



# Koła naukowe etapem rozwoju kompetencji zawodowych i naukowych

*redakcja  
Barbara Buraczyńska*



MONOGRAFIE

Koła naukowe etapem rozwoju  
kompetencji zawodowych i naukowych

# Monografie – Politechnika Lubelska

## Rada Naukowa Wydawnictwa Politechniki Lubelskiej

**Przewodnicząca:**  
Agnieszka RZEPKA

**Dyrektor CIN-T:**  
Katarzyna WEINPER

### **Wydawnictwo Politechniki Lubelskiej:**

Magdalena CHOŁOJCZYK  
Karolina FAMULSKA-CIESIELSKA  
Jarosław GAJDA  
Anna KOŁTUNOWSKA  
Katarzyna PEŁKA-SMĘTEK  
Anna STROJEK

### **Przedstawiciele Dyscyplin Naukowych Politechniki Lubelskiej:**

Marzenna DUDZIŃSKA  
Małgorzata FRANUS  
Arkadiusz GOLA  
Paweł KARCZMAREK  
Beata KOWALSKA  
Anna KUCZMASZEWSKA  
Jarosław LATAJSKI  
Tomasz LIPECKI  
Zbigniew ŁAGODOWSKI  
Joanna PAWŁAT  
Lucjan PAWŁOWSKI  
Natalia PRZESMYCKA  
Magdalena RZEMIENIAK  
Mariusz ŚNIADKOWSKI

### **Przedstawiciele honorowi:**

Zhihong CAO, Chiny  
Miroslav GEJDOŠ, Słowacja  
Karol HENSEL, Słowacja  
Hrvoje KOZMAR, Chorwacja  
Frantisek KRCMA, Czechy  
Sergio Lujan MORA, Hiszpania  
Dilbar MUKHAMEDOVA, Uzbekistan  
Sirgii PAWŁOW, Ukraina  
Natalia SAVINA, Ukraina  
Natia SHENGELIA, Gruzja  
Daniele ZULLI, Włochy

# Koła naukowe etapem rozwoju kompetencji zawodowych i naukowych

redakcja  
Barbara Buraczyńska



POLITECHNIKA  
LUBELSKA  
WYDAWNICTWO

Lublin 2023

**Recenzenci:**

prof. dr hab. Stepan Shakhno, Narodowy Uniwersytet im. Iwana Franki  
we Lwowie, Ukraina

prof. dr hab. inż. Andrzej Chochowski, SGGW w Warszawie

Autorka fotografii na okładce: Ewa Organista

Publikacja wydana za zgodą Rektora Politechniki Lubelskiej

ISBN: 978-83-7947-571-1

Wydawca: Wydawnictwo Politechniki Lubelskiej

[www.wpl.pollub.pl](http://www.wpl.pollub.pl)

ul. Nadbystrzycka 36C, 20-618 Lublin

tel. (81) 538-46-59

Druk: Agencja Reklamowa TOP Agnieszka Łuczak