as a key to understanding the petroleum system in regions of complex geology – Polish Outer Carpathians*, Marine and Petroleum Geology, vol. 170, art. no. 107155.
- Uliasz-Misiak B., Lewandowska-Śmierzchalska J., Matuła R., 2019, *Emisja ditlenku węgla i metanu związana z poszukiwaniem i wydobywaniem węglowodorów w Polsce*, Przemysł Chemiczny, t. 88, nr 5, s. 813–816.
- Waleńczak Z., 1987, *Geochemia organiczna*, Wydawnictwa Geologiczne, Warszawa.
- Yudhowijoyo A., Rafati R., Haddad A.S., Raja M.S., Hamidi H., 2018, *Subsurface methane leakage in unconventional shale gas reservoirs: A review of leakage pathways and current sealing techniques*, Journal of Natural Gas Science and Engineering, vol. 54, s. 309–319.
- Zawisza L., Macuda J., 2006, *Ocena szczelności struktur geologicznych przeznaczonych do składowania odpadów płynnych w górotworze*, Wiertnictwo, Nafta, Gaz, t. 23, nr 1, s. 619–626.
- Żelaźniewicz A., Aleksandrowski P., 2008, *Regionalizacja tektoniczna Polski – Polska południowo-zachodnia*, Przegląd Geologiczny, t. 56, nr 10, s. 904–911.

Polskie środowiska emigracyjne – różne strategie działania w obliczu klęski idei suwerennej Polski w 1945 roku

Anna Siwik

Akademia Górniczo-Hutnicza, Wydział Humanistyczny, Kraków

Streszczenie: Artykuł analizuje strategie działania polskich środowisk emigracyjnych po klęsce idei suwerennej Polski w 1945 roku, w kontekście dominacji ZSRR i USA w powojennym porządku świata. Opierając się na rozróżnieniu realizmu politycznego (Realpolitik) i idealizmu, autorka omawia perspektywę interesów mocarstw (Stalina, Churchilla i Roosevelta), które marginalizowały Rząd RP w Londynie. Wyróżnia trzy orientacje emigracyjne: „niezłomnych” (odrzucających Jałtę, oczekujących konfliktu Wschód-Zachód, np. Arciszewski, Anders); „umiarkowanych realistów” (akceptujących kompromis, np. powracający do kraju Mikołajczyk) oraz „aktywnych realistów” (godzących się z zależnością od ZSRR). Artykuł dowodzi, że historyczna rola emigracji wykraczała poza doraźną politykę, przejawiając się w trzech kluczowych płaszczyznach. W sferze politycznej był to wysiłek na rzecz symbolicznego podtrzymania ciągłości państwowości przez instytucje „państwa na wygnaniu”. W wymiarze kulturowym – fundamentalny wkład w zachowanie i rozwój niezależnej myśli oraz dorobku narodowego m.in. poprzez takie ośrodki jak paryska „Kultura” i londyńskie „Wiadomości”. Wreszcie, w sferze materialnej – wsparcie dla opozycji demokratycznej w kraju. W konkluzji autorka stwierdza, że choć bezpośrednie cele polityczne z 1945 roku były nieosiągalne, to długofalowy dorobek emigracji, zwłaszcza w sferze kultury i symbolicznego trwania państwa, okazał się bezcenny dla odzyskania pełnej suwerenności w 1989 roku.

Słowa kluczowe: polska emigracja polityczna, uchodźstwo, kultura emigracyjna

POLISH EMIGRATION COMMUNITIES – DIFFERENT STRATEGIES OF ACTION IN THE FACE OF THE COLLAPSE OF THE IDEA OF A SOVEREIGN POLAND IN 1945

Abstract: The article analyzes the strategies adopted by Polish émigré communities after the collapse of the idea of a sovereign Poland in 1945, within the context of the postwar world order dominated by the USSR and the USA. Drawing on the distinction between political realism (Realpolitik) and idealism, the author examines the perspectives of the major powers (Stalin, Churchill, Roosevelt), whose policies led to the marginalization of the Polish government-in-exile in London. Three main émigré orientations are identified: the “uncompromising” group – rejecting the Yalta agreements and anticipating an East-West conflict (e.g. Arciszewski, Anders); the “moderate realists” – accepting partial compromise (e.g. Mikołajczyk and his supporters who returned to Poland); and the “active realists” – accepting dependence on the Soviet Union. The article argues that the historical role of Polish emigration extended far beyond short-term politics and manifested itself in three key spheres: political – through the symbolic preservation of state continuity (“the state-in-exile”); cultural – through a crucial contribution to sustaining and developing independent Polish thought and national heritage (notably the Paris-based “Kultura” and London’s “Wiadomości”); and material – through assistance provided to the democratic opposition in Poland. In conclusion, the author states that although the immediate political objectives of 1945 proved unattainable, the long-term achievements of Polish émigrés, particularly in the realms of culture and symbolic state continuity, were invaluable for the eventual restoration of full sovereignty in 1989.

Keywords: Polish political emigration, exile, émigré culture

https://doi.org/10.7494/978-83-68219-80-7_10

1. Nowy porządek świata

W stosunkach międzynarodowych istnieją dwie szkoły myślenia: **realizm polityczny** (Realpolitik) – uznający, że polityka oparta jest na kalkulacji siły i interesów narodowych oraz **idealizm polityczny** – widzący esencję stosunków międzynarodowych w „wyższych” wartościach, uznający wpływ czynników ideowych i moralnych na politykę.

Realiści nie odrzucają moralności ani nie szerzą demoralizacji, lecz opisują i wyjaśniają deficyt etyki w polityce międzynarodowej (Radziejewski 2023, s. 40). Brutalność i bezwzględność polityki światowej wynika ze struktury międzynarodowej, w której nie można ignorować faktów zastanych, niemożliwych do pomijania bez poważnych konsekwencji. Kompasem poznawczym dla tego typu analiz jest klasyfikacja Maksa Webera, który mówił o etyce przekonania i etyce odpowiedzialności. Realizowanie w polityce zasad moralnej słuszności może mieć negatywne, a nawet tragiczne konsekwencje dla zbiorowości.

Idealizm polityczny odwołujący się do wartości takich jak pokój, sprawiedliwość, ład demokratyczny etc. jest wyrazem niezgody na odwieczną brutalność i tragizm stosunków międzynarodowych oraz próbą ich przezwyciężenia. Niektórzy uważają, że idealizm polityczny jest bronią słabszych.

Wybitny teoretyk stosunków międzynarodowych i ekspert od realizmu politycznego Kannath Waltz podkreślał, jak ważna jest umiejętność trafnego rozpoznania istoty struktury międzynarodowej. W ramach europocentrycznej polityki ostatnich trzech stuleci co najmniej pięć wielkich mocarstw dążyło do pokojowego współistnienia, a niekiedy wikało się w walkę o dominację. Druga wojna światowa położyła kres temu modelowi i wyłoniła się nowa struktura międzynarodowa, w której dominację przejęły dwa supermocarstwa: USA i ZSRR (Waltz 2010, s. 147).

Niewątpliwie rok 1944 był tym, w którym rodził się nowy porządek świata, choć należałoby przyjąć, że jego zręby zarysowywały się już na konferencji w Teheranie. Rola decydentów należała do Wielkiej Trójki, a faktycznie do USA i ZSRR. Każda ze stron patrzyła na porządek świata z własnej perspektywy.

Józef Stalin, w miarę jak jego sukcesy militarne w walce z Niemcami rosły, stawał się głównym partnerem koalicji określającym swoje warunki pokoju. Definiował je tak, jak od wieków robili to rosyjscy mężowie stanu – wokół granic ZSRR należało stworzyć możliwie najszerszy pas bezpieczeństwa. Od każdego z chwilowych partnerów chciał uzyskać wszystko, co tylko było możliwe, działaniami dyplomatycznymi, a to, czego nie uzyskalby tą drogą, zabrałby siłą, nie podejmując ryzyka wojny. Po bitwie stalingradzkiej zaczął nabierać pewności, że gdy wojna się zakończy, większość terytoriów, które staną się przedmiotem dyskusji, znajdzie się w posiadaniu ZSRR. Można by rzec: Stalin pozostawiał utworzenie kształtów powojennego świata postępowi swojej armii (Kissinger 1996, s. 433, Roberts 2006).

Wielka Brytania występowała w tej wojnie z pozycji słabszego. Winston Churchill robił, co mógł, by znajdując się w takiej pozycji, realizować tradycyjną politykę swego kraju, a więc jeśli świat nie miał pozostać na łasce najsilniejszego i najbardziej bezwzględnego gracza, pokój musiał zostać oparty na jakiejś równowadze „checks and balances” (Gargarella 2022). Szkoła ta zakłada, że w świecie przeciwstawnych interesów i konfliktów między nimi, zasad moralnych nie można nigdy w pełni zrealizować, ale co najwyżej zbliżyć się do nich poprzez zawsze czasowe równoważenie interesów i niepewne rozwiązywanie konfliktów.

Rozumiał on również dobrze, że po zakończeniu wojny Wielka Brytania nie będzie mogła bronić swoich interesów, a tym bardziej decydować o równowadze sił. Robiąc dobrą minę, rozumiał, że rola, jaką jego naród odegrał w czasie wojny, była ostatnim występem Wielkiej Brytanii w charakterze rzeczywiście niezależnego mocarstwa światowego. Dlatego dla Churchilla żaden aspekt sojuszniczej dyplomacji nie był ważniejszy od tego, by stworzyć solidne podstawy przyjaźni z Ameryką, dające pewność, że Wielka Brytania nie stanie wobec powojennego świata samotnie.

Stany Zjednoczone również realizowały swoje interesy. W czasie wojny konsekwentnie odmawiały podejmowania dyskusji nad powojennym kształtem ładu międzynarodowego. Roosevelt słusznie zakładał, że dyskusja o warunkach pokoju z Niemcami może prowadzić do podziałów między aliantami, a chciał, aby cała energia była skierowana na wygraną w wojnie z Niemcami i ich bezwarunkową kapitulację. Jednak odmowa podjęcia dyskusji na temat kształtu powojennego świata w czasie wojny nie spowodowała, że sam Roosevelt nie miał wpływu na jego kształt. Wprost przeciwnie, elementy tego porządku zostały wypracowane na kolejnych konferencjach: w Dambar-ton Oaks w sprawie powołania Organizacji Narodów Zjednoczonych czy w Bretton Woods dla systemu finansowego świata. Na razie jednak najważniejszy był cel najbliższy: pokonanie wspólnego wroga, czyli hitlerowskich Niemiec i na tym głównie się koncentrował, odkładając na plan dalszy sprawę ładu międzynarodowego (Dallek 1995, Kissinger 1996, s. 441, Kimball 2002).

2. Polska w nowym układzie sił

Gdy w czerwcu 1944 roku Alianci wylądowali w Normandii i posuwali się od zachodu, los Niemiec był przesądzony. Z drugiej strony na froncie wschodnim, w miarę jak sytuacja stawała się coraz bardziej korzystna, Stalin zaczął podnosić stawkę – jej elementem składowym była przyszłość Polski. W 1941 r. domagał się uznania granic z 1941 roku i wykazywał gotowość do uznania Rządu RP w Londynie. W 1942 roku zaczął uskarżać się na skład tego rządu, w 1943 roku zerwał stosunki dyplomatyczne pod pretekstem sprawy katyńskiej i intensywnie zaczął wspierać komunistów w Polsce,

aż wreszcie w 1944 roku powołał tzw. Komitet Lubelski, który wkrótce przekształcił w Tymczasowy Rząd Jedności Narodowej. W 1941 roku głównym problemem Stalina były granice, a w 1945 – polityczna kontrola terenów poza tymi granicami.

Jeśli z tak zarysowanej perspektywy interesów trzech mocarstw spojrzymy na „sprawę Polski”, to łatwo dostrzeżemy, że pole manewru Rządu RP w Londynie było dramatycznie małe, a wybór między realizmem a idealizmem politycznym był trudny do zdefiniowania.

Przypomnieć należy, że rząd RP kierowany przez Stanisława Mikołajczyka do 22 listopada 1944 roku posiadał w zasadzie poparcie czterech reprezentowanych w Radzie Narodowej stronnictw politycznych (Stronnictwa Ludowego, Polskiej Partii Socjalistycznej, Stronnictwa Narodowego i Stronnictwa Pracy). W miarę jak druga wojna światowa wchodziła w swoje decydujące stadium i zbliżała się klęska III Rzeszy, pozycja Rządu RP w Londynie ulegała stałemu pogorszeniu. Armia Czerwona zajmowała coraz większe tereny wchodzące w skład II Rzeczypospolitej, władze sowieckie przejmowały administrację obszarów leżących na wschód od linii Curzona, likwidując działające tam oddziały AK, a w województwach leżących na zachód od tej linii przekazywały władzę lewicowemu Polskiemu Komitetowi Wyzwolenia Narodowego, pozostającemu pod kontrolą i opieką Moskwy. Pozycję polskiego rządu dodatkowo osłabiały wewnętrzne spory¹.

Wobec braku porozumienia z pozostałymi stronnictwami w sprawie Wilna i Lwowa 24 listopada 1944 roku Mikołajczyk złożył dymisję. Jego zabiegi o porozumienie z Moskwą były krytykowane przez opozycję zarówno na emigracji, jak i w kraju. Na czoło opozycji wysuwał się Tadeusz Bielecki ze Stronnictwa Narodowego, ale nie tylko. Mikołajczyka atakowały też „Wiadomości Polskie”, „Orzeł Biały” oraz tacy politycy socjalistyczni jak Adam Ciołkosz i Adam Pragier, Tadeusz Tomaszewski, a także z jego własnego stronnictwa Jerzy Kuncewicz i Zygmunt Nagórski.

Nowym premierem w końcu listopada 1944 roku został przerzucony z kraju działacz PPS Tomasz Arciszewski. W nowym rządzie zabrakło przedstawicieli Stronnictwa Ludowego, którzy w większości solidaryzowali się z Mikołajczykiem. Rząd Arciszewskiego, określany jako „niezłomny” czy „rząd protestu”, nie mógł więc liczyć na poparcie Wielkiej Brytanii i Stanów Zjednoczonych, które były zainteresowane utrzymaniem sojuszu z ZSRR. Wprawdzie Churchill był zaniepokojony wzrostem potęgi Związku Sowieckiego i sukcesami Armii Czerwonej (zwłaszcza na Bałkanach), ale zdawał sobie sprawę z osłabienia Wielkiej Brytanii, ograniczonych możliwości jej działania i rosnącego uzależnienia od Stanów Zjednoczonych.

Rząd brytyjski niechętnie więc odnosił się do gabinetu Arciszewskiego, kontakty z nim utrzymywał przez urzędników niższych szczebli i za pośrednictwem ambasadora

¹ Na temat polskiej emigracji politycznej istnieje bogata literatura, m.in.: Friszke (1999), Szczerbiński i Wolsza (2003), Duraczyński i Turkowski (1997), Urban (1998), Sudziński (2007).

Edwarda Raczyńskiego. Dokąd trwała wojna, Brytyjczycy byli natomiast zainteresowani utrzymaniem Polskich Sił Zbrojnych pozostających w dyspozycji naczelnego dowództwa brytyjskiego. Udzielali więc nadal pomocy i kredytów polskiego rządu, akceptując nawet powiększenie Polskich Sił Zbrojnych.

Coraz bardziej natomiast rosła pozycja Stalina, który metodą faktów dokonanych, bez większych protestów Zachodu, przejmował kontrolę nad Polską. Przypomnijmy tylko najważniejsze fakty.

31 grudnia 1944 roku Krajowa Rada Narodowa w miejsce PKWN powołała Rząd Tymczasowy z premierem Edwardem Osóbką-Morawskim. Już 4 stycznia 1945 roku ZSRR uznał Rząd Tymczasowy i wywarł presję na prezydenta Edvarda Beneša, aby uczyniła to również Czechosłowacja.

Na konferencji w Jałcie w lutym 1945 roku w sprawie Polski zapadły decyzje uwzględniające punkt widzenia Stalina, gdzie Rząd RP w Londynie został de facto wyeliminowany z gry.

W marcu zostali aresztowani przywódcy polskiego podziemia. Proces odbywał się równolegle z trwającymi w czerwcu 1945 roku rozmowami nad kształtem nowego rządu, w którym rolę listka figowego pełnili tzw. demokraci z Londynu z Mikołajczykiem na czele.

Te fakty dobitnie pokazywały, jaka jest realna siła Stalina w kształtowaniu nowego ładu w Europie. Legalne władze RP w Londynie już od dłuższego czasu były zupełnie wyłączone z procesu decyzyjnego w polskich sprawach. Mogły jedynie protestować, co też uczynił zarówno rząd Tomasza Arciszewskiego, jak i prezydent Władysław Raczkiewicz. Oczywiście protesty te nie odniosły żadnego skutku.

W dniu 5 lipca 1945 roku rządy mocarstw zachodnich uznały ukonstytuowany w Warszawie Tymczasowy Rząd Jedności Narodowej, co oznaczało również cofnięcie uznania dla Rządu RP w Londynie. W tym samym dniu premier Tomasz Arciszewski przyjął ambasadorów Wielkiej Brytanii i Stanów Zjednoczonych, którzy złożyli mu oświadczenia obu rządów, że utworzenie Tymczasowego Rządu Jedności Narodowej w Warszawie jest spełnieniem postanowień krymskich, co oznacza cofnięcie uznania dla Rządu RP w Londynie (Kacewicz 1979, Lucas 1978, 1982, Mason 2018).

Jak pisał ambasador Raczyński: „Mikołajczyk i jego czterej towarzysze, udając się do Polski, wzięli ze sobą platoniczne błogosławieństwo Waszyngtonu i Londynu, i to mocarstwa zachodnie uznały za ostateczne wywiązanie się z odpowiedzialności za dalsze losy polskiego sojusznika” (Raczyński 1989, s. 351).

3. Emigracja – strategie działania

Jakie zatem polityczne narzędzia pozostały ośrodkom politycznej emigracji w Londynie? Można powiedzieć, że znikome. Jedynie konflikt pomiędzy dwoma politycznymi graczami: Ameryką i Rosją mógł odwrócić kartę polskiej historii. Jednakże liczenie na

nowy konflikt w warunkach dopiero co zakończonej wojny, która pochłonęła miliony ofiar i przyniosła dramatyczne zniszczenia, graniczyło z szaleństwem, zwłaszcza w momencie gdy społeczeństwa zachodnie, szczególnie społeczeństwo amerykańskie, domagały się jak najszybszego powrotu żołnierzy do domów. Podobne nastroje panowały w Polsce. Na tym polegał dramatyzm polskiego losu.

Wybitny przedstawiciel szkoły realistycznego myślenia na emigracji i współtwórca linii politycznej „Kultury” Juliusz Mieroszewski pisał: „Rozbitek na bezludnej wyspie wymachujący chusteczką do nosa w nadziei, że dojrzy go jakiś okręt czy samolot – jest niewątpliwie figurą romantyczną i patetyczną. Niemniej rozbitek ten uprawia jedyną realistyczną politykę, jaka pozostała mu do dyspozycji – ponieważ innych możliwości nie ma” (Mieroszewski 1973, s. 70).

Postawy środowisk emigracyjnych wobec beznadziejnej sytuacji, w jakiej znalazła się Polska na przełomie lat 1944/1945, były zróżnicowane i można byłoby je ująć w trzy kategorie.

„**Niezlomni**” – nie godzili się na kompromis z komunistami i odrzucali ustalenia konferencji w Jałcie w sprawie polskiej. Uważali, że powojenny układ sił nie zapewni pokoju, lecz stanie się źródłem nowych konfliktów między Wschodem i Zachodem, w wyniku czego może dojść do rozpadu ZSRR i odzyskania przez Polskę niepodległości. Do głównych rzeczników tej orientacji można zaliczyć: prezydenta Władysława Raczkiewicza i rząd Tomasza Arciszewskiego, część generalicji na czele z Władysławem Andersem i Kazimierzem Sosnkowskim. Zapleczem dla tej orientacji zarówno w kraju, jak i zagranicą było Stronnictwo Narodowe, ale opcja ta miała zwolenników również w innych ugrupowaniach politycznych, w tym wśród socjalistów (Pragier 1966, s. 905–920, Raczyński 1989, s. 293–350, Woźniczka 2007, s. 190–191).

Drugą orientację, którą można nazwać „**umiarkowanymi realistami**”, reprezentował Mikołajczyk. Odrzucał on prawdopodobieństwo nowej wojny. Polska i Europa Środkowo-Wschodnia miały być strefą ścierania się wpływów Wschodu i Zachodu. Poglądy te podzielał m.in. gen. Stanisław Tatar i płk Jan Rzepecki, szef BIP-Komendy Głównej AK, Delegat Sił Zbrojnych na Kraj oraz część polityków emigracyjnych. Związany z Mikołajczykiem Stanisław Grabski, przewodniczący Rady Narodowej, opublikował w grudniu 1944 roku broszurę *Nil desperandum*, w której opowiadał się za granicą na Odrze i Nysie Łużyckiej i akceptował linię Curzona (Grabski 1945). Pod koniec 1945 roku prezydent Raczkiewicz na wniosek premiera Arciszewskiego rozwiązał Radę Narodową pozostającą w opozycji do polityki rządu. W tej sytuacji Grabski zainicjował deklarację jej 14 członków, w której potępiono politykę „protestu narodowego” prowadzoną przez rząd i postulowano, aby punkt ciężkości polityki narodowej przenieść do kraju i utworzyć tam rząd, opierając się na ustaleniach konferencji jałtańskiej. Wspomnianą deklarację podpisali m.in. Stanisław Grabski, Arka Bożek i Elżbieta Kor-

fanty. Wielu wraz z Mikołajczykiem powróciło do kraju, a niektórzy objęli ważne funkcje. Gdy przestali być potrzebni, zostali odsunięci i represjonowani.

Trzecia orientacja, pełna niechęci wobec komunistów i obaw wobec ZSRR, godziła się z faktem zależności od wschodniego sąsiada oraz z tym, że Stalin wyznacza zakres wolności i suwerenności. Można by ją określić jako „aktywni realiści”. Stanowisko takie zajmowali niektórzy socjaliści w Wielkiej Brytanii (np. Ludwik Grosfeld i tzw. grupa zjednoczonych socjalistów) czy we Francji (np. Henryk Jabłoński). Osobną grupę stanowili komuniści lub ich sympatycy, którzy byli nieliczni na Zachodzie, głównie we Francji (Kersten 1989, 1993).

Te zróżnicowane „strategie” były wynikiem dość realistycznej oceny sytuacji geopolitycznej wynikającej z rosnącej siły Moskwy, która zamierzała obszar Europy Środkowo-Wschodniej uczynić strefą swoich wpływów i stopniowo sowietyzować. W tym zakresie środowiska emigracyjne, mimo dzielących je różnic, miały dość jednolity pogląd.

Przewidywano, że proces sowietyzowania Europy Środkowo-Wschodniej będzie długotrwały i wieloetapowy. Zjawisko to trafnie opisano w analizach dokonanych na łamach emigracyjnej prasy, m.in. przez dra Stefana Mękarskiego z Ligii Niepodległości Polski² czy dra Adama Pragiera z PPS³. Przykładem może być *Raport o sytuacji wewnętrznej w Polsce 1945 roku*. Pogląd taki wyrażała zdecydowana większość ugrupowań politycznych na emigracji. W styczniu 1946 roku działacze Rady Polskich Stronnictw Politycznych (Polska Partia Socjalistyczna, Polskie Stronnictwo Demokratyczne, Stronnictwo Pracy, Stronnictwo Ludowe „Wolność” i Stronnictwo Narodowe) pisali: „Godność człowieka w Polsce jest podeptana. Nie istnieje wolność religii, sumienia i przekonań. Znikły niezależne sądy. Nie ma wolności słowa, prasy, zrzeszeń i zgromadzeń. Bezpieczeństwo osobiste nie istnieje. Nad całym życiem kraju ciąży totalistyczna władza rządu, który narzucono narodowi polskiemu z zewnątrz i który utrzymuje się na powierzchni jedynie dlatego, że kraj jest okupowany przez obce armie i rządzony jest systemem policyjnym” (Wolsza 2007, s. 28).

Emigranci polityczni bez względu na opcje polityczne uważali, że przyszłość Polski zależeć będzie od relacji Waszyngton – Moskwa. W konsekwencji przyjmowano, że

² Stefan Mękarski (1895–1985), poseł na Sejm III kadencji, uczestnik kampanii wrześniowej, pozostał na emigracji w Anglii, był współzałożycielem Instytutu Badań Zagadnień Krajowych, publikował na łamach „Dziennika Polskiego i Dziennika Żołnierza” m.in. *Fatamorgana* (nr 74 z 27 marca 1945 r.), *Nowe esperanto* (nr 70 z 22 marca 1945 r.), *Diariusz kultury polskiej 1945* (nr 97 z 23 kwietnia 1946 r.) czy *Kalkulacje zawiedzione*, „Listy z Londynu” nr 22/70 z 27 października 1944 r.

³ Adam Pragier (1886–1976), poseł na Sejm II kadencji, polityk i publicysta, aktywnie uczestniczył w życiu polskiego wychodźstwa od pierwszych dni drugiej wojny światowej, członek Rady Narodowej RP. Publikował w „Dzienniku Polski i Dzienniku Żołnierza”, a następnie w „Wiadomościach”, gdzie wraz ze Stefanią Zahorską prowadził stałą rubrykę „Puszka Pandory”.

Polska, mimo rozczarowań, nie ma wyboru i musi orientować się na Amerykę. Istniała także powszechna zgoda odnośnie do statusu Polski jako kraju, który znajduje się pod okupacją sowiecką. W związku z tym odpowiedzialność za walkę o pełną wolność i niezależność Polski spoczywać musi na emigracji. Jak to określił jeden z socjalistycznych działaczy, rolą emigracji było „trzymanie kwestii niepodległości Polski” w kursie spraw międzynarodowych. Kraj nazywany „wielkim niemową” miał zakneblowane usta i nie mógł przemawiać w swoim imieniu. Jego główną rolą było przetrwanie i zachowanie substancji biologicznej.

4. Rok 1947

Rok 1947 stanowił pewną istotną cezurę dla emigracji politycznej. Wprawdzie w sensie formalnym rząd Tomasza Arciszewskiego nie był już podmiotem w stosunkach międzynarodowych, ale nadal istniał jako centralny ośrodek politycznej emigracji. Jednak w kwietniu śmiertelnie chory prezydent Raczkiewicz powziął decyzję o wyznaczeniu na swojego następcę Augusta Zaleskiego, pomimo że od sierpnia 1944 roku nominowany był na to stanowisko Tomasz Arciszewski. Doprowadziło to do dymisji premiera Arciszewskiego i powołania na to stanowisko Tadeusza Bora Komorowskiego, a w konsekwencji wywołało rozłam wśród politycznych ugrupowań (Siwik 1994a).

Rok 1947 jeszcze z innego powodu jest ważny dla polskiej emigracji. Totalną klęską zakończyła się próba demokratycznego dojścia do władzy przez polityków reprezentujących opcję z Mikołajczykiem na czele. Władze komunistyczne nie dopuściły do wolnych wyborów. Obóz Mikołajczyka uzyskał w wyborach do sejmu tylko 10%, a Mikołajczyk zagrożony aresztowaniem z pomocą Amerykanów uciekł z kraju i ponownie znalazł się na emigracji (Paczkowski 1991).

Można by powiedzieć, że rację mieli „niezłomni”, którzy przewidzieli taki obrót sytuacji. Mikołajczyk spotkał się z powszechną infamią środowisk londyńskich. Nazywano go „kawalerem jałtańskim”, uważano za zdrajcę, który sprzeniewierzył się moralnym zasadom i wszedł w kolaborację z akolitami Moskwy, legitymizując ich rządy.

Ale Mikołajczyk i podobnie myślący obóz też miał mocne argumenty przemawiające za obraną strategią. Z jednej strony był popierany przez zachodnich aliantów z Ameryką na czele, którzy nie chcieli konfliktu z Moskwą, z drugiej – dla rodaków w kraju był ostatnią nadzieją na to, że Polska nie zostanie zsowietyzowana. Podjął próbę i poniósł klęskę, ponieważ w tej grze nie było wygranych. Podobną klęską zakończyła się gra prezydenta Edvarda Beneša w Pradze. Moskwa wszystkim krajom Europy Środkowo-Wschodniej narzucała swój model, a przeciwnicy byli eliminowani. Mikołajczyk przegrał, ale też dał dowód, że kooperacja z władzami komunistycznymi nie jest możliwa. Polityka jest osiąganiem rzeczy możliwych, ale tym różni się od teorii politycznych, że weryfikowana jest w praktyce. W tym sensie strategia obozu Mikołajczyka

miała znamiona racjonalności. Po 1948 roku było jasne, że emigracja musi na nowo zdefiniować swoją rolę.

5. Realizm vs. idealizm

Spróbujmy przeanalizować z dzisiejszej perspektywy – już po upadku systemu komunistycznego – która ze strategii przyjętych przez emigrację okazała się optymalna pod względem interesu Polski.

Dwubiegunowy układ międzynarodowy nie pozostawiał zbyt wielkiego pola manewru. Polska z woli międzynarodowych graczy znalazła się w strefie wpływów Moskwy. Brakowało narzędzi politycznego oddziaływania, aby to zmienić. W podobnej sytuacji znaleźli się zresztą emigranci polityczni z pozostałych krajów zza żelaznej kurtyny. Po pierwszym szoku związanym z utratą ojczyzny oraz adaptacją do nowych warunków pojawiła się tendencja do tworzenia ponadpartyjnych struktur reprezentujących ujarzmione narody. Łączyły ich głównie mniej lub bardziej radykalny antykomunizm, uznanie zainstalowanych rządów jako marionetki Moskwy oraz przeświadczenie, że w dwubiegunowym świecie, jaki ukształtował się po Jałcie, należy orientować się na Stany Zjednoczone. Konsekwencją tego było uzależnienie uchodźstwa politycznego krajów zza żelaznej kurtyny od strategii politycznej Stanów Zjednoczonych w stosunku do Moskwy. Polityka amerykańska wobec Europy Środkowo-Wschodniej nie tylko była uwarunkowana relacjami amerykańsko-radzieckimi, ale stanowiła ich pochodną.

Ewolucja polityki USA po 1947 roku, która przyjęła nazwę doktryny Trumana, zakładała globalne zaangażowanie Stanów Zjednoczonych w obronę wolnego świata przed komunizmem (Judt 2008). W tej doktrynie Europa Wschodnia, jak konsekwentnie nazywano tę część kontynentu europejskiego, uznana została za obszar już zdominowany przez komunizm. W praktyce oznaczało to niedopuszczenie do dalszej ekspansji komunizmu na Zachód, a nie wyzwalamie Wschodu. W dokumencie *Aide Memoire Polish Emigration Possibilities of Its Participation in the Anti-Communist Front* z czerwca 1948 roku wskazywano na możliwość wykorzystania emigracji do antykomunistycznych działań (Mania 1994, s. 73).

Jak powiedział zastępca dyrektora Biura ds. Europejskich (Bureau of European Affairs) w Departamencie Stanu Llewellyn Thompson, „wyzwalanie i przywracanie niepodległości” miało nastąpić w dalszej perspektywie, a na razie celem Amerykanów było podtrzymywanie nadziei i morale Polaków oraz okazywanie stałego zainteresowania Polską i sympatii jej mieszkańcom (Mazurkiewicz 2016, s. 28). Z amerykańskiego punktu widzenia największym problemem utrudniającym polityczne wykorzystanie emigracji było jej rozbitcie. W raporcie dla Waszyngtonu radca amerykańskiej ambasady w Londynie donosił, że polskich emigrantów łączą trzy silne uczucia: „1) gorący patriotyzm 2) pragnienie powrotu do domu 3) nienawiść wobec komunizmu. Te czynniki

umacniają ich poczucie wspólnoty, dzięki czemu mogą stać się zimnowojennym atutem (*cole war asset*) po naszej stronie” (Machcewicz 1999, s. 45).

Szczególnie silne naciski Stanów Zjednoczonych w kierunku zjednoczenia polskiej emigracji podjęto wiosną 1949 roku, kiedy finalizowały się przygotowania do powołania Komitetu Wolnej Europy, który miał przejąć ciężar kontaktów z uchodźstwem zza żelaznej kurtyny. Amerykanie podkreślali wyraźnie, że Komitet Wolnej Europy podejmie współpracę z wyłonioną spośród szerokiego spektrum polityków emigracyjnych reprezentacją (komitet/ komisja), ale nie może to być instytucja powiązana z emigracyjnym rządem. Amerykanie uznawali rząd w Warszawie i nie chcieli wchodzić w kolizję na tym polu.

Wśród grup uchodźców z Europy Środkowo-Wschodniej środowisko polskie odgrywało kluczową rolę nie tylko ze względu na liczebność, ale także pozycję Polski pośród radzieckich satelitów. Uchodźcy z tych krajów powołali struktury organizacyjne aspirujące do roli politycznej reprezentacji własnych narodów. Starali się jednak unikać formuły opartej na „prawnej kontynuacji struktur państwowych”, tak jak robili to Polacy.

Mimo nacisków amerykańskich stronie polskiej nie udało się powołać jednego ośrodka reprezentującego uchodźstwo polityczne. Nie oznaczało to jednak, że w polityce *containment* nie wykorzystano możliwości oddziaływania na społeczeństwo w Polsce poprzez ośrodki emigracyjne. Dotyczyło to zwłaszcza kanałów komunikacyjno-kurierskich, jakimi dysponowały siły polityczne „obozu zjednoczenia”, szczególnie Stronnictwo Narodowe, ale także PPS. Akcja ta zakończyła się kompromitującą „afery Bergu”, którą propaganda w PRL wykorzystała do ostatecznego rozprawienia się z opozycją w kraju⁴.

Amerykanie dość instrumentalnie używali emigracji do własnych politycznych celów, zwłaszcza w okresie zimnej wojny. W warunkach odprężenia w stosunkach Waszyngton – Moskwa polityczne znaczenie emigracji zmalało.

Mimo wszystko rola uchodźstwa dla Polski jest trudna do przecenienia. W sytuacji, gdy kraj pozbawiony był pełnej suwerenności, emigracja odgrywała „rolę zastępczą”, mogła wobec Zachodu prezentować polski punkt widzenia. Bez względu na barwy polityczne, społeczność emigracyjna pełniła rolę służebną wobec kraju.

6. Płaszczyzny aktywności

Działalność emigracji można rozpatrywać w trzech płaszczyznach aktywności: politycznej, kulturowej i materialnej.

⁴ Tzw. afera Bergu wybuchła na emigracji i była spowodowana ujawnieniem na przełomie 1952/1953 przez władze PRL – w związku z prowadzoną przez Urząd Bezpieczeństwa operacją pod kryptonimem „Cezary” – faktu finansowania przez zachodnie służby specjalne ośrodka szkoleniowego Rady Politycznej w Bergu koło Monachium. Szerzej w pozycji Siwik (2001).

W warstwie politycznej różnice w strategii dotyczyły form organizacyjnych. W największym skrócie można powiedzieć, że część środowisk politycznej emigracji uważała, że sytuacja w Polsce nabierze cech trwałości i wobec tego należy stworzyć nowe formy organizacyjne emigracji. Szczególnie silnie te tendencje ujawniły się po tzw. kryzysie prezydenckim w 1947 roku. Część działaczy pozostała konsekwentnie przy formule „państwa na wygnaniu”. August Zaleski jako prezydent był przedstawicielem tej grupy. Ich zdaniem w warunkach przemocy i dyktatu jedynie odwołanie się do imponderabiliów narodowych miało sens. Dość sarkastycznie pisał o tej grupie Gustaw Herling-Grudziński w liście do Zygmunta Zaremby: „Panuje tam ogólne przekonanie, że wystarczy cierpliwie reprezentować i odcinać kupony od renty legalności” (Herling-Grudziński, 1947). Próby zażegnania rozłamu zakończyły się fiaskiem, mimo że misja zjednoczeniowa, której podjął się gen. Kazimierz Sosnkowski, była bliska pomyślnego zakończenia. W konsekwencji w 1954 roku utrwalił się podział na „obóz zamkowy” z prezydentem Augustem Zaleskim i „obóz zjednoczenia”, w skład którego wchodziły: Rada Trzech, Egzekutywa Zjednoczenia Narodowego i Tymczasowa Rada Jedności Narodowej (Siwik 1994b). Obozy te połączyły się dopiero w 1972 roku po śmierci Augusta Zaleskiego, zachowując jednak formułę „państwa na wygnaniu”.

Taka koncepcja z biegiem lat wydawała się coraz bardziej anachroniczna, zwłaszcza wobec braku oznak upadku systemu komunistycznego. Kolejne dorastające pokolenia nie interesowały się „politycznymi strukturami emigracji”. Wobec przedstawicieli „obozu zamkowego” używano ironicznego określenia „emeryci minionej epoki”. Sytuacja uległa zmianie po upadku komunizmu. Pierwsze po drugiej wojnie wolne wybory i wyłonienie Lecha Wałęsy na prezydenta Polski zrodziły potrzebę symbolicznego nawiązania do ciągłości niepodległego państwa polskiego. Patrząc z tej perspektywy, można powiedzieć, że to jednak „obóz zamkowy” miał „historyczną rację”, wybierając strategię trwania przy imponderabiliach. W warstwie symbolicznej proces zachowania ciągłości niepodległego państwa domknął się, kiedy na Zamku Królewskim w Warszawie prezydent RP na uchodźstwie Ryszard Kaczorowski przekazał symbole II Rzeczypospolitej prezydentowi Lechowi Wałęsie. Wówczas, 22 grudnia 1990 roku, ostatni trwający na straży symboli władzy prezydenckiej prezydent na emigracji przekazał wybranemu przez naród prezydentowi insygnia władzy: chorągiew RP, pieczęć Kancelarii Prezydenta, oryginał konstytucji z 1935 roku, a także order Orła Białego i Polonia Restituta.

Drugą niezwykle istotną sferą była kultura. Kulturotwórczą rolę emigracji można porównać jedynie do dorobku Wielkiej Emigracji w XIX wieku (Wejs-Milewska i Rogalewska 2009, 2016). Rozwój kultury polskiej po drugiej wojnie światowej miał szczególny charakter. Czas wojny i zmiany polityczne spowodowały, iż część polskiego dorobku kulturalnego i środowisk kulturotwórczych została rozproszona, a przede wszystkim bezpowrotnie zniszczona. Przez blisko pół wieku wiele dóbr kultury narodowej, a nawet

podstawowych faktów historycznych, było pracowicie niszczonej i zacieranym w świadomości narodowej. Z programów szkolnych, z programów teatralnych czy wydawnictw na pół wieku zniknęły nazwiska wielkich twórców kultury i nauki oraz wybitne dzieła narodowej kultury. Cenzura skutecznie ograniczała do nich dostęp. Do Polski dzieła wydawane zagranicą mogły być tylko przemycane, a w najlepszym wypadku dostępne były tylko za specjalnym pozwoleniem dla naukowców tzw. resy (mogły być czytane tylko w bibliotekach naukowych na miejscu i nie było zgody na kopiowanie nawet fragmentów).

Szczególną rolę w rozwoju polskiej kultury zagranicą odegrała „Kultura”. Jerzy Giedroyc stworzył w Maisons-Lafitte koło Paryża ośrodek koncepcyjny, programowy i organizacyjny, którego zadaniem było wpływanie na społeczeństwo w Polsce i środowiska emigracji politycznej. Mimo że swoim pojmowaniem racji stanu był związany z II Rzeczpospolitą, umiał ustrzec się od wszelkiego epigoństwa i potrafił odczytywać „znaki czasu” nowej epoki z jej przemianami i prądami. Stworzył wydawnictwo zrozumiałe dla Polaków w kraju, jednocześnie nie zrywając z tym, co w tradycji było ważne na przyszłość (Kudelski 1995, s. 32, Mencwel 2024). Jak pisał Krzysztof Pomian, postanowił rozwijać działania „przede wszystkim na obszarze kultury: literatury, filozofii, historii, nauki, sztuki – słowem tych dziedzin, które kształtują wierzenia zbiorowe i postawy jednostek” (Pomian 2000, s. 24). Giedroyc skupił wokół „Kultury” i Instytutu Literackiego wielu autorów. Najbardziej znaczącymi byli: Juliusz Mieroszewski, Andrzej Bobkowski, Jerzy Stępowski, Józef Czapski, Konstanty Jeleński, Bohdan Osadczuk i Leopold Unger. Upowszechnił dorobek wybitnych pisarzy emigracyjnych: Witolda Gombrowicza, Czesława Miłosza i Gustawa Herlinga-Grudzińskiego, dysydentów ze Wschodu i prace autorów z kraju. Organizował też przekłady i wysyłkę do kraju najnowszych prac autorów zachodnich: Alberta Camusa, Raymonda Arona, George’a Orwella i Jamesa Burnhama (Hofman 2017, 2020, Łastawski 2019, s. 6). Na emigracji było wiele pism, które broniły narodowych wartości, spuścizny kulturalnej oraz walczyły z komunizmem, ale jedynie „Kultura” stała się symbolem niezależnej myśli wolnej Polski. Budowała mosty wzajemnego porozumienia narodów mimo przeciwności politycznych i materialnych oraz społecznych uprzedzeń.

„Kultura” była wydawana przez Instytut Literacki założony przez Jerzego Giedroycia, Zofię Hertz, Zygmunta Hertza i Józefa Czapskiego. Instytut był jednym z najbardziej zasłużonych polskich wydawnictw emigracyjnych. W serii „Biblioteka Kultury” do 2000 roku ukazało się 512 tomów literatury polskiej i obcej. W 1962 roku Instytut Literacki w Paryżu rozpoczął wydawanie „Zeszytów Historycznych”, w których publikowano dokumenty, relacje, pamiętniki, wspomnienia i opracowania poświęcone najnowszej historii Polski. Ich wersję „podziemną” wydawało w Polsce Krakowskie Towarzystwo Wydawnicze. Ostatni numer pisma ukazał się w czerwcu 2010 roku (Supruniuk 2011).

Równie silny ośrodek powstał w Londynie pod kierunkiem dawnego redaktora „Wiadomości Literackich” i „Skamandra” Mieczysława Grydzewskiego, który skupił

środowiska intelektualne wokół „Wiadomości Polskich” (Terlecki 1957, Grydzewski 2014, Supruniuk 2023). W 1957 roku Grydzewski ogłosił w „Wiadomościach” ankietę nt. popularności polskich pisarzy na emigracji. Na podstawie wyników powołał emigracyjną Akademię Literatury, zwaną „Akademią Grydzewskiego” i ustanowił doroczną nagrodę „Wiadomości” za najlepszą książkę wydaną na obczyźnie i najlepszy utwór publikowany w tygodniku. W „Wiadomościach Polskich” publikował Zygmunt Nowakowski, Tymon Terlecki, Stanisław i Józef Mackiewiczowie, Wacław Zbyszewski, Wojciech Wasutyński, ale także Gustaw Herling-Grudziński czy Jerzy Stempowski.

Paryska „Kultura” i londyńskie „Wiadomości Polskie”, mówiąc metaforycznie, tworzyły dwa płuca polskiego, suwerennego oddechu intelektualnego i kulturowego, które wzajemnie się dopełniały. Różnica polegała na tym, że „Kultura” miała ambicje wpływania na życie polityczne w kraju, a „Wiadomości” chciały bardziej oddziaływać na emigrację. Dla Giedroycia „Kultura” była narzędziem politycznym i w tym sensie miała charakter bardziej ekskluzywny, podczas gdy „Wiadomości” pozostawały bardziej inkluzywne. Ciągłość, tożsamość, trwanie kultury polskiej znajdowały wyraz w „Wiadomościach”. W pewnym sensie „Wiadomości” wyrażały ideowe przesłanie obozu politycznego „niezłomnych”: „trwać na posterunku” i przechować polskość. Juliusz Mieroszewski ze swoją stałą rubryką „Listy z Wyspy” w „Kulturze” i polemizujący z nim przez ćwierć wieku na łamach „Wiadomości” w rubryce „Puszka Pandory” nie mniej wybitny publicysta Adam Pragier symbolizowali te dwa ośrodki kultury emigracyjnej. Jak podkreśla Jakub Moroz, „w przeciwieństwie do dobrze już przyswojonego dziedzictwa paryskiej »Kultury«” „londyńczycy to wielki nieobecny polskiej świadomości narodowej” (Moroz 2012).

W Londynie działały wydawnictwa: Gryf Publication (1946–1976), Katolicki Ośrodek Wydawniczy „Veritas”, Oficyna Malarzy i Poetów (1966–1980), Polska Fundacja Kulturalna, Polonia Book Fund Ltd. (1959–1991), „Kontra”, w Paryżu: Księgarnia Polska, Libella (1946–1993), Éditions du Dialogue, w Kanadzie: Polski Fundusz Wydawniczy, w USA: Roy Publishers (Danilewicz-Zielińska 1999).

Ważną rolę dla kultury polskiej odegrał Polski Ośrodek Społeczno-Kulturalny (POSK) założony w Londynie przez społeczność polską w Wielkiej Brytanii z inicjatywy prof. inż. Romana Wajdy. Budynek został ukończony w 1974 roku. W Londynie powstał także Instytut Polski i Muzeum im. gen. Sikorskiego placówka kulturalna, stawiająca sobie za cel gromadzenie, opiekę, opracowanie naukowe oraz udostępnianie archiwów polskiego rządu na uchodźstwie i Polskich Sił Zbrojnych na Zachodzie.

Również w Londynie powstały inicjatywy zorganizowania własnych placówek edukacyjnych i powołania do życia polskich galerii oraz stowarzyszeń. W drugiej połowie XX wieku w Anglii działały szkoły Mariana Bohusza-Szyszki, Henryka Gotliba, Mariana Kratochwila, a karierę na artystycznej scenie Londynu robiły „Drian Gallery” Haliny Nałęcz, „Grabowski Gallery” Mateusza Grabowskiego i „Centaur Gallery” małżeństwa

Dinah i Jana Wieliczków. W latach 50. powołano również do życia Zrzeszenie Polskich Artystów Plastyków w Wielkiej Brytanii.

W Stanach Zjednoczonych w Nowym Jorku od 1942 roku działał Polski Instytut Naukowy (Polish Institute of Arts and Sciences of America). Jego założycielami byli znani polscy naukowcy, m.in. antropolog Bronisław Malinowski, historycy Oskar Halecki i Jan Kucharzewski, papirolog Rafał Taubenschlag, geograf i podróżnik Henryk Arctowski i poeta Jan Lechoń. Także w Nowym Jorku w 1943 roku z inicjatywy działaczy Komitetu Narodowego Amerykanów Polskiego Pochodzenia powstał Instytut Józefa Piłsudskiego – instytucja naukowa, archiwum, biblioteka i muzeum. Kulturowa rola emigracji powojennej jest wyjątkowa.

Wreszcie trzecia, niezwykle istotna działalność społeczności emigracyjnej, to pomoc materialna dla prześladowanych rodaków w kraju. Gdy w połowie lat 70. pojawiły się w Polsce ośrodki zinstytucjonalizowanej opozycji z ruchem „Solidarność” włącznie, emigracja udzielała wszelkiej pomocy, także materialnej (Lewicki 1995). Zakres i skala były ogromne, ale temat wykracza poza ramy niniejszego artykułu i dlatego nie będzie omawiany.

Na koniec należy podkreślić, że nie można rozpatrywać strategii działania elit emigracyjnych w krytycznym dla Polski 1945 roku bez rozszerzenia analizy o całokształt dorobek emigracji powojennej. Wynika to z faktu, że historia tego wychodźstwa dobiegła końca w 1989 roku. Takiego szczęścia nie mieli ich antecesorzy. Uchodźstwo powojenne wpisuje się w tradycję Wielkiej Emigracji XIX wieku, która odegrała szczególną rolę w kształtowaniu zbiorowej świadomości (Żurawski vel Grajewski 2015). Już w okresie zaborów utrwalił się stereotyp emigranta politycznego jako niezłomnego patrioty, który nie godząc się na zniewolenie, wybiera wygnanie, aby toczyć walkę o wolność kraju. Emigracja wytworzyła pewien kanon politycznych zachowań, z których na czoło wysuwa się tworzenie pod polską komendą sił zbrojnych, które nie tylko są formacją zbrojną, ale symbolem trwania („Jeszcze Polska nie zginęła póki my żyjemy”...). Drugim charakterystycznym elementem było polityczne życie ogólnonarodowe przejawiające się w strukturach organizacyjnych i nawet, gdy cofnięto uznanie legalnym władzom, kontynuowano model „państwa na wygnaniu”. Trzecim – tworzenie instytucji kultury będących skarbnicami narodowej pamięci. Należy oczywiście mieć na względzie fakt, że dotyczyło to jedynie najbardziej zaangażowanych, zdecydowana większość była pochłonięta walką z codziennymi kłopotami bytowymi. Jednak nawet ci mobilizowali się w sytuacjach krytycznych, często także finansowo wspierali akcje na rzecz kraju.

Literatura

- Dallek R., 1995, *Franklin D. Roosevelt and American Foreign Policy, 1932–1945*, Second Edition, Oxford University Press.
- Danilewicz-Zielińska M., 1999, *Szkice o literaturze emigracyjnej półwiecza 1939–1989*, Zakład Narodowy im. Ossolińskich, Wrocław.

- Duraczyński E., Turkowski R., 1997, *O Polsce na uchodźstwie. Rada Narodowa Rzeczypospolitej Polskiej 1939–1945*, Wydawnictwo Sejmowe, Warszawa.
- Friszke A., 1999, *Życie polityczne emigracji. Druga wielka emigracja 1945–1990*, cz. 1, „Biblioteka Więzi”, Warszawa.
- Gaddis J.L., 2007, *Zimna wojna. Historia podzielonego świata*, tłum. B. Pierzyk, „Znak”, Kraków.
- Gargarella R., 2022, *Checks and Balances: Combining “Institutional Means and Personal Motives”*, [w:] Gargarella R., *The Law As a Conversation among Equals*, Cambridge University Press, s. 124–135. <https://www.cambridge.org/core/books/abs/law-as-a-conversation-among-equals/checks-and-balances/7C2FBFAA7BE49DA918722-BFF0C010C63> [dostęp: 20.07.2024].
- Grabski S., 1945, *Nil desperandum*, Londyn (bez wyd.).
- Grydzewski M., 2014, *Silva rerum*, „Iskry”, Warszawa.
- Herling-Grudziński G., 1947, list do Zygmunta Zaremby z dnia 9 marca 1947 r., Archiwum Polskiego Instytutu Naukowego w Nowym Jorku (APIN), ZZ6/1.
- Hofman I. (red. nauk.), (2017), *„Kultura” i jej odbiór w Polsce*, Wydawnictwo Adam Marszałek, Toruń, Instytut Literacki Kultura, Paryż.
- Hofman I. (red.), 2020, *Redaktor: 20 lat bez „Kultury”*, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin.
- Judt T., 2008, *Powojnie. Historia Europy od roku 1945*, tłum. R. Bartoń, Dom Wydawniczy „Rebis”, Poznań.
- Kaciewicz G.V., 1979, *Great Britain, The Soviet Union and the Polish Government in Exile (1939–1945)*, Springer Netherlands, Dordrecht.
- Kersten K., 1989, *Jałta w polskiej perspektywie*, „Aneks”, Londyn, Niezależna Oficyna Wydawnicza, Warszawa.
- Kersten K., 1993, *Między wyzwoleniem a zniewoleniem. Polska 1944–1956*, „Aneks”, Londyn.
- Kimball W.F., 2002, *Forged in War: Roosevelt, Churchill, and the Second World War*, Ivan. R. Dee, Chicago.
- Kissinger H., 1996, *Dyplomacja*, tłum. S. Głąbiński, G. Woźniak, I. Zych, Philip Wilson, Warszawa.
- Kudelski Z., 1995, *Spotkania z paryską „Kulturą”*, Towarzystwo Opieki nad Archiwum Instytutu Literackiego w Paryżu, Oficyna Wydawnicza Pomost, Warszawa.
- Lewicki R. (red.), 1995, *Pomoc Krajowi przez niepodległe uchodźstwo 1945–1990*, Polskie Towarzystwo Naukowe na Obczyźnie, Londyn.
- Lukas R.C., 1978, *The Strange Allies: Poland and the United States 1941–1945*, University of Tennessee Press, Knoxville.
- Lukas R.C., 1982, *Bitter Legacy: Polish-American Relations in the Wake of World War II*, University Press of Kentucky, Lexington.

- Łastawski K., 2019, *Polityka wschodnia Polski w publicystyce paryskiej „Kultury” i działalności Instytutu Literackiego*, Studia Wschodnioeuropejskie, nr 11, s. 5–15. <https://swe.uw.edu.pl/11.1.pdf> [dostęp: 20.06.2025].
- Machcewicz P., 1999, *Emigracja w polityce międzynarodowej. Druga Wielka Emigracja 1945–1990*, cz. 2, „Biblioteka Więzi”, Warszawa.
- Mania A., 1994, *The National Security Council i amerykańska polityka wobec Europy Wschodniej 1945–1950*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
- Mason A., 2018, *British Policy Towards Poland, 1944–1956*, Springer International Publishing, Cham, imprint: Palgrave Macmillan.
- Mazurkiewicz A., 2016, *Uchodźcy polityczni z Europy Środkowo-Wschodniej w amerykańskiej polityce zimnowojennej (1948–1954)*, Instytut Pamięci Narodowej Komisja Ścigania Zbrodni przeciwko Narodowi Polskiemu, Warszawa, Uniwersytet Gdański, Gdańsk.
- Mencwel A., 2024, *Książka o „Kulturze”*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Mieroszewski J., 1973, *Polska Ostpolitik*, „Kultura”, 6/609, s. 68–79. <https://static.kulturaparyska.com/attachments/f1/38/5605f797c23229775b6250f501020ce033e4ee7a.pdf#page=36> [dostęp: 20.07.2024].
- Moroz J., 2012, *Czułość i nieprzejednanie*, „Pressje”, 30–31, s. 28–40.
- Paczkowski A., 1991, *Stanisław Mikołajczyk czyli klęska realisty*, Omnipress, Warszawa.
- Pomian K., 2000, *W kręgu Giedroycia*, Czytelnik, Warszawa.
- Pragier A., 1966, *Czas przeszły dokonany*, B. Świderski, Londyn.
- Raczyński E., 1989, *W sojuszniczym Londynie. Dziennik ambasadora Edwarda Raczyńskiego 1939–1945*, Niezależna Oficyna Wydawnicza, Warszawa.
- Radziejewski B., 2023, *Nowy porządek globalny. Mocarstwa, średniacy i niewidzialne siły kierujące światem*, Wydawnictwo Nowej Konfederacji, Warszawa.
- Roberts G., 2006, *Stalin's Wars: From World War to Cold War, 1939–1953*, Yale University Press, New Haven.
- Siwik A., 1994a, *Polska emigracja polityczna lat sześćdziesiątych i siedemdziesiątych. Problem pojednania*, Dzieje Najnowsze, nr 1, s. 77–89.
- Siwik A., 1994b, *Władze emigracyjne „Obozu Zjednoczenia” w latach 1954–1972*, Studia Historyczne, z. 4, s. 519–539.
- Siwik A., 2001, *Sprawa Bergu: współpraca Emigracyjnej Rady Politycznej z zachodnimi wywiadami*, Studia Historyczne, t. 44, z. 2, s. 267–288.
- Sudziński R. (red.), 2007, *Emigracja polska wobec sowietyzacji kraju*, Wydawnictwo Adam Marszałek, Toruń.
- Supruniuk M.A., 2011, *Zobaczyć inną Polskę: pomoc paryskiego Instytutu Literackiego dla Polski w latach 1946–1990: koncepcje i realizacja*, Towarzystwo Przyjaciół Archiwum Emigracji, Toruń.

- Supruniuk M.A., 2023, *Mieczysław Grydzewski – uparty czytelnik książek historycznych*, Archiwum Emigracji, nr 31, 2023, s. 11–37. <https://doi.org/10.12775/AE.2023.001>.
- Szczerbiński M., Wolsza T. (red.), 2003, *Z dziejów Polski i emigracji (1939–1989)*, Instytut Kultury Fizycznej Poznańskiej AWF, Gorzów Wielkopolski.
- Terlecki T., 1957, *XXX-lecie Wiadomości*, „Wiadomości”, Londyn.
- Urban A., 1998, *Emigracyjny dramat*, „Bellona”, Warszawa.
- Waltz K., 2010, *Struktura teorii stosunków międzynarodowych*, tłum. R. Włoch, Wydawnictwo Naukowe Scholar, Warszawa.
- Wejs-Milewska V., Rogalewska E. (red.), 2009, *Paryż. Londyn. Monachium. Nowy Jork. Powrześniowa emigracja niepodległościowa na mapie kultury nie tylko polskiej*, Oddział Instytutu Pamięci Narodowej w Białymstoku, Wydawnictwo Uniwersyteckie Trans Humana, Białystok.
- Wejs-Milewska V., Rogalewska E. (red.), 2016, *Paryż. Londyn. Monachium. Nowy Jork. Powrześniowa emigracja niepodległościowa na mapie kultury nie tylko polskiej*, t. 2, Oddział Instytutu Pamięci Narodowej w Białymstoku, Wydawnictwo Uniwersyteckie Trans Humana, Białystok.
- Wnuk M., 2014, *Emigracja z „beznadziei”. Motywacje wyjazdów z Polski w latach osiemdziesiątych*, Polska 1944/45–1989: Studia i Materiały, nr 12, s. 331–348.
- Wolsza T., 2007, *Sowietyzacja Polski i Europy Środkowo-Wschodniej w ujęciu polskich kręgów emigracyjnych na wyspach brytyjskich (1944–1950). Zagadnienia teoretyczne, zakres, skala zjawiska oraz chronologia wydarzeń*, [w:] Sudziński R. (red.), *Emigracja polska wobec sowietyzacji kraju*, Wydawnictwo Adam Marszałek, Toruń, s. 19–31.
- Woźniczka Z. 2007, *Emigracja wobec podziemia w kraju po 1945 r.*, [w:] Sudziński R. (red.), *Emigracja polska wobec sowietyzacji kraju*, Wydawnictwo Adam Marszałek, Toruń, s. 187–209.
- Żurawski vel Grajewski R.P., 2015, *Polskie emigracje 1931–1918*, [w:] Nowak A. (red.), Barańska A., Weeks T.R., Żurawski vel Grajewski R.P., Orłowski B., Miknys R., Janowski M., *Historie Polski XIX wieku*, t. 4. *Narody, wyznania, emigracje, porównania*, Wydawnictwo DiG, Warszawa, s. 115–276.

Zastosowanie kryptografii postkwantowej w sieciach IoT

Tomasz Styn , Michał Pitera , Mateusz Okoń , Patryk Harasik 

Akademia Górniczo-Hutnicza, Wydział Informatyki, Elektroniki i Telekomunikacji, Koło Naukowe Cipher, Kraków

Streszczenie: Artykuł porusza kwestie wprowadzenia postkwantowego algorytmu wymiany klucza kryptograficznego w kontekście sieci Internet of Things (IoT) na podstawie protokołu ZigBee. Coraz szersze zastosowanie internetu rzeczy powoduje konieczność zabezpieczenia urządzeń wchodzących w jego skład oraz komunikacji między nimi. Jest to kluczowa kwestia zarówno dla bezpieczeństwa organizacji wdrażających te technologie jako część infrastruktury, jak i użytkowników końcowych stosujących je w swoich domach. Bezpieczeństwo rozwiązania opiera się na implementacji mechanizmu ML-KEM-512 (Kyber) w celu enkapsulacji klucza symetrycznego stosowanego w szyfrowaniu komunikacji za pomocą AES-256 w trybie CTR. Za pomocą demonstracyjnej sieci IoT przeprowadzone zostały pomiary narzutów czasowych w komunikacji i wymianie klucza. Został oceniony poziom bezpieczeństwa zarówno standardowego, jak i proponowanego rozwiązania. Pozyskano wyniki porównania wydajności komunikacji przed wprowadzeniem zmodyfikowanego rozwiązania i po jego wprowadzeniu w celu oceny potencjału jego powszechnego wdrożenia. Zaobserwowana zmienność czasów przesyłu pakietów została powiązana z funkcjonowaniem sprzętu, co obaliło pierwotną hipotezę retransmisji. Badania wskazują na znaczące obciążenie zasobów sieci i urządzeń, a to pokazuje, że rozwiązanie w tej postaci znajduje zastosowanie między kluczowymi węzłami o większych zasobach obliczeniowych lub w przypadku, gdy bezpieczeństwo przesyłanej informacji ma większy priorytet od prędkości jej transferu.

Słowa kluczowe: cyberbezpieczeństwo, kryptografia postkwantowa, internet rzeczy, ZigBee

APPLYING POST-QUANTUM CRYPTOGRAPHY IN IoT NETWORKS

Abstract: The advent of powerful quantum computing brings risks to the security of cryptographic systems as we know them today. Continuous growth of the already widespread IoT sphere emphasizes the need for securing devices it consists of, and the communication between them. This paper describes the process of implementing ML-KEM-512 (Kyber) – a post-quantum key encapsulation mechanism in the ZigBee protocol. Furthermore, an encryption layer utilizing AES-256 is introduced in a way that is compatible with existing infrastructure, as we feel that this aspect is the key to rapid and widespread adoption among consumers and businesses. Through the use of a demonstrative IoT network based on ZigBee, efficiency measurements were taken before and after implementing the proposed solution. Their analysis was carried out to judge the solution's viability. Transmission time fluctuations were found likely to be connected with hardware, rather than network retransmissions, as was initially theorized. The security level of the paper's proposed solution was evaluated in the context of the existing standard. Results show a significant increase in network and device resource usage, which leads us to believe the solution can find adoption between crucial IoT network nodes, where security of transmitted information takes priority over the transfer's speed.

Keywords: cybersecurity, post-quantum cryptography, Internet of Things, ZigBee

https://doi.org/10.7494/978-83-68219-80-7_11

1. Wstęp

W związku z wieloma korzyściami płynącymi z opracowywanych technologii komputerów kwantowych rozwój tej dziedziny nieustannie postępuje. Właśnie z tego powodu należy przyjrzeć się również powstałym zagrożeniom.

Najpoważniejszym ze względu na cyberbezpieczeństwo zagrożeniem jest możliwość realizacji algorytmu Shora (Shor 1997). Algorytm ten może zostać skutecznie wykorzystany do zaatakowania kryptosystemów RSA oraz El Gamal. Gwarantem bezpieczeństwa w przypadku pierwszego kryptosystemu jest złożoność obliczeniowa operacji rozkładu liczb na czynniki pierwsze. W RSA podczas generowania kluczy na jednym z etapów wykorzystywany jest iloczyn dużych liczb pierwszych p i q (Rivest i in. 1978). O ile mnożenie przez siebie dwóch czynników jest operacją stosunkowo prostą do wykonania, o tyle rozłożenie znanego iloczynu na czynniki – nazywane faktoryzacją – wymaga dużych zasobów obliczeniowych. Funkcje mające własność łatwego obliczenia i trudnego odwrócenia nazywane są funkcjami jednokierunkowymi lub zapadkowymi (Paar i Pelzl 2010). W przypadku kryptosystemu El Gamal funkcją jednokierunkową jest rozwiązanie równania $g^r \equiv x \pmod{p}$ dla r (El Gamal 1985).

Oba opisane gwaranty, po wykorzystaniu algorytmu Shora, są poddane redukcji złożoności obliczeniowej z obecnie wykładniczej do wielomianowej. Algorytm Shora pozwala z wysokim prawdopodobieństwem dokonać faktoryzacji iloczynu dużych liczb pierwszych – stosowanego w RSA – oraz, w przypadku kryptosystemu El Gamal, odzyskać z klucza publicznego element r .

Za pomocą algorytmu Shora można rozwiązać powyższe problemy dzięki wykorzystaniu funkcji pomocniczych. W przypadku RSA funkcją tą jest $f(a) = x^a \pmod{n}$, gdzie n jest iloczynem p i q , a x jest losowo wybieraną liczbą względnie pierwszą z n . Dla El Gamal funkcją tą jest $f(a, b) = g^a x^{(-b)} \pmod{p-1}$. Obydwie funkcje są okresowe, a odnalezienie ich okresów odbywa się za pomocą kwantowej transformaty Fouriera. Znajomość okresu potrzebna jest do przeprowadzenia dalszych obliczeń wyłaniających prawdopodobną liczbę rozwiązującą problem (Shor 1997).

W odpowiedzi na nadchodzące zagrożenia w celu zabezpieczenia dotychczas stosowanych kryptosystemów tworzone i wprowadzane są algorytmy kryptografii postkwantowej, które kosztem wymogu większej mocy obliczeniowej mają za zadanie zapewnić bezpieczeństwo w obliczu możliwości realizacji algorytmu Shora.

W rozwiązaniu przedstawionym w niniejszej publikacji zastosowano mechanizm ML-KEM-512, potocznie zwany Kyber512, który został wyłoniony jako zwycięzca w kwietniu 2024 roku w ogłoszonym w roku 2016 konkursie NIST Post-Quantum Cryptography Standardization. Konkurs został powołany w celu ustandaryzowania kryptosystemów kryptografii postkwantowej oraz zbadania bezpieczeństwa nominowanych do konkursu algorytmów. Rywalizacja została podzielona na cztery rundy, w wyniku czego

wyłoniono łącznie pięć algorytmów w dwóch kategoriach – KEM: CRYSTALS-KYBER, HQC oraz DSA: CRYSTALS-DILITHIUM, FALCON, SPHINCS+.

KEM – *Key Encapsulation Mechanism* – jest podejściem, w którym jedna ze stron komunikacji jest odpowiedzialna za wygenerowanie losowego klucza symetrycznego i przekazanie go drugiej w postaci kapsułki kanałem publicznym. Oznacza to utworzenie szyfrogramu za pomocą klucza publicznego odbiorcy. Istnieje również inne niż zastosowane w implementacji ML-KEM-512 rozwiązanie z rodziny KEM – RSA-KEM. Odmiernym podejściem jest mechanizm KEX, w którym każda ze stron przyczynia się w pewnym stopniu do uzgodnienia klucza symetrycznego. W wyniku jawnej wymiany odpowiednio obliczonych składników klucza kanałem publicznym tworzony jest współdzielony przez obie strony wymiany sekret. Przykładem tego podejścia jest protokół Diffiego–Hellmana.

Rodzina kryptosystemów ML-KEM została ustandaryzowana w dokumencie FIPS 203 (NIST 2024). Należą do niej ML-KEM-512, ML-KEM-768 oraz ML-KEM-1024. Każda z wersji ML-KEM wykorzystuje ten sam algorytm z innymi wartościami parametrów, których rekomendacje również umieszczono w FIPS 203. Wybrany przez autorów ML-KEM-512 zapewnia bezpieczeństwo bitowe na poziomie przynajmniej 128 bitów, a rekomendowanymi wartościami parametrów są $n = 256$, $q = 3329$, $k = 2$, $\eta_1 = 3$, $\eta_2 = 2$, $d_1 = 10$, $d_2 = 4$ (NIST 2024).

Algorytm ML-KEM jest oparty na problemie MLWE (*Module Learning with Errors*), który z kolei jest oparty na problemach kratowych (Avanzi i in. 2021). Wykorzystanie problemów opartych na kratkach gwarantuje bezpieczeństwo przed atakami z pomocą komputera kwantowego, ponieważ aktualnie nie istnieje kwantowy algorytm zdolny do przyspieszenia rozwiązywania tego typu problemów.

W ML-KEM klucz symetryczny jest kapsułkowany kluczem publicznym z dodatkiem wektora błędu o niewielkich wartościach. Zadaniem odbiorcy kapsułkowanego klucza jest jego odszyfrowanie. Ze względu na wektor błędu proces ten obciążony jest prawdopodobieństwem błędu, które w przypadku poszczególnych ML-KEM ma następujące wartości: ML-KEM-512 – $2^{-138.8\%}$, ML-KEM-768 – $2^{-164.8\%}$ i ML-KEM-1024 – $2^{-1748.8\%}$.

Celem prowadzonych badań jest analiza narzutów czasowych i wydajnościowych rozwiązań postkwantowych w istniejącej infrastrukturze IoT na podstawie powszechnie stosowanego protokołu ZigBee (ZigBee Alliance 2015). Można uznać, że rozwiązanie rozszerzające stosowany już protokół oraz możliwe do wdrożenia w aktualnie stosowanych urządzeniach stanowi kluczowy czynnik związany z jego rozpowszechnieniem.

Wybrany protokół ZigBee umożliwia tworzenie sieci typu *mesh* oraz zapewnia energooszczędność, co czyni go idealnym rozwiązaniem w zastosowaniach, gdzie różnorodne urządzenia zasilane bateriami o niskich pojemnościach mają komunikować

się w zróżnicowanych środowiskach. Jest w domach, w służbie zdrowia, a także w przemyśle. Z powodu wielkoskalowych aplikacji IoT wymaga odpowiedniego, długoterminowego zabezpieczenia.

Z technicznego punktu widzenia ZigBee opiera się na standardzie IEEE 802.15.4, który definiuje dwie najniższe warstwy modelu komunikacyjnego: warstwę fizyczną (PHY) – odpowiedzialną za transmisję danych w medium – oraz warstwę kontroli dostępu do medium (MAC) – odpowiedzialną za kontrolę dostępu do kanału radiowego za pomocą mechanizmu CSMA-CA lub LBT, a także synchronizację i zapewnienie niezawodnego mechanizmu transmisji. ZigBee rozszerza ten model o dodatkowe poziomy – warstwę sieciową (NWK) oraz warstwę aplikacji (APL).

W ramach sieci ZigBee urządzenia mogą pracować w jednym z trzech trybów, które określają ich rolę w topologii sieci:

- **Koordynator** – węzeł centralny lub nadrzędny w stosunku do innych węzłów w sieci, odpowiedzialny za inicjalizację i zarządzanie siecią. Przechowuje listę powiązanych urządzeń, obsługuje procesy asocjacji, dysocjacji oraz skanowania dostępnych urządzeń.
- **Router** – urządzenie pośrednie, odpowiedzialne za przekazywanie pakietów między urządzeniami końcowymi lub między urządzeniem końcowym a koordynatorem.
- **Urządzenie końcowe** – proste urządzenie, odpowiedzialne za monitorowanie i zbieranie danych lub podejmowanie określonych działań na podstawie poleceń użytkownika.

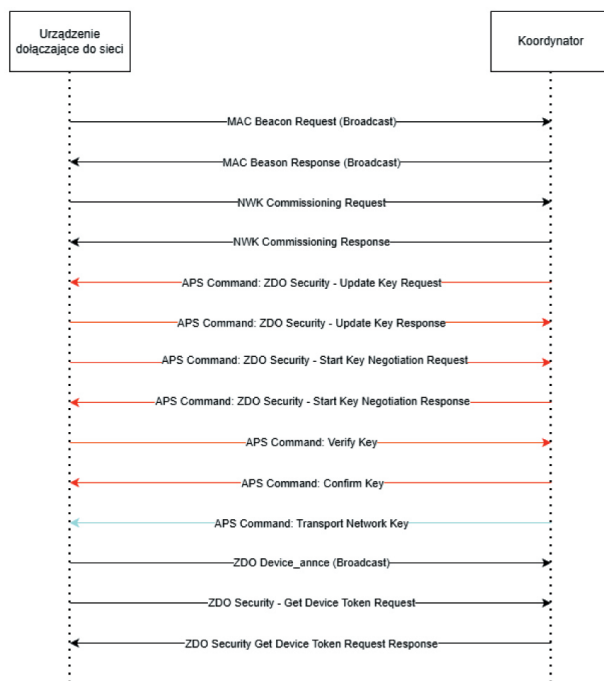
Mimo że standard ZigBee zapewnia szyfrowaną komunikację, zastosowane w nim protokoły uzgadniania klucza – ECDHE oraz SPEKE – nie są odporne na ataki z wykorzystaniem komputerów kwantowych. Do szyfrowania danych ZigBee stosuje algorytm AES-128.

W kwestii bezpieczeństwa protokołów ZigBee opiera się na dwóch podstawowych rodzajach kluczy kryptograficznych: *link key* oraz *network key*. Klucze typu *link key* służą do szyfrowania ruchu w warstwie aplikacyjnej (APS), w szczególności komunikacji związanej z transportem klucza m.in. podczas procesu dołączania urządzenia do sieci. Z kolei klucze typu *network key* służą do szyfrowania komunikacji w ramach całej sieci ZigBee.

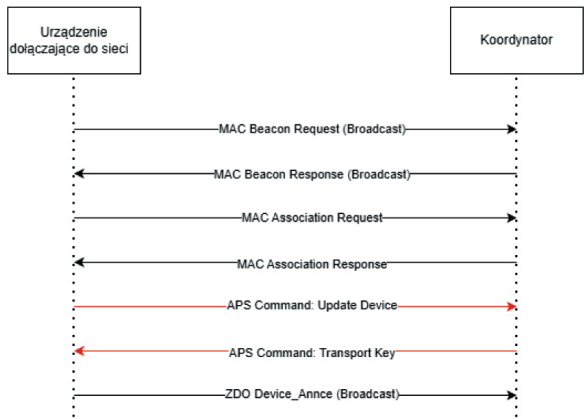
Proces dołączania w trybie *dynamic key negotiation* urządzenia do sieci ZigBee składa się z kilku etapów, rozpoczynających się od wykrycia sieci i uzyskania jej podstawowych parametrów za pomocą zapytania *Beacon Request*. Koordynator odpowiada na zapytanie komunikatem *Beacon Response*, a następnie urządzenie inicjuje ustanawianie połączenia na poziomie warstwy sieciowej dzięki wymianie komunikatów *NWK Commissioning Request/Response*. Po zestawieniu połączenia urządzenie i koordynator

przechodzą do fazy uzgadniania klucza – *link key* – na poziomie warstwy aplikacji (APS), w której wykorzystywane są komunikaty *Update Key Request/Response*, *Start Key Negotiation Request/Response* oraz proces weryfikacji i potwierdzania klucza. Po pomyślnym zakończeniu negocjacji klucza koordynator przesyła urządzeniu klucz sieciowy (*network key*), który umożliwia szyfrowanie dalszej komunikacji sieciowej. Ostatnim etapem procesu dołączenia jest wymiana komunikatów odpowiedzialnych za finalizację rejestracji urządzenia w sieci.

Dynamic Key Negotiation (rys. 1) to nowszy sposób na wymianę klucza *link key*. W przypadku gdy jedno z urządzeń nie wspiera tego rozwiązania lub urządzenia ZigBee działają w sieci w różnych wersjach specyfikacji, wymiana kluczy następuje według starej statycznej metody (klucze *pre-configured*). Natomiast gdy klucze preinstalowane nie są ustawione, kluczem domyślnym jest łańcuch *ZigBeeAlliance09*, co wpływa znacząco na zmniejszenie bezpieczeństwa protokołu. Proces dołączania w podanej statycznej metodzie polega na analogicznym uzyskaniu podstawowych parametrów sieci oraz na procesie asocjacji. Po zakończeniu asocjacji koordynator wysyła klucz sieciowy przy użyciu preinstalowanego klucza *link key* oraz wysyła komunikat *Device Announce*, który działa w ten sam sposób jak w przypadku wcześniejszego sposobu *dynamic* (rys. 2).



Rys. 1. Schemat wymiany komunikatów między urządzeniem a koordynatorem w procesie dołączania do sieci ZigBee (na podstawie specyfikacji ZigBee R23) (Connectivity Standards Alliance 2023)



Rys. 2. Schemat wymiany komunikatów między urządzeniem a koordynatorem w procesie dołączania do sieci ZigBee (w przypadku braku obsługi wersji specyfikacji ZigBee R23) (Connectivity Standards Alliance 2023)

2. *State-of-the-art*

Prace łączące kryptografię postkwantową z internetem rzeczy (np. Sajimon i in. 2022) skupiają się z reguły na wymaganiach obliczeniowych nowych algorytmów i ich osiągnięciach w kontekście urządzeń IoT z niską mocą obliczeniową. T.M. Fernández-Caramés przedstawia szerokie spojrzenie na zagadnienie kryptografii postkwantowej w internecie rzeczy, porównując algorytmy z drugiej rundy konkursu PQC NIST i sprawdzając różne konfiguracje na kilku procesorach używanych w urządzeniach IoT (Fernández-Caramés 2020). L. Malina wraz ze współpracownikami przedstawił rezultaty implementacji różnych algorytmów postkwantowych w urządzeniach o niskiej mocy obliczeniowej (Malina i in. 2018). Badania te prowadzone były jednak przed ogłoszeniem dokumentu FIPS 203, który wyłonił CRYSTALS-Kyber i CRYSTALS-Dilithium jako najlepiej sprawdzające się algorytmy szyfrowania i podpisów cyfrowych oraz podał rekomendowane parametry dla ML-KEM. W badaniach skupiono się na aspekcie narzutów czasowych przy implementacji ML-KEM-512 w protokole ZigBee do wymiany klucza oraz w komunikacji sieciowej z szyfrowaniem AES-256 w trybie CTR na podstawie pomiarów wykonanych z pomocą funkcjonującej sieci urządzeń.

3. *Metoda*

W tym podrozdziale opisano metody badawcze, rozpoczynając od omówienia wykorzystanego sprzętu i oprogramowania, a kończąc na metodologii prowadzenia pomiarów.

3.1. Narzędzia i urządzenia

Do przeprowadzenia doświadczenia wykorzystane zostały narzędzia zigpy-cli (Open Home Foundation b.d.), esp.idf (Espressif Systems b.d.), pyenv, nrfutil, Wireshark i liboqs oraz poniższe urządzenia:

- **Bramka ZigBee 3.0 SONOFF ZBDongle-E.** Urządzenie pełniące rolę koordynatora sieci ZigBee. Umożliwia tworzenie i zarządzanie siecią ZigBee 3.0. Za jego pomocą możliwe było zestawienie sieci oraz komunikacja z urządzeniami końcowymi.
- **SparkFun Thing Plus – ESP32-C6.** Mikrokontroler z obsługą standardu ZigBee 3.0, używany jako urządzenie końcowe (*end device*). W projekcie realizował funkcje nadawcy danych w sieci ZigBee, wykorzystując stos ESP-IDF oraz przykładowe aplikacje ZigBee.
- **nRF52840 Dongle.** Moduł oparty na układzie firmy Nordic Semiconductor, wykorzystywany jako *sniffer* (monitor) ruchu ZigBee. Dzięki odpowiedniemu firmware'owi i integracji z Wiresharkiem umożliwiał przechwytywanie ramek 802.15.4 oraz analizę protokołu ZigBee w czasie rzeczywistym.
- **Komputer z systemem Kali Linux (2025.1).** Służył jako stanowisko monitorujące – zarządzał snifferem nRF52840, konfigurował jego interfejs i przeprowadzał przechwytywanie oraz analizę ruchu w Wiresharku.
- **Komputer z systemem Ubuntu 24.04.2 LTS.** Pełnił rolę terminala kontrolnego w przypadku bramki ZigBee oraz ESP32-C6. Umożliwiał programowanie ESP32-C6 za pomocą ESP-IDF, konfigurację sieci ZigBee przy użyciu biblioteki zigpy-cli, a także zbieranie i przetwarzanie danych pomiarowych.

3.2. Konfiguracja sieci bazowej

Na pierwszym etapie zestawiono standardową sieć ZigBee bez modyfikacji kryptograficznych. Do konfiguracji bramki zastosowano bibliotekę zigpy-cli w wersji 1.1.0, umożliwiającą ręczne zestawienie sieci i odczyt jej parametrów (w tym klucza sieciowego i kanału operacyjnego). Urządzenie końcowe ESP32-C6 zaprogramowano przy użyciu pakietu ESP-IDF oraz oficjalnego repozytorium esp-zigbee-sdk, wykorzystując dostarczone przykładowe implementacje węzłów końcowych ZigBee.

3.3. Monitoring ruchu sieciowego

Do analizy ruchu użyto sniffera nRF52840 oraz oprogramowania Wireshark. *Sniffer* został skonfigurowany z wykorzystaniem narzędzia *nrfutil* i odpowiedniego firmware'u. Następnie zintegrowano go z Wiresharkiem jako urządzenie przechwytyjące, umożliwiając dekodowanie ruchu ZigBee przy użyciu klucza sieciowego pozyskanego wcześniej

z konfiguratora. Przechwycony ruch posłużył do weryfikacji poprawności komunikacji oraz stanowił podstawę do pomiarów czasów transmisji i wymiany kluczy.

3.4. Procedura pomiarowa

W każdym cyklu pomiarowym inicjalizowano pełny proces komunikacyjny, obejmujący zestawienie połączenia, wymianę danych i ich deszyfrowanie. W przypadku każdego scenariusza zebrano 1000 niezależnych pomiarów czasu trwania komunikacji (od momentu wysłania pakietu do jego pełnego odbioru i rozszyfrowania). Czas mierzone na poziomie systemu operacyjnego z dokładnością do milisekundy.

3.5. Zmodyfikowana sieć ZigBee z postkwantową wymianą klucza

W odpowiedzi na rosnące wymagania w zakresie bezpieczeństwa komunikacji w sieciach IoT opracowano autorską modyfikację warstwy komunikacyjnej sieci ZigBee, wykorzystującą postkwantowy kryptosystem ML-KEM-512 (Kyber) do wymiany kluczy szyfrujących. Koordynator sieci, oparty na bibliotekach zigpy i zaimplementowany w języku Python, uruchamia sieć ZigBee oraz tworzy lokalną bazę danych `zigbee.db`, w której przechowywane są informacje o połączonych urządzeniach. Po uruchomieniu sieci następuje okres pozwalający na dołączenie nowych urządzeń końcowych. Każde z nich przesyła koordynatorowi swój klucz publiczny ML-KEM-512. Następnie koordynator generuje unikatowy klucz symetryczny (256-bitowy), który szyfruje przy użyciu otrzymanego klucza publicznego i przesyła z powrotem do urządzenia końcowego. Dalsza komunikacja między urządzeniem a koordynatorem odbywa się z dodatkową warstwą szyfrowania za pomocą algorytmu AES w trybie CTR i z użyciem wymienionego klucza symetrycznego. Ze względu na potrzeby badania komunikacja między urządzeniami zawsze była szyfrowana.

3.5.1. Implementacja klastra ZigBee

Protokół ZigBee wspiera tworzenie własnych aplikacji, nazywanych klastrami. Każdy klaster ma wersję przeznaczoną dla klienta i serwera oraz umożliwia komunikację za pomocą zdefiniowanych przez twórcę komend. Zarówno serwer, jak i klient mogą wysyłać komunikaty oraz przechowywać atrybuty – przykładem zastosowania może być domowy system ogrzewania. Piec grzewczy (będący serwerem klastra „ogrzewanie”) otrzymuje od klientów (termostatów) żądanie ogrzewania różnych miejsc za pomocą komendy `request_heating`. Za pomocą innego klastra piec może sterować zaworami w celu docieplenia tylko potrzebujących tego pomieszczeń.

Aby zintegrować mechanizm wymiany klucza, utworzono specjalny klaster ZigBee o identyfikatorze `0xfcfc`, nazwany *security cluster*. Klaster ten zawiera atrybut identyfi-

kowany przez 0x0000 (`key_version`), który informuje o wersji klucza symetrycznego, oraz trzy podstawowe komendy:

- **0xFF – `send_key`** (wykorzystywana przez koordynatora do przesyłania zakodowanego klucza AES; końcowe urządzenie odpowiada tą samą komendą w formie potwierdzenia odbioru (`ACK`));
- **0xFC – `request_key`** (używana do zażądania klucza publicznego od urządzenia końcowego w przypadku jego braku po stronie koordynatora);
- **0xFA – `send_pub_key`** (wysyłana przez klienta w odpowiedzi na żądanie `request_key`, zawiera klucz publiczny ML-KEM-512).

3.5.2. Ruch w sieci

Proces wymiany zaszyfrowanego klucza symetrycznego, składającego się z 768 bajtów, ze względu na specyfikę działania protokołu ZigBee jest realizowany z wykorzystaniem fragmentaryzacji na 12 pakietów po 64 bajty, które następnie przesyłane są do urządzenia końcowego. Po otrzymaniu wszystkich fragmentów szyfrogramu urządzenie końcowe wysyła potwierdzenie otrzymania całej wiadomości.

Procedura przesyłania klucza publicznego do koordynatora jest rozpoczynana przez koordynator wysłaniem do urządzenia końcowego komendy 0xFC. W odpowiedzi urządzenie końcowe przesyła swój klucz publiczny. Składa się on z 800 bajtów, co wymusza fragmentaryzację – jest przesyłany za pomocą 13 pakietów, pierwsze 12 po 64 bajty oraz ostatni 32-bajtowy. W celu symulowania komunikacji danych urządzenie końcowe (ESP32-C6) generuje pakiety zawierające losowe dane, które następnie przesyła do koordynatora.

3.6. Metodologia testów

W ramach testów wydajnościowych przeanalizowano kluczowe komponenty zmodyfikowanego protokołu komunikacyjnego. Przedmiotem pomiarów były czas generowania i transmisji 64-bajtowego szyfrogramu zawierającego klucz symetryczny oraz opóźnienie związane z procesem jego fragmentacji i przesyłania przez sieć ZigBee. Zostało przeprowadzonych 1000 niezależnych pomiarów czasu transmisji szyfrogramu.

Został również zmierzony narzut czasowy spowodowany przez implementowaną warstwę szyfrowania AES-256 (tryb CTR) w ruchu sieciowym, a także przeprowadzono po 1000 niezależnych pomiarów czasów przesłania 64-bajtowego pakietu ZigBee w sieci bez modyfikacji z użyciem szyfrowania AES-256, co pozwoliło oszacować wpływ warstwy szyfrowania AES na opóźnienie transmisji.

Dodatkowo zmierzono czasy wykonania funkcji odpowiedzialnych za szyfrowanie pakietów za pomocą AES. Testy wykonano na 64-bajtowych ciągach losowych danych.

4. Dyskusja

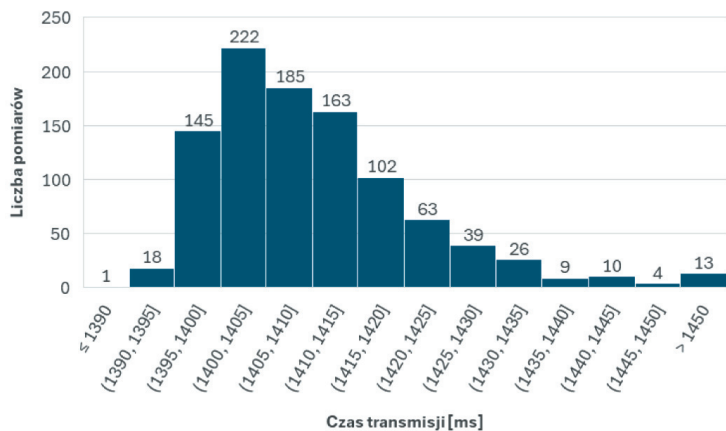
Poniżej omówiono wyniki pomiarów czasu wymiany klucza za pomocą ML-KEM-512 oraz porównano zmierzone czasy przesyłu danych w standardowej sieci ZigBee z wynikami otrzymanymi w sieci implementującej opisaną w artykule rozwiązanie szyfrowania ruchu za pomocą AES-256.

4.1. Wydajność wymiany klucza symetrycznego

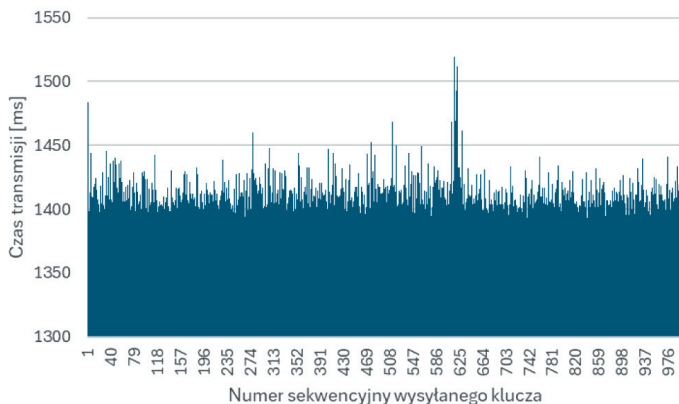
Po przeprowadzeniu wyżej opisanych testów prędkości przesyłania klucza symetrycznego otrzymano wyniki przedstawione na rysunkach 3–5.

Jak pokazano na rysunku 3, najwięcej pomiarów (222) mieści się w przedziale [1400, 1405) ms. Większość pomiarów skupia się w zakresie od 1395 ms do 1420 ms. Rozkład danych jest prawostronnie skośny, występuje wyraźny „ogon” po prawej stronie – są też pomiary z czasami dochodzącymi nawet do ponad 1450 ms, ale nieliczne. Sugeruje to obecność sporadycznych opóźnień, które mogą wynikać z obciążenia sieci, przeciążeń sprzętowych lub zakłóceń transmisji. Wartości skrajnie niskie (<1390 ms) i skrajnie wysokie (>1450 ms) są rzadkie (łącznie 14 pomiarów). Wskazuje to na relatywną stabilność działania protokołu (zdarzają się przypadki wyjątkowe). Rysunek 4 przedstawia graficznie czasy wymiany kolejnych kluczy. Obserwowany w okolicy 625 pomiaru wzrost czasu wymiany jest artefaktem działania sieci w rzeczywistym środowisku.

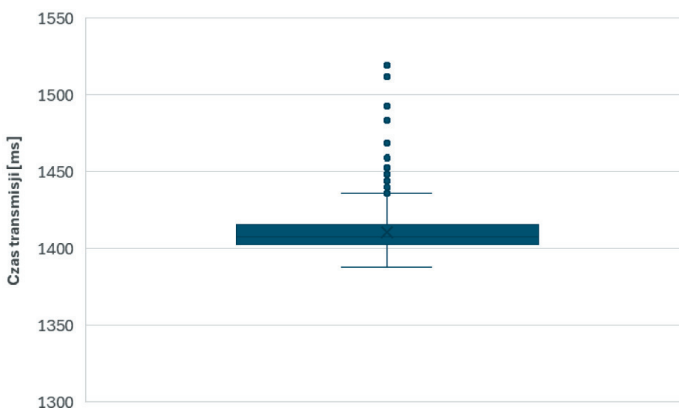
Na podstawie 1000 pomiarów czasu transmisji klucza symetrycznego można stwierdzić, że średnia wartość wyniosła 1410,51 ms, natomiast mediana była nieco niższa i wyniosła 1407,65 ms. Odchylenie standardowe próbki na poziomie 12,97 ms wskazuje na niewielką zmienność wyników, a niski błąd standardowy średniej wynosi 0,41% (rys. 5).



Rys. 3. Rozkład czasów transmisji klucza symetrycznego



Rys. 4. Graficzna reprezentacja czasów przesłania kolejnych kluczy

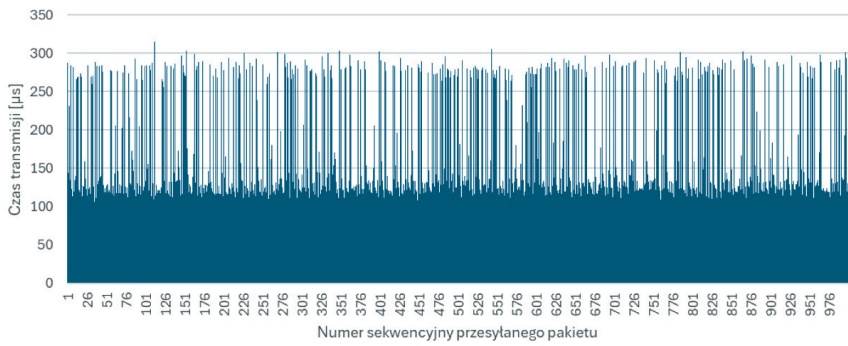


Rys. 5. Wykres pudełkowy z wąsami czasów transmisji klucza symetrycznego

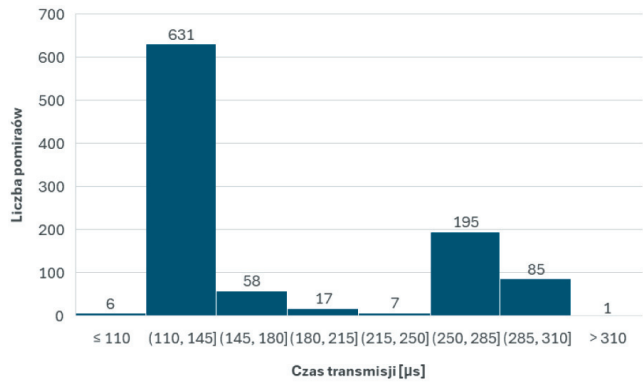
4.2. Wyniki pomiarów w sieci podstawowej

Pomiary czasu transmisji pakietów w bazowej sieci oraz w sieci korzystającej z proponowanego rozwiązania zostały przeprowadzone w seriach 1000 przesyłanych pakietów o jednakowej długości 64 bajtów i losowej zawartości. Między transmisją kolejnych pakietów został wprowadzony odstęp 100 ms, który nie jest uwzględniany w pomiarze. Rysunek 6 ilustruje graficzną reprezentację czasu transmisji w sieci bazowej na podstawie 1000 pomiarów. Widoczne są dwie przeważające wartości – około 125 μ s i 275 μ s. Tę obserwację potwierdza histogram przedstawiony na rysunku 7.

Można przypuszczać, że obserwowany podział jest skutkiem konieczności retransmisji niektórych pakietów, co powoduje około dwukrotny wzrost czasu przesyłu. Wydłużenie czasu transmisji wydaje się związane z koniecznością odczekania na potwierdzenie odbioru pakietu ze strony odbiorcy oraz podjęciem decyzji o retransmisji.

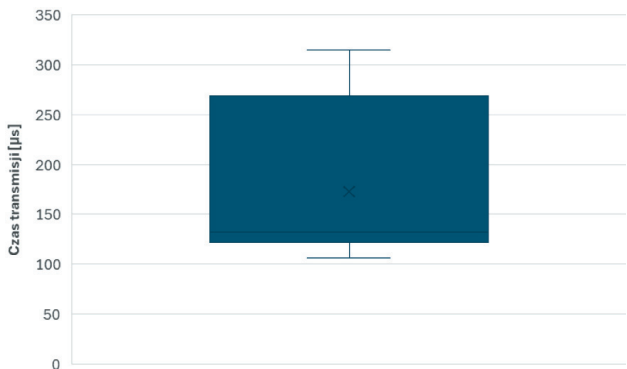


Rys. 6. Graficzna reprezentacja czasów przesłania pakietów z serii w sieci, która nie korzysta z szyfrowania



Rys. 7. Histogram czasów przesyłu pakietu bez dodatkowej warstwy szyfrowania

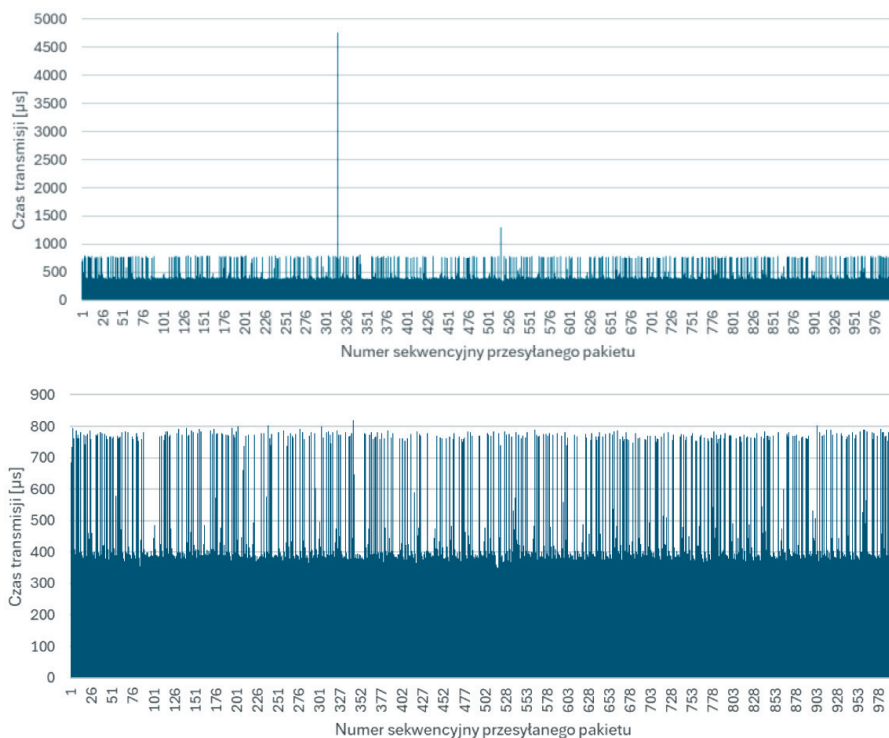
Maksymalnym obserwowanym czasem przesyłu było 315 μs, natomiast najkrótszy czas wynosił 106 μs. Średni czas wynosił 172,69 μs, a mediana czasu przesyłania ma wartość 132 μs – te dane ilustruje rysunek 8.



Rys. 8. Wykres pudełkowy z wąsami czasów przesyłu pakietu bez użycia szyfrowania AES-256

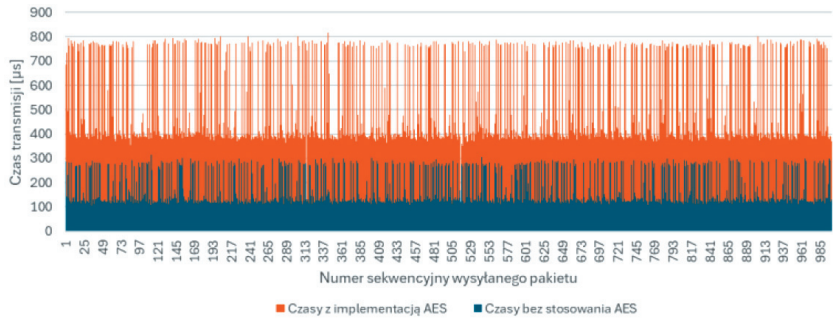
4.3. Wyniki pomiarów w przypadku sieci implementującej szyfrowanie AES-256

Wyniki pomiarów czasu przesyłu kolejnych pakietów w sieci implementującej szyfrowanie za pomocą AES ilustruje rysunek 9. W pomiarach pojawiły się dwa wyniki znacznie odbiegające od reszty – uznane zostały one za wyjątki. Tak samo jak na rysunku 6, również na rysunku 9 można zaobserwować dwie wyraźne serie wartości czasów.

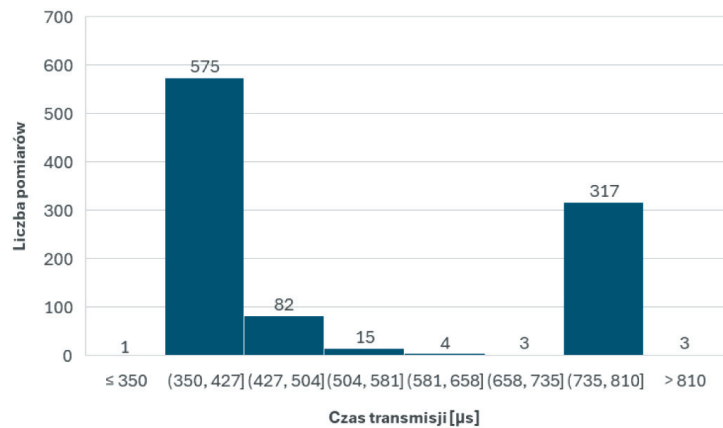


Rys. 9. Graficzna reprezentacja czasów przesyłu pakietów z serii w sieci, która implementuje szyfrowanie za pomocą AES. Górny wykres przedstawia dane bez modyfikacji, natomiast na dolnym wykresie ilustrowane są czasy bez uwzględnienia dwóch odstających pomiarów

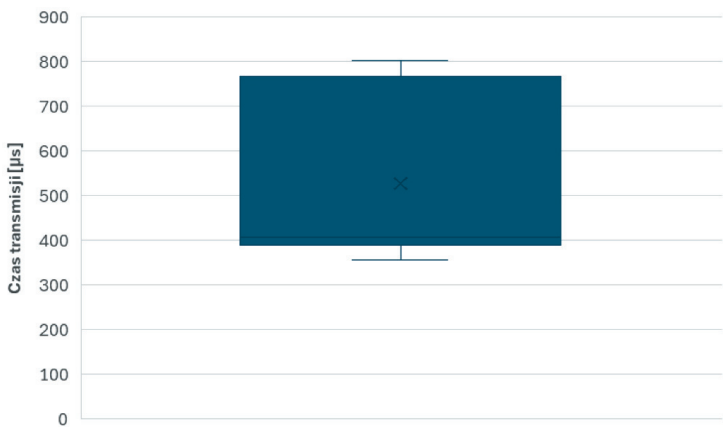
Rysunek 10 ilustruje porównanie serii pomiarów w przypadku bazowej sieci oraz po implementacji szyfrowania AES. Histogram przedstawiony na rysunku 11 również odzwierciedla występujące w poprzednich pomiarach dwie wyraźne serie. Najdłuższy wynik pomiaru liczył 818 μs, najkrótszy natomiast wyniósł 349 μs – jest to wartość wyższa od maksymalnego odnotowanego czasu w sieci bazowej. Średni czas przesyłu z szyfrowaniem wynosił w przybliżeniu 522,36 μs, a wartość mediany dla tej serii to 406 μs. Te dane przedstawia rysunek 12.



Rys. 10. Wykres porównujący czasy transmisji bez szyfrowania (niebieski) oraz po implementacji AES (pomarańczowy).
Odrzucone zostały pomiary drastycznie różniące się od pozostałych



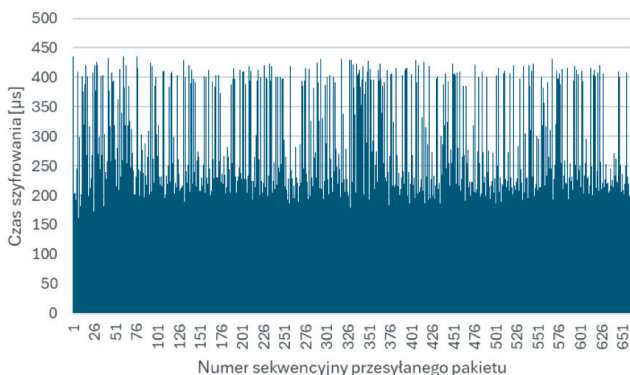
Rys. 11. Histogram czasów przesyłu pakietu z użyciem dodatkowej warstwy szyfrowania



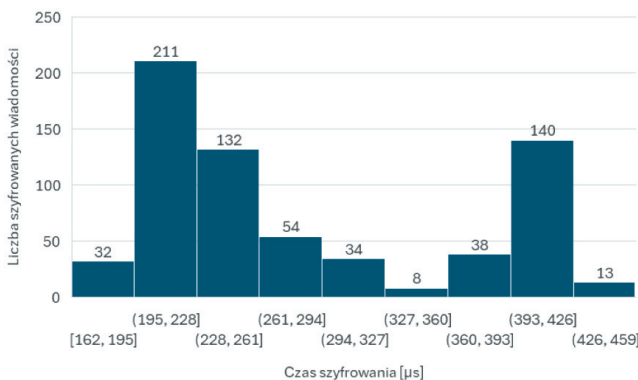
Rys. 12. Wykres pudełkowy z wąsami czasów przesyłu pakietu z użyciem dodatkowej warstwy szyfrowania, nieuwzględniający pomiarów drastycznie różniących się od pozostałych

4.4. Wyniki pomiarów czasów szyfrowania AES

W poszukiwaniu przyczyn zmienności czasów transmisji zostało przeprowadzonych 660 pomiarów czasów wykonania funkcji szyfrującej. Wyniki pomiarów są ilustrowane na rysunku 13. Ponownie obserwowany jest podział na dwie serie czasów. Wartości wyższej serii stanowią mniej więcej dwukrotność tych z serii niższej. Ten trend odzwierciedlony jest na histogramie przedstawionym na rysunku 14. Podobieństwo tych wyników do pomiarów transmisji prowadzi do wniosku, że zmienności rezultatów pomiarów są powodowane funkcjonowaniem sprzętu.



Rys. 13. Wykres czasów szyfrowania kolejnych pakietów



Rys. 14. Histogram czasów szyfrowania pakietów

4.5. Interpretacja wyników

Wyniki pomiarów czasów przesłania szyfrogramu klucza AES sugerują znaczący wpływ wdrożenia wymiany klucza na czas inicjalizacji infrastruktury. Pokazują one narzut czasowy rzędu około 1,4 s w przypadku każdego urządzenia będącego częścią sieci, który znacząco ogranicza skalowalność proponowanego w artykule rozwiązania.

Analiza wyników przesyłu pakietów zarówno w sieci bez szyfrowania, jak i w sieci implementującej szyfrowanie wykazuje obecność dwóch dominujących przedziałów czasowych, odpowiednio: od 110 μ s do 145 μ s i od 250 μ s do 285 μ s w sieci bez szyfrowania oraz od 350 μ s do 427 μ s i od 735 μ s do 810 μ s w sieci z dodatkową warstwą szyfrującą komunikację. Ten trend jest również widoczny w pomiarach czasów samego szyfrowania AES – odpowiednio od 195 μ s do 261 μ s oraz od 393 μ s do 426 μ s. Na podstawie widocznej zależności między pomiarami można wnioskować, że zmienność czasów transmisji jest zależna od funkcjonowania sprzętu użytego w badaniach – konkretnie ESP32-C6. Prawdopodobnie ten fakt związany jest z ograniczonymi zasobami pamięciowymi oraz obliczeniowymi urządzenia lub mechanizmami zarządzania tymi zasobami. Analiza funkcjonowania rozwiązania w konfiguracji różniącej się od prezentowanej w artykule może pomóc wyjaśnić tę obserwację. Porównując czasy z głównych grup widocznych na histogramach z rysunków 7 i 11, można zobaczyć, że przedział czasu obejmujący większość niższych wyników w sieci bez szyfrowania – (110,145] μ s – drastycznie wzrasta przy implementacji AES – do (350,427] μ s. Ten przedział czasu jest znacznie wyższy od maksymalnego wyniku pomiarów wykonanych w przypadku bazowej sieci, który wynosił 315 μ s. Druga grupa, o dłuższych czasach transmisji, obserwowana na histogramach, przeniosła się z przedziału (250,310] do przedziału (735,810] w przypadku sieci implementującej przedstawiane rozwiązanie. Natomiast wzrost mierzonego czasu w przypadku grupy z niższymi wartościami sięga około 261 μ s, czyli około 205%. W wyższej grupie ta zmiana wynosiła 492,5 μ s, czyli w przybliżeniu 176%. Mediany czasów wynosiły 132 μ s w serii niestosującej szyfrowania oraz 406 μ s w serii pomiarów z wykorzystaniem opisanego rozwiązania. Oznacza to różnicę wynoszącą 274 μ s, co przekłada się na 207% wzrostu czasu przesłania pakietu.

5. Wnioski

Bezpieczeństwo szyfrowania uzależnione jest od skutecznej wymiany klucza symetrycznego. W zaproponowanym rozwiązaniu proces ten realizowany jest z wykorzystaniem algorytmu ML-KEM-512 odpornego na ataki i komputerów kwantowych.

Poziom bezpieczeństwa standardowego ZigBee a poziom bezpieczeństwa proponowanego rozwiązania

Rozwiązania kryptograficzne w standardowym protokole ZigBee wykorzystują protokół ECDHE, oparty na krzywej Curve25519 i zapewniający 128 bitów bezpieczeństwa. W przypadku ataków za pomocą komputera kwantowego bezpieczeństwo bitowe kombinacji ECDH + AES256 praktycznie spada do zera, ze względu na całkowite złamanie protokołu ECDH algorytmem Shora. Proponowane rozwiązanie imple-

mentuje algorytm ML-KEM-512 o 128 bitach bezpieczeństwa (NIST 2024) oraz o udowodnionej odporności na znane techniki kryptoanalizy z użyciem komputera kwantowego. Jest on stosowany podczas wymiany klucza AES-256 gwarantującego 256 bitów bezpieczeństwa bez udziału komputera kwantowego. Algorytm Grovera pozwala jednak na redukcję bezpieczeństwa do 128 bitów (Grover 1996). Oznacza to, że jako całość proponowane rozwiązanie zapewnia bezpieczeństwo na poziomie 128 bitów. Uzyskany poziom jest akceptowany przez aktualne standardy.

Wprowadzenie takiego mechanizmu wiąże się jednak z dodatkowymi wymaganiami dotyczącymi infrastruktury systemu. W szczególności istotnym czynnikiem ograniczającym jest rozmiar generowanego szyfrogramu, który w przypadku ML-KEM-512 wynosi 768 bajtów. Może to stanowić istotne obciążenie dla sieci o ograniczonej przepustowości, zwłaszcza w środowiskach o dużej liczbie urządzeń końcowych. Czas inicjalizacji rośnie w tym przypadku liniowo wraz z liczbą urządzeń, ponieważ koordynator musi przeprowadzić indywidualną wymianę klucza z każdym z nich.

Zastosowanie wzmocnionego szyfrowania z wykorzystaniem algorytmu AES-256, pomimo dostępności sprzętowych akceleratorów wspierających jego działanie, wiąże się ze wzrostem zapotrzebowania na zasoby systemowe. Może to negatywnie wpłynąć na działanie systemu w przypadku urządzeń o ograniczonej mocy obliczeniowej, powodując spadek efektywności komunikacji.

Z przeprowadzonych analiz wynika, że zaproponowane rozwiązanie nie jest optymalne w zastosowaniach wymagających obsługi dużej liczby urządzeń końcowych generujących intensywny ruch sieciowy. Może ono natomiast być przydatne w systemach o ograniczonej skali lub w sieciach złożonych z urządzeń dysponujących odpowiednią mocą obliczeniową. Rozwiązanie to może być szczególnie skutecznie stosowane w komunikacji pomiędzy wybranymi węzłami, w których przypadku zapewnienie wysokiego poziomu bezpieczeństwa transmisji danych ma kluczowe znaczenie. Ze względu na narzuty związane z inicjalizacją sieci koordynatorem powinno być urządzenie o wysokiej dostępności. Pozwoli to ograniczyć występowanie zdarzeń powodujących, że cała sieć musi zostać zbudowana od nowa.

Literatura

- Avanzi R., Bos J., Ducas L., Kiltz E., Lepoint T., Lyubashevsky V., Schanck J.M. et al., 2021, *CRYSTALS-Kyber. Algorithm specifications and supporting documentation (version 3.02)*. <https://pq-crystals.org/kyber/data/kyber-specification-round3-20210804.pdf> [dostęp: 13 maja 2025].
- Connectivity Standards Alliance, 2023, *ZigBee Specification – Revision 23*. <https://csa-iot.org/wp-content/uploads/2023/04/05-3474-23-csg-zigbee-specification-compressed.pdf> [dostęp: 13 maja 2025].

- El Gamal, T., 1985, *A public key cryptosystem and a signature scheme based on discrete logarithms*, [w:] Blakley G.R., Chaum D. (eds.), *Advances in cryptology. Proceedings of CRYPTO 84*, Lecture Notes in Computer Science, vol. 196, s. 10–18, Springer, Berlin–Heidelberg–New York–Tokyo.
- Espressif Systems, *esp-idf*, b.d. <https://github.com/espressif/esp-idf> [dostęp: 11 maja 2025].
- Fernández-Caramés T.M., 2020, *From pre-quantum to post-quantum IoT security: A survey on quantum-resistant cryptosystems for the Internet of Things*, IEEE Internet of Things Journal, vol. 7, no. 7, s. 6457–6480. <https://doi.org/10.1109/JIOT.2019.2958788>.
- Grover L.K., 1996, *A fast quantum mechanical algorithm for database search*, [w:] *STOC '96. Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, Association for Computing Machinery, New York, USA, s. 212–219. <https://doi.org/10.1145/237814.237866>.
- Malina L., Popelova L., Dzurenda P., Hajny J., Martinasek Z., 2018, *On feasibility of post-quantum cryptography on small devices*, IFAC-PapersOnLine, vol. 51, s. 462–467. <https://doi.org/10.1016/j.ifacol.2018.07.104>.
- NIST (National Institute of Standards and Technology), 2024, *Module-Lattice-Based Key-Encapsulation Mechanism Standard*, FIPS 203. Federal Information Processing Standards Publication. <https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.203.pdf> [dostęp: 13 maja 2025].
- Open Home Foundation, b.d., *zigpy*. <https://github.com/zigpy/zigpy> [dostęp: 11.05.2025].
- Paar C., Pelzl J., 2010, *The RSA cryptosystem*, [w:] Paar C., Pelzl J., *Understanding cryptography: A Textbook for students and practitioners*, Springer, Heidelberg–Dordrecht–London–New York.
- Rivest R.L., Shamir A., Adleman L., 1978, *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM, vol. 21, no. 2, s. 120–126. <https://doi.org/10.1145/359340.359342>.
- Sajimon P.C., Jain K., Krishnan P., 2022, *Analysis of post-quantum cryptography for Internet of Things*, [w:] 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, Indie, s. 387–394. <https://doi.org/10.1109/ICICCS53718.2022.9787987>.
- Shor P.W., 1997, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Journal on Computing, vol. 26, no. 5, s. 1484–1509. <https://doi.org/10.48550/arXiv.quant-ph/9508027>.
- ZigBee Alliance, 2015, *ZigBee Specification*. <https://zigbeealliance.org/wp-content/uploads/2019/11/docs-05-3474-21-0csg-zigbee-specification.pdf> [dostęp: 3.12.20125].

Niestandardowe możliwości zastosowania technologii horyzontalnego przewiertu sterowanego – HDD

Jan Ziaja , Robert Rado 

Akademia Górniczo-Hutnicza, Wydział Wiertnictwa, Nafty i Gazu, Kraków

Streszczenie: W artykule przedstawiono krótką historię powstania i rozwoju technologii HDD na świecie i w Polsce. Jej zalety, wady oraz ograniczenia techniczne, a także rekordowe osiągnięcia. Jednak głównym celem było opisanie niestandardowych możliwości zastosowania tej technologii dla różnych aplikacji. Na podstawie zrealizowanych projektów przemysłowych przedstawiono techniczne możliwości wykonywania otworów za pomocą technologii HDD. Opisano jej wykorzystanie do wykonywania drenażu i oczyszczania gruntów, uszczelniania składowisk odpadów, budowy horyzontalnych otworowych wymienników ciepła, eksploatacji podziemnych złóż wód oraz przekraczania linii brzegowej w celu instalacji kabli energetycznych z morskich farm wiatrowych.

Słowa kluczowe: technologie bezwykopowe, horyzontalny przewiert sterowany, geoenergetyka, energetyka wiatrowa, ochrona środowiska

UNCONVENTIONAL APPLICATIONS OF HORIZONTAL DIRECTIONAL DRILLING (HDD)

Abstract: This article presents a brief history of the emergence and development of HDD technology globally and in Poland. Its advantages, disadvantages, and technical limitations, along with record-breaking achievements, are presented. However, the primary goal is to describe the unconventional applications of this technology for various applications. The technical feasibility of drilling holes using HDD technology is presented based on completed industrial projects. The application of HDD technology for drainage and soil remediation, landfill sealing, horizontal borehole heat exchangers, underground water extraction, and shoreline crossings for the installation of offshore wind farm power cables is described.

Keywords: trenchless technologies, horizontal directional drilling, geoenergetics, wind energy, environmental protection

https://doi.org/10.7494/978-83-68219-80-7_12

1. Wprowadzenie

Technologia HDD jest znana na świecie od ponad 50 lat. Jednak jej aplikacje ciągle się poszerzają i są wykorzystywane w coraz to nowych branżach przemysłu. HDD to już nie tylko liniowa instalacja gazociągów i wodociągów pod naturalnymi przeszkodami terenowymi. Ale ma już też swoje pierwsze aplikacje w budowie podziemnych ujęć wody pitnej. Służy do remediacji, neutralizacji i oczyszczania zdegradowanych gruntów, a także wzmacniania podłoża gruntowego i jego uszczelniania. W ostatnich latach jest również podstawową metodą przekraczania morskich linii brzegowych w celu instalacji kabli energetycznych doprowadzających energię elektryczną z morskich farm wiatrowych na ląd. Zrównoważony rozwój oraz wszechobecna transformacja energetyczna pozwalają na znajdowanie nowych zastosowań technologii HDD.

2. Historia technologii HDD

Można jednoznacznie stwierdzić, że technologia HDD bezpośrednio wywodzi się z przemysłu naftowego, a konkretnie: naftowych wierceń kierunkowych. Skrót HDD pochodzi z języka angielskiego i oznacza *horizontal directional drilling* (Bennett i Ariaratnam 2017). Jej polska nazwa, preferowana przez autora niniejszego artykułu, to horyzontalny przewiert sterowany – HPS. Technika HDD, jak każda z przełomowych metod bezwykopowych, ma ciekawą genezę. Titan Contractors to firma powstała w 1964 roku w Stanach Zjednoczonych z inicjatywy twórcy technologii HDD Martina Cherringtona, który uważany jest za pioniera oraz ikonę rynku technologii bezwykopowych. Już w 1971 roku firma zrealizowała pierwsze przekroczenie rzeki Pajaro w Kalifornii metodą zbliżoną do HDD. Za pomocą tej technologii zainstalowany został 4-calowy stalowy gazociąg o długości 180 m (Belani i in. 2014).

W kolejnych latach właśnie w USA zaczęły pojawiać się na rynku wiertnice z przewodem wiertniczym zaadaptowanym z wiertnictwa naftowego, co umożliwiło stopniowe zwiększanie długości wierceń, osiągając dystans do 500 m. W latach 1971–1979 Martin Cherrington wraz z firmą Titan Contractors zrealizował aż 40 instalacji (Belani i in. 2014). Sztandarowym projektem była instalacja rurociągu pod kanałem w Houston. Zrealizowany z sukcesem w 1979 roku rurociąg o średnicy 40 cali (1016 mm) przyniósł firmie rozgłos jako pierwszy przewiert o tak dużej średnicy (Najafi i in. 2021). W 1984 roku w Sacramento Cherrington założył firmę Cherrington Corp., w której akronim HDD został spopularyzowany, zyskał status oficjalnego terminu branżowego i został przyjęty do pełnoprawnej funkcjonalności. Dopiero w drugiej połowie lat 80. XX wieku na rynek wierceń kierunkowych zaczęły wkraczać inne firmy – takie jak: Michels, Laney, Canadian Horizontal Drilling czy w Europie Horizontal Drilling International – które w kolejnych latach stały się jego liderami (Willoughby 2005).

Początkowo wiercenia prowadzono wyłącznie w terenach nieurbanizowanych, co eliminowało problemy z istniejącą już infrastrukturą podziemną. Jednak wraz ze wzrostem

zapotrzebowania na realizację projektów w obszarach miejskich rozpoczęto poszukiwania bezpiecznego systemu nawigacji, który umożliwiałby skuteczne omijanie istniejących już instalacji podziemnych. U schyłku lat 80. XX wieku rozpoczęto prace nad skonstruowaniem innowacyjnego magnetycznego wgłębnego urządzenia pomiarowego. Działania te wdrożyła firma Tensor, która zrewolucjonizowała rynek technologii HDD, przyczyniając się do całkowitej modernizacji przewiertów. System pomiarowy tej firmy z dużą dokładnością lokalizował sondę, a zarazem dokładne miejsce położenia używanego do wiercenia narzędzia. W bardzo dużym stopniu ułatwiło to korygowanie zaplanowanej trajektorii i uzyskiwanie dokładniejszych przewiertów (Wiśniowski i in. 2020, Jin i Wei 2022).

Niewątpliwie był to kamień milowy w rozwoju technologii HDD. Możliwość stałej kontroli trajektorii wykonywanego przewiertu wpłynęła na znaczne rozpowszechnienie tej technologii przy przekraczaniu innych przeszkód terenowych, takich jak: jeziora, zbiorniki wodne, wiercenia w dnie morskim. Obecnie do kontroli trajektorii osi otworu stosuje się systemy hybrydowe i cyfrowe. Jest to połączenie systemu GPS, żyrokompasów i czujników inercyjnych (IMU). Pozwala to na bardzo precyzyjne określanie pozycji narzędzia wierzącego w przestrzeni 3D (Wang i in. 2017). W Europie również rozwijał się rynek technologii bezwykopowych. Od początku lat 80. projekty HDD realizowane były kolejno w Holandii, Niemczech, Wielkiej Brytanii, Francji. W następnej dekadzie przyszła kolej na Polskę, gdzie w 1991 roku doszło do debiutu technologii HDD przy przekraczaniu Wisły we Włocławku. Operatorem wiertniczym była niemiecka firma LMR Drilling. W latach 80. XX wieku powstało wiele europejskich firm odgrywających później pierwszoplanowe role na rynku HDD, takich jak: Nacap, Horizontal Drilling International, Visser & Smit Hanab czy wyżej wspomniane LMR Drilling. Przedsiębiorstwa te dysponowały urządzeniami pełnowymiarowymi, tzw. klasy *maxi rig*.

Urządzenia mobilne i kompaktowe pojawiały się u schyłku lat 80. oraz na początku lat 90. ubiegłego wieku. Były to urządzenia o sile uciągu sięgającej maksymalnie 500 kN, wprowadzone na rynek przez firmy takie jak Vermeer, Ditch Witch i Tracto-Technik. Do chwili obecnej wyprodukowano je w ogromnych ilościach: do 2018 roku liczba egzemplarzy przekroczyła pięćdziesiąt tysięcy na świecie (Carpender 2018). Wynikało to z dużego zapotrzebowania na krótkie instalacje z tworzyw sztucznych, które zazwyczaj nie przekraczały długości 500 m.

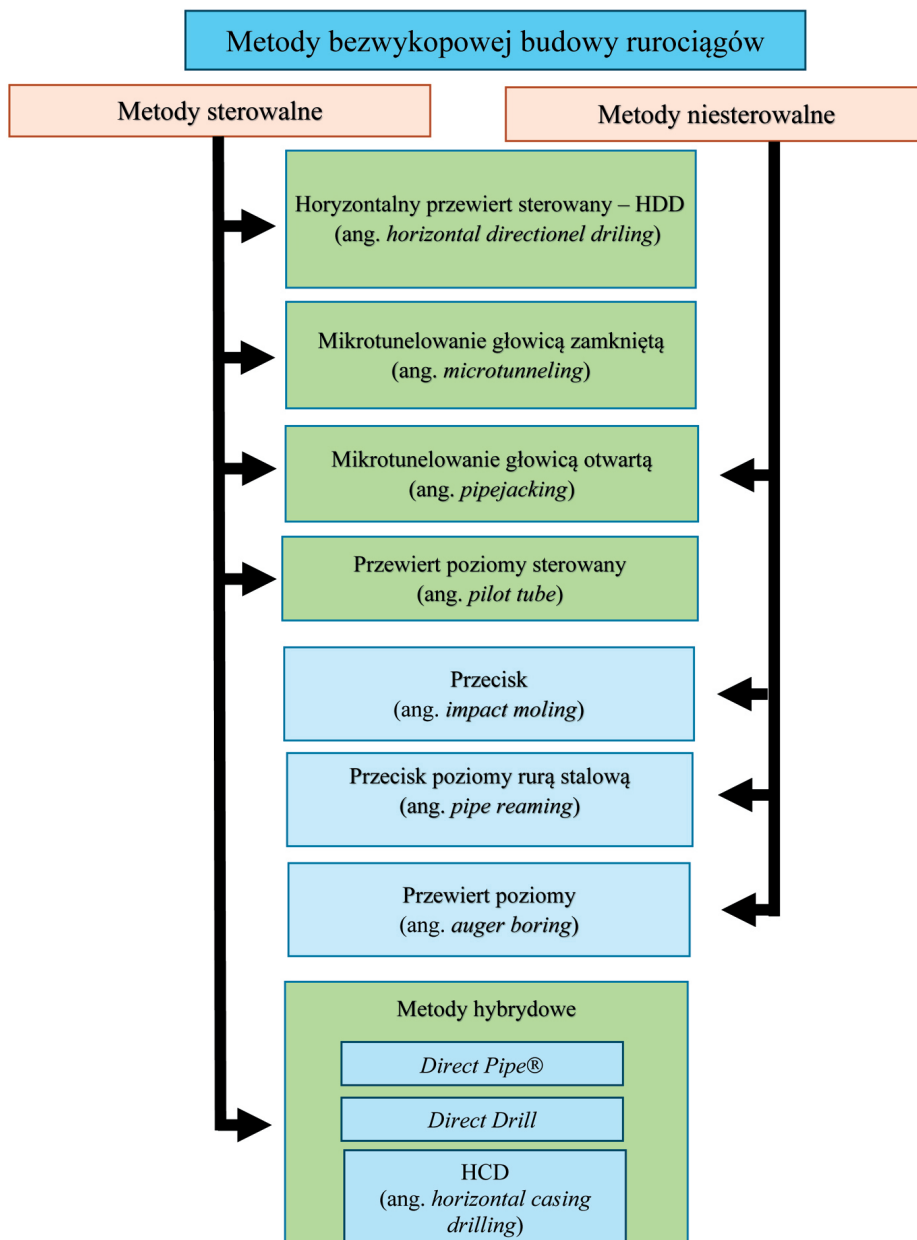
Z biegiem lat zmieniało się podejście do technologii HDD. Na początku lat 90. wszystkie przewiertu sięgające 1000 m były dużym osiągnięciem, a 10 lat później zostały uznane za standardowe. Coraz popularniejsze stało się instalowanie gazociągów przesyłowych o średnicy 40 cali (1016 mm). Przełomem w wykonywaniu otworów pilotowych została metoda *intersect* zastosowana po raz pierwszy w Wielkiej Brytanii w 2000 roku (Willoughby 2005). Polega ona na równoczesnym wykonywaniu otworu pilotowego po obydwóch stronach przekroczenia, do momentu przecięcia się trajektorii w ustalonym punkcie. Co pozwoliło wiercić coraz to większe długości wykonywanych przekroczeń, osiągając w 2018 roku magiczne (Jin i Wei 2022). W tabeli 1 zamieszczono dziesięć najdłuższych co do dystansu projektów wykonanych na świecie w technologii HDD.

Tabela 1
Dziesięć najdłuższych przewiertów HDD na świecie

Nazwa Projektu	Firma wykonawcza	Dystans	Średnica rurociągu	Rodzaj instalacji	Lokalizacja	Rok
Hongkong International Airport/ Aviation Fuel Pipeline	Huayuan (Chiny)	5205 m/ 5187 m	20"	rurociągi na paliwo lotnicze	Hongkong	2018
The North Bakken Expansion Sakakawea Lake	Michels (USA)	4702 m	24"	gazociąg	USA	2021
Mawan Power Plant – Dachan Island	Huayuan (Chiny)	4635 m	24"	gazociąg	Chiny	2023
Marsdiep Crossing Den Heleder – Texel	LMR Drilling (Niemcy)	4608 m	12"	wodociąg	Holandia	2017
Jiaxing Port LPG Transmission Pipeline	CPP (Chiny)	4150 m	12¾"	gazociąg	Chiny	2023
Jiaxing Port LPG Transmission Pipeline	CPP (Chiny)	4150 m	12¾"	gazociąg	Chiny	2023
Brahmaputra River North East Gas Grid Pipeline Project	NRP Projects (Indie)	4102 m	24"	gazociąg	Indie	2022
The Tongming Bay/ Sino-Kuwait Refinery	Sinopec (Chiny)	4071 m/ 4060 m	12"	ropociąg/ gazociąg	Chiny	2019

Źródło: opracowanie własne na podstawie (Osikowicz 2023)

Historia technologii HDD sięga ponad 50 lat wstecz. Od samego początku technologia była awangardowa i zaliczana jest do sterowanych technologii bezwykopowej budowy rurociągów (rys. 1).

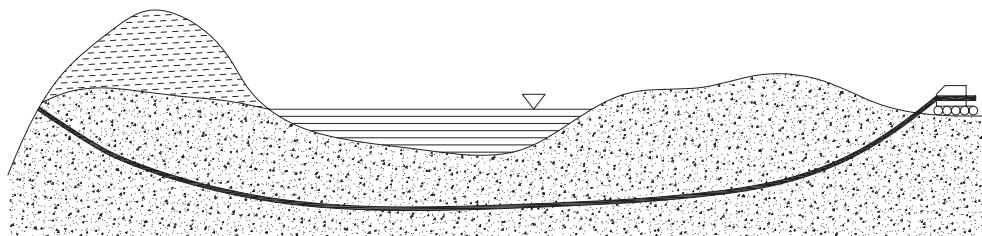


Rys. 1. Podział technologii bezwykopowej budowy rurociągów wg standardu Izby Gospodarczej Gazownictwa ST-IGG-3301:2021

Na przestrzeni lat można zaobserwować naturalną ewolucję tej metody, dzięki której rozwinęły się kolejne technologie bezwykopowej budowy rurociągów. Można śmiało stwierdzić, że HDD to technologia pierwotna, która została stworzona na potrzeby zrównoważonego rozwoju i wpisuje się w oczekiwania społeczne względem ochrony środowiska (Wiśniowski 2003, Ziaja 2018, Osikowicz 2018, Kuliczkowski 2019).

3. Zastosowania przemysłowe technologii HDD

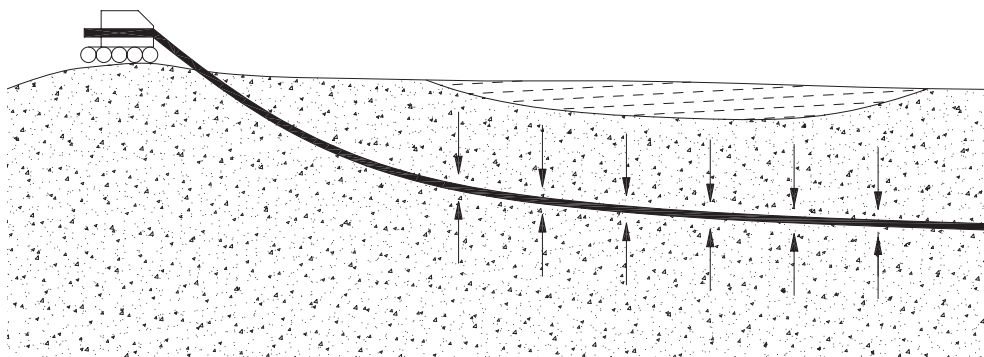
Choć technologia HDD została stworzona na potrzeby transportu rurociągami gazu, to jednak ze względu na swoje zalety bardzo szybko znalazła zastosowanie także w innych dziedzinach gospodarki. Obecnie oprócz przekraczania naturalnych przeszkód terenowych (rys. 2) technologię tę wykorzystuje się także do innych celów (rys. 3–6).



Rys. 2. Budowa nowych podziemnych instalacji rurowych dla transportu różnych rodzajów mediów: gaz, ropa, woda, ścieki, energia elektryczna itp.

Źródło: Wiśniowski (2003)

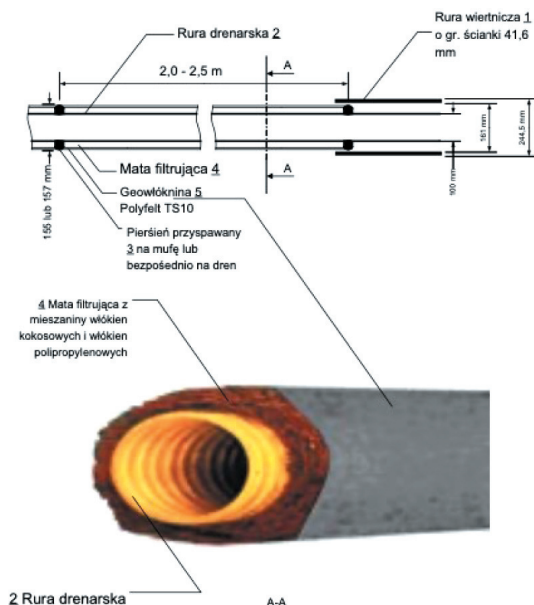
Jedną z możliwości jest wykonywanie drenów poziomych do odwodnień lub remediacji gruntów (rys. 3).



Rys. 3. Melioracja gruntów oraz oczyszczający drenaż gruntów

Źródło: Wiśniowski (2003)

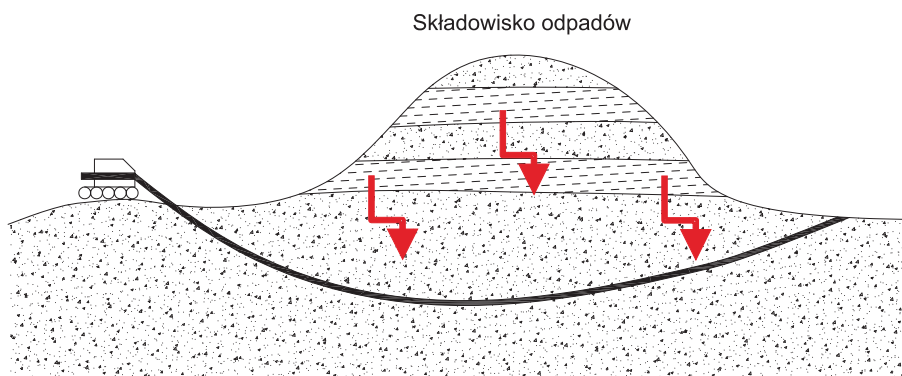
Na rysunku 4 przedstawiono specjalną rurę drenarską, opracowaną (Kosinkiewicz 2006) do drenażu poziomego wykonywanego w technologii HDD. Ponieważ jest ona instalowana przez wciąganie do otworu, niekiedy na znaczne głębokości, posiada specjalną konstrukcję, aby zapewnić jej odpowiednie parametry wytrzymałościowe i technologiczne.



Rys. 4. Schemat – rura drenarska do wykonywania drenażu w technologii bezwykopowej patent nr P-371067

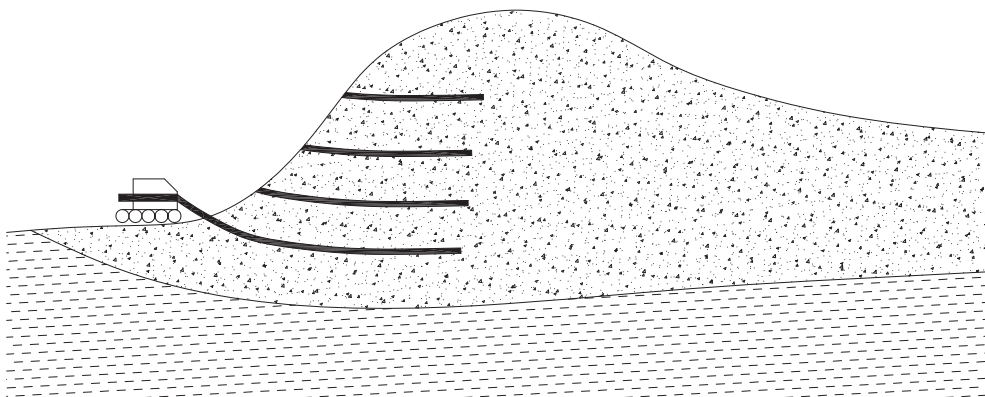
Źródło: Kosinkiewicz (2006)

Technologia HDD może także z powodzeniem być stosowana do wykonywania kanałów drenażowych do odbioru skażonych odcieków ze składowisk odpadów (rys. 5).



Rys. 5. Tworzenie bariery ochronnej pod składowiskiem odpadów w celu odbioru odcieku

Innym zastosowaniem technologii HDD jest wzmacnianie skarp, górotworu lub nawet pewnych konstrukcji inżynierskich, jak to pokazano na rysunku 6.



Rys. 6. Wzmacnianie skarp lub nasypów

Źródło: Wiśniowski (2003)

Jako przykład przemysłowy na rysunkach 7 i 8 pokazano budowę wiaduktu w ciągu drogi krajowej nr 79, przechodzącej nad linią PKP w miejscowości Krzeszowice w województwie małopolskim. W projekcie tym wykorzystano technologię HDD w celu wzmocnienia jego konstrukcji.

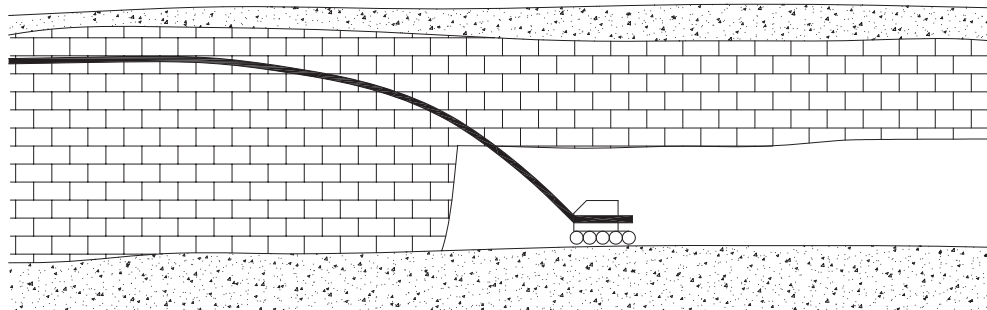


Rys. 7. Wiertnica Vermeer D16x20 podczas montażu ściągów metodą HDD
(fot. A. Jaroń)



Rys. 8. Bloki kotwiące na płytach drogowych
(fot. A. Jaroń)

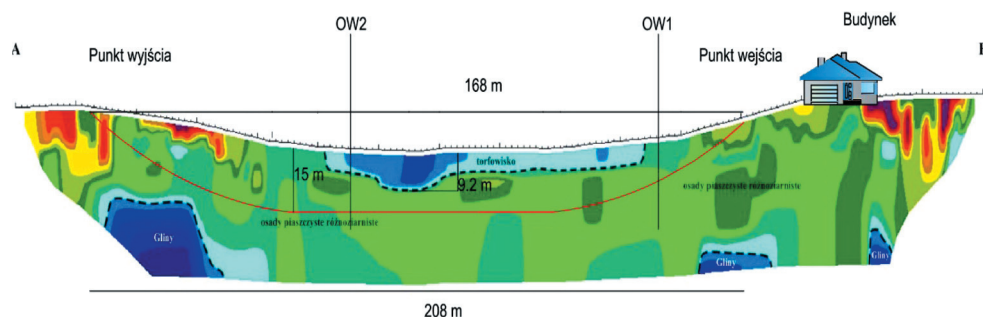
Jak pokazują doświadczenia przemysłowe, wykorzystując wiertnice HDD, można także wykonywać lokalne i precyzyjne wzmocnienia fundamentów budynków, tuneli lub innych konstrukcji przez iniekcję otworową (rys. 9).



Rys. 9. Wzmacnianie budynków, fundamentów i wykonywanie przyłączy
Źródło: Wiśniowski (2003)

4. Horyzontalny otworowy wymiennik ciepła – HOWC

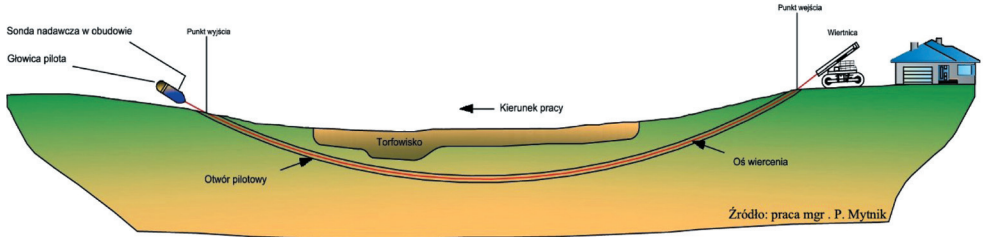
Jednym z ostatnio opracowanych (m.in. przez autora niniejszego artykułu) zastosowań technologii HDD jest wykonywanie przy jej użyciu horyzontalnych otworowych wymienników ciepła HOWC (rys. 10).



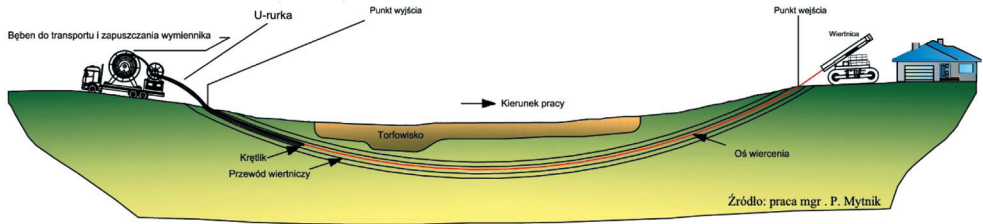
Rys. 10. Przykładowy projekt trajektorii HOWC
Źródło: Mytnik (2020)

Na kolejnych rysunkach przedstawiono fazę wiercenia otworu pilotowego (rys. 11) oraz instalację wymiennika ciepła (np. U-rurki) podczas wykonywania HOWC (rys. 12).

W opisaney powyżej technice zaprojektowano i wykonano w AGH podczas rozbudowy laboratorium Geoenergetyki II na Wydziale Wiertnictwa, Nafty i Gazu HOWC. Jego projekt przedstawiono na rysunku 13.

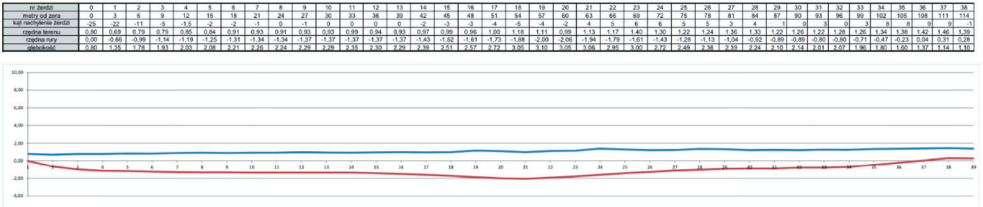


Rys. 11. Faza I: wiercenie otworu pilotowego
Źródło: Mytnik (2020)



Rys. 12. Faza II: instalacja U-rurki
Źródło: Mytnik (2020)

Przewiert sterowany rurą $\varnothing$ 5440 mm o długości 117 m
Zleceniodawca - AGH KRAKÓW
Lokalizacja - Kraków, ul. Reymonta 7
termin realizacji - 08.08.2023



Rys. 13. Projekt przewiertu pod HOWC.

Linia niebieska odwzorowuje poziom terenu, a czerwona wykonaną trajektorię

Otwór pilotowy o średnicy 4" (101,6 mm) został wykonany na podstawie zaprojektowanej trajektorii przewiertu (rys. 13) za pomocą maszyny przewiertowej firmy Vermeer (rys. 14). Następnie przystąpiono do jego rozwiercania do średnicy 6" (152,4 mm) (rys. 15), z równoczesną instalacją wiązki rur wymiennika (rys. 16). Oczywiście przed przystąpieniem do wykonania prac przewiertowych wszystkie trasy istniejących już pod ziemią instalacji zostały rozpoznane, a urządzenia podziemne zlokalizowane i w miarę możliwości odkryte. Otwór pilotowy między studzienkami został wykonany od strony ul. Czarnowiejskiej (budynek C-4). Następnie zainstalowano komplet pięciu rur o średnicy 40 mm (dodatkowa rura jako rezerwa).



Rys. 14. Maszyna HDD



Rys. 15. Wiercenie pilota



Rys. 16. Instalacja rurek

Wykonanie horyzontalnego otworowego wymiennika ciepła za pomocą technologii HDD ma następujące zalety w porównaniu z techniką wykopową służącą standardowo do budowy płytłych wymienników poziomych:

- ułożenie instalacji jest możliwe na głębokościach eliminujących możliwość przemrożenia przypowierzchniowej warstwy gruntu (praktycznie nie ma limitu głębokości ułożenia);
- brak konieczności prowadzenia w jednym otworze przewodu z chłodnym i ogrzanym nośnikiem ciepła;
- brak oddziaływania na roślinność ponad przewiertem;
- brak wpływu czynników pogodowych w fazie budowy;
- krótki czas realizacji projektu – opisany projekt wiertniczy wykonano w ciągu jednej zmiany;
- możliwość układania instalacji pod terenami zielonymi, zurbanizowanymi lub niedostępnymi z powierzchni, udostępniającymi tzw. lokalne wyspy ciepła;
- udostępniona jest znaczna pojemność cieplna skał otaczających przewód wymiennika.

Do wad można zaliczyć:

- ubogie doświadczenia w wykonywaniu tego typu instalacji, mała konkurencyjność na rynku;
- wymagany specjalistyczny sprzęt do prac wiertniczych;
- konieczność zabezpieczenia dostępu dla prac po dwóch stronach przewiertu (punkt wejścia i wyjścia).

Ograniczenia techniczne to:

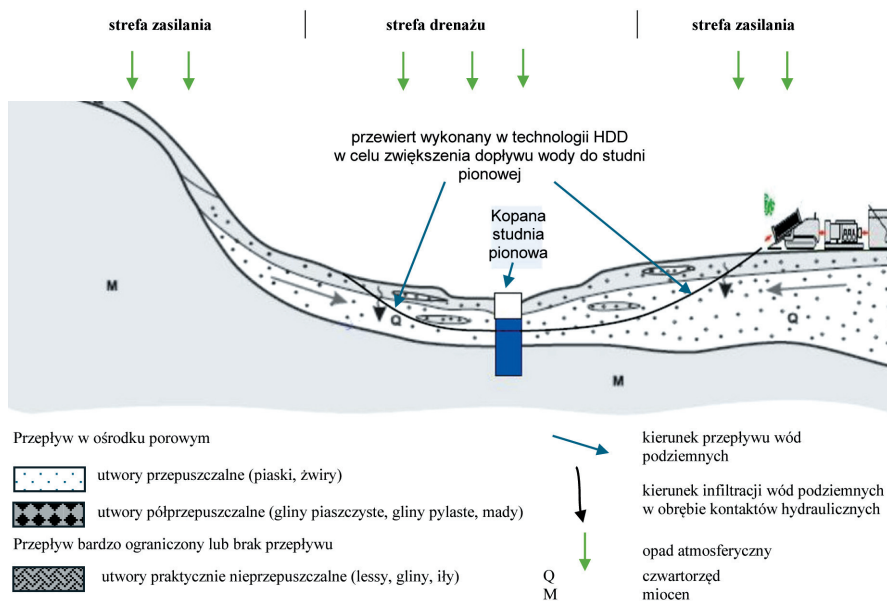
- wymagana stosunkowo długa działka (raczej prostokątna);
- odpowiednia wytrzymałość na rozciąganie (etap instalacji) i zgniatanie (etap eksploatacji) rurek wymiennika.

5. Innowacyjne zastosowanie technologii HDD w celu zwiększenia dopływu wód gruntowych do studni pionowej

Wykonanie studni głębinowych coraz częściej nie pozwala na dowiercenie się do warstw wodonośnych o wystarczającej miąższości do zaopatrzenia odbiorców w wodę, a wynika to często z samej budowy geologicznej (mała miąższość warstwy wodonośnej) w danej lokalizacji. Dlatego proponuje się wykorzystanie technologii HDD do zwiększenia dopływu wód podziemnych do istniejącej studni pionowej. Schemat takiego projektu przedstawiono na rysunku 17.

Podczas wykonywania przewiertów pod studnie poziome, przewody odwodnieniowe i drenażowe korzystne jest używanie płuczek biodegradowalnych skonstruowanych na bazie naturalnych polimerów. Takie płyny wiertnicze pozwalają na kontrolę parametrów reologicznych i czasu rozkładu pozostawionej w otworze płuczki dzięki dodatkom biocydów i enzymów ją degradujących.

Właściwy dobór płuczki wiertniczej oraz prawidłowa jej aplikacja zdecydowanie zmniejszają możliwość kolmatacji wykonanej studni czy drenu. Zwiększając efektywność dopływu wody. Płyn wiertniczy jest charakteryzowany parametrami, które muszą być za każdym razem analizowane w celu jak najlepszej optymalizacji procesu wiercenia. Kolejnym problemem technicznym jest instalacja i zabudowa filtra w tego typu otworach poziomych. W tym celu zespół pracowników WWNiG opracował i opatentował: „Sposób instalacji modułowego filtra wody podziemnej w przewiercie oraz moduł filtra wody podziemnej” (nr patentu P.442369, Solecki i in. 2022). Wynalazek ten pozwala na sprawny i tani montaż filtra w przewiercie horyzontalnym.



Rys. 17. Schemat koncepcji zwiększenia dopływu wody do studni pionowej z wykorzystaniem technologii HDD

Źródło: Gudź (2016)

6. Uszczelnianie składowisk odpadów metodą HDD

Początek XXI wieku bez wątpienia w Europie można nazwać wiekiem ochrony środowiska. Najpierw dziura ozonowa, a teraz ocieplenie klimatu i transformacja energetyczna są języczkiem u wagi dużej grupy naukowców. Jednak nie można zapominać o pozostałościach minionej epoki, XX wieku i niepoliczalnej liczbie legalnych i nielegalnych składowisk różnego rodzaju odpadów. W Polsce nadal są to kwestie problematyczne (rys. 18).

a)



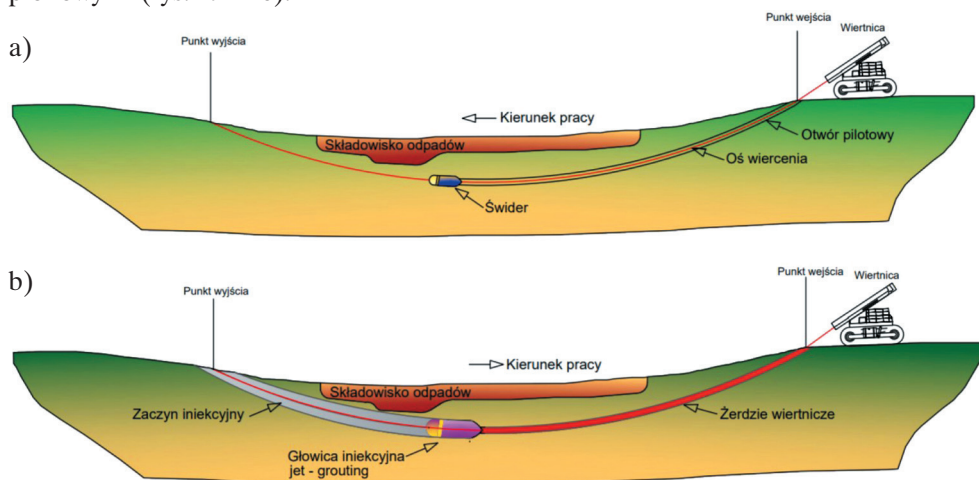
b)



Rys. 18. Widok przykładowej komory z odpadami w okresie deszczowym (a) i suchym (b) zlokalizowanej w centralnej Polsce

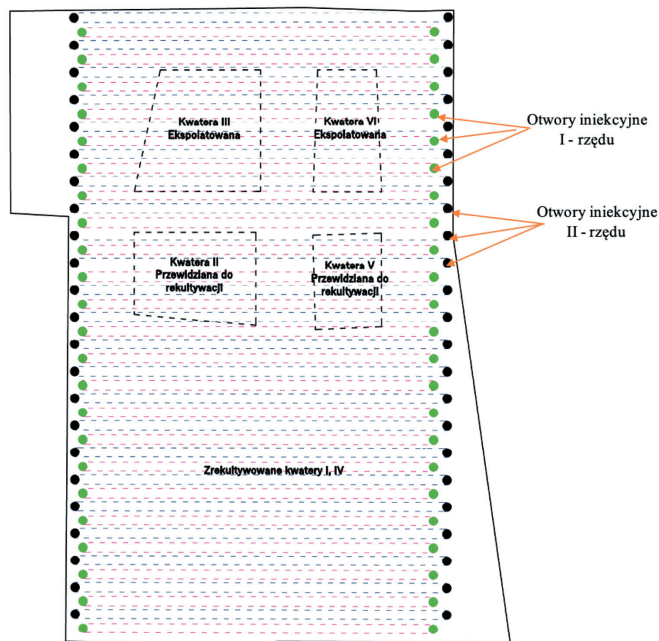
(fot. A. Jamrozik)

Jednak nawet z tak trudnym i toksycznym składowiskiem można sobie poradzić, stosując technologię HDD i wykorzystując doświadczenia iniekcji górotworu otworami pionowymi (rys. 19 i 20).



Rys. 19. Schemat wykonywania otworu do iniekcji w technologii HDD:

- wiercenie otworu pilotowego pod iniekcję;
- wykonanie prac iniekcyjnych przy marszu powrotnym



Rys. 20. Doszczelnianie składowiska odpadów metodą iniekcji otworowej za pomocą technologii HDD

Główne zalety tego rozwiązania to, że:

- otrzymuje się uszczelnienie nie tylko pionowe wokół składowiska, ale także poziome pod nim – powstaje szczelna wanna;
- przesłona może być wykonana jako iniekcja klasyczna lub strumieniowa.

Główną wadą tej koncepcji są wysokie koszty wykonywanych prac. Z tego powodu opracowany przez WWNiG AGH projekt jeszcze nie doczekał się realizacji.

7. Przekraczanie linii brzegowej instalacjami rurowymi

Transformacja energetyczna z założenia opiera się na odnawialnych źródłach energii. Morska energetyka wiatrowa (MEW) jest obecnie jedną z najszybciej rozwijających się technologii OZE na świecie, ale także w Europie i Polsce.

W Europie Zachodniej 12 państw posiada ponad 120 farm, które mają już blisko 30 GW mocy zainstalowanej w MFW. Do roku 2030 planowane jest uzyskanie mocy 111 GW dla wszystkich projektów europejskich. W Polsce właśnie rozpoczynamy realizację tego ambitnego projektu. Zgodnie z zapisami w „Polityce energetycznej Polski 2040” (Raport PEP2040 2022), do 2030 roku morska energetyka wiatrowa ma charakteryzować się mocą blisko 6,0 GW, a pierwsze czynne instalacje wiatrowe na polskim obszarze morskim powstać mają już w 2026 roku. Natomiast do roku 2040 osiągnięty ma zostać poziom ok. 11 GW (tab. 2).

Tabela 2
Moc planowanych MFW w Polsce dla pierwszych 10 koncesji

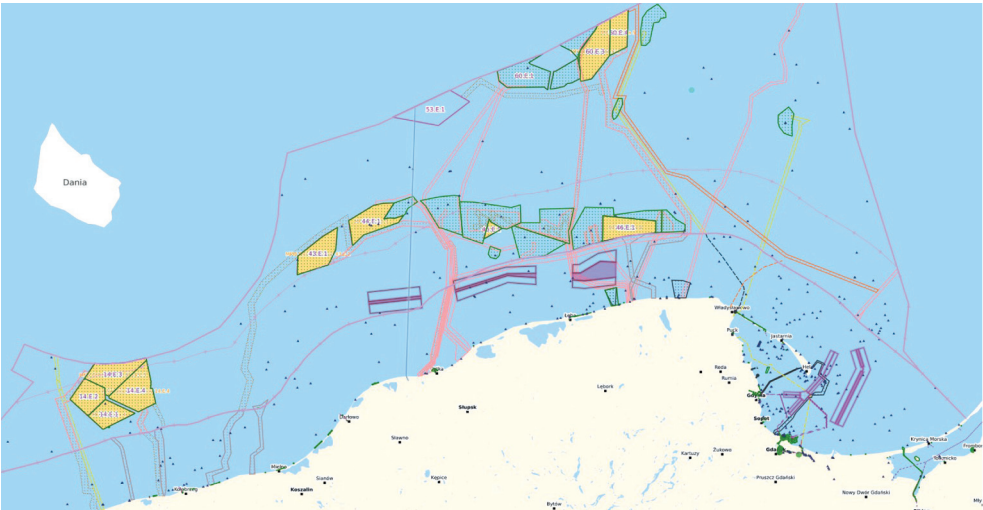
Nazwa projektu	Powierzchnia farmy	Planowana moc przyłączeniowa
Polenergia – Bałtyk I	128 km ²	1,560 GW
Polenergia/Equinor – Bałtyk II	122 km ²	600 MW
Polenergia/Equinor – Bałtyk III	116 km ²	1,200 GW
PGE Baltica 2	189 km ²	1,498 GW
PGE Baltica 3	131 km ²	1,045 GW
PGE Baltica 1	108 km ²	b.d
RWE Renewables	42 km ²	350 MW
PKN Orlen – Baltica Power	131 km ²	1,200 GW
EDPR – B-Wind	42 km ²	b.d
EDPR – C-Wind	42 km ²	b.d
Razem	1084 km ²	7,093 GW

Przekroczenia linii brzegowej mórz i oceanów są bardzo aktualnym tematem ze względu na potrzebę wprowadzenia kabli elektrycznych z farm wiatrowych na ląd. Początkowa faza budowy MEW w Polsce pociągnie za sobą konieczność wykonania 28 wejścia kablem w głąb lądu (tab. 3). Na chwilę obecną (stan na wrzesień 2025) wykonano w Polsce już 5 z tych przekroczeń, w tym 4 dla projektu Baltic Power.

Tabela 3
Planowana liczba wejść w dno morza kablem energetycznym

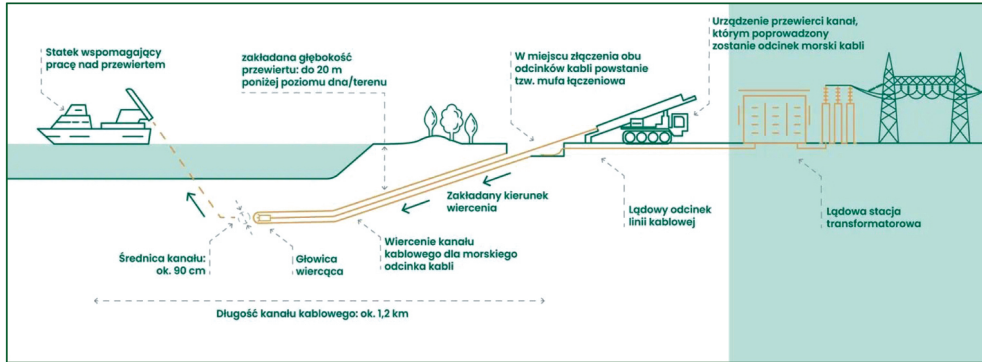
Projekt	Liczba przekroczeń
Bałtyk I	3
Bałtyk II	4
Bałtyk III	3
Baltica 1	4
Baltica 2	4
Baltica 3	3
RWE	1
Baltic Power	4
B-Wind / C-Wind	2
Suma	28

Transformacja energetyczna w Polsce nabiera tempa, a farma wiatrowa Baltic Power staje się jej wizytówką. Po jej uruchomieniu w 2026 roku do polskiego systemu energetycznego zostanie włączone dodatkowe niemal 1,2 GW mocy. Farma Baltic Power zlokalizowana jest 23 km od wybrzeża Bałtyku (rys. 21). Stoją tam turbiny o wysokości ponad 230 m. Do końca 2025 roku zaplanowano, że na farmie zostanie zamontowane 76 turbin wiatrowych Vestas o mocy 15 MW każda (ORLEN 2025). W tabeli 2 przedstawiono parametry techniczne planowanych inwestycji oraz liczbę zaprojektowanych przekroczeń linii brzegowej (tab. 3).



Rys. 21. Zaplanowane obszary budowy MFR w polskiej strefie Morza Bałtyckiego
Źródło: System Informacji Przestrzennej Administracji Morskiej

Na rysunku 22 przedstawiono typowy schemat wykonywania przekroczeń morskiej linii brzegowej w technologii HDD, natomiast na rysunku 23 przedstawiono jedną z pierwszych instalacji rury osłonowej do kabla energetycznego w Polsce.



Rys. 22. Możliwy przebieg instalacji rurociągu w przejściu brzegowym
Źródło: Gdalewski (2024)

a)



b)



Rys. 23. Początek (a) i koniec (b) procesu instalacji rury osłonowej pod kable energetyczne w projekcie Baltic Power
(fot. J. Janicki, 2024)

8. Podsumowanie

Pomimo stopniowego kończenia rozbudowy sieci przesyłowych gazu wysokiego ciśnienia (MPO 8,4 MPa) przez firmę GAZSYSTEM do 2027 roku, przyszłość technologii HDD wydaje się niezagrażona w dobie transformacji energetycznej. Można nawet stwierdzić, że dzięki niej technologia HDD zyska nowe obszary zastosowań. Szczególnie obiecująco wygląda rynek przesyłu energii elektrycznej, który stwarza duże perspektywy rozwoju. Równie duży potencjał tkwi w wykorzystaniu HDD do budowy horyzontalnych otworowych wymienników ciepła. Z kolei w obliczu narastającego w szybkim tempie deficytu wody pitnej w Polsce, technologia HDD staje się doskonałym rozwiązaniem do udostępniania zasobów wód podziemnych.

Obecnie dostępne systemy sterowania i pomiaru położenia trajektorii wierconego otworu, oparte na żyrokompasach, zapewniają bardzo dużą dokładność jej wykonania, sięgającą rzędu kilku centymetrów.

Literatura

- Belani D., Pitroda J., Bhavsar J.J., 2014, *A Review On Trenchless Technology: State Of Art Technology For Underground Utility Services*, [w:] *National Conference on "Trends and Challenges of Civil Engineering in Today's Transforming World"*, 29 March 2014, Civil Engineering Department S.N.P.I.T. & R.C., Umrah.
- Bennett D., Ariaratnam S., 2017, *Horizontal Directional Drilling (HDD). Good Practices Guidelines*, The North American Society for Trenchless Technology, HDD Consortium, USA, 4th Edition.
- Gdalewski K., 2024, *Modelowanie numeryczne zachowania się rurociągu wykonanego z rur PE podczas ich instalacji w technologii HDD* [praca magisterska], Akademia Górniczo-Hutnicza, Kraków.
- Gudz T., 2016, *Analiza możliwości niestandardowego zastosowania sterowanych przewierć horyzontalnych* [praca magisterska], Akademia Górniczo-Hutnicza, Kraków.
- Jin L., Wei J., 2022, *Research on Horizontal Directional Drilling (HDD) Trajectory Design and Optimization Using Improved Radial Movement Optimization*, *Applied Sciences*, vol. 12(23), art. 12207. <https://doi.org/10.3390/app122312207>.
- Kosinkiewicz Z., 2006, *Drenaż poziomy wykonany bezwykopowo*, *Inżynieria Bezwykopowa*, t. 4, s. 86–87.
- Kuliczkowski A. (red.), 2019, *Technologie bezwykopowe w inżynierii środowiska*, wyd. 2, Wydawnictwo Seidel-Przywecki, Warszawa.
- Raport PEP2040, 2022, Ministerstwo Klimatu i Środowiska, *Polityka energetyczna Polski 2040 (PEP2040)*. <https://www.gov.pl/web/ia/polityka-energetyczna-polski-do-2040-r-pep2040> [dostęp: 10.08.2025].

- Mytnik P., 2020, *Możliwości aplikacji technologii bezwykopowych do budowy otworowych wymienników ciepła* [praca magisterska], Akademia Górniczo-Hutnicza, Kraków.
- Najafi M., Gokhale S., Calderon D., Ma B., 2021, *Trenchless Technology: Pipeline and Utility Design, Construction, and Renewal*, 2nd Edition. McGraw Hill.
- ORLEN, 2025, *MEW w Grupie ORLEN*. <https://neptun.orlen.pl/pl/morska-energetyka-wiatrowa/mew-w-grupie-orlen> [dostęp: 19.08.2025].
- Osikowicz R., 2018, *Planowanie i realizacja projektów HDD. Część I: przegląd literatury przedmiotu*, Inżynieria Bezwykopowa, nr 2, s. 102–111.
- Osikowicz R., 2023, *Long and Large Diameter HDD Crossings. Top 100 światowych projektów*, Inżynieria Bezwykopowa, nr 1, s. 70–96.
- System Informacji Przestrzennej Administracji Morskiej. <https://sipam.gov.pl/geoportal> [Wydane decyzje lokalizacyjne – mapa dostępna dla zalogowanych użytkowników] [dostęp: 10.06.2022].
- Solecki T., Solecki M., Stopa J., Wiśniowski R., Ziaja J., 2022, *Sposób instalacji modułowego filtra wody podziemnej w przewiercie oraz moduł filtra wody podziemnej*, Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, wynalazca Int.Cl.: E21B 43/10, Urząd Patentowy RP, opis zgłoszeniowy PL442369A1, opubl. 2.04.2024, zgłosz. nr P.442369 z dn. 27.09.2022.
- Wang Y., Wen G., Chen H., 2017, *Horizontal Directional Drilling-Length Detection Technology While Drilling Based on Bi-Electro-Magnetic Sensing*, Sensors, vol. 17(5), art. 967. <https://doi.org/10.3390/s17050967>.
- Willoughby D.A., 2005, *Horizontal Directional Drilling: Utility and Pipeline Applications*, McGraw-Hill, New York.
- Wiśniowski R., 2003, *Otworki kierunkowe – propozycja klasyfikacji, pojęcia podstawowe*, Techniki i Technologie Bezwykopowe, nr 1, s. 36–41.
- Wiśniowski R., Łopata P., Orłowicz G., 2020, *Numerical Methods for Optimization of the Horizontal Directional Drilling (HDD) Well Path Trajectory*, Energies, vol. 13(15), art. 3806. <https://doi.org/10.3390/en13153806>.
- Yan J., Ariaratnam S., Dong S., Zeng C., 2018, *Horizontal Directional Drilling: State-of-the-art Review of Theory and Applications*, Tunnelling and Underground Incorporating Trenchless Technology, vol. 72, s. 162–173. <https://doi.org/10.1016/j.tust.2017.10.005>.
- Ziaja J., Wiśniowski R., Jamrozik A., Knez D., 2018, *Modern Construction Technologies of Gas Pipelines and Oil Pipelines*, [w:] *SGEM 2018 – 18th International Multidisciplinary Scientific Geoconference*, 2–8 July 2018, Albena, Bułgaria, s. 613–620.

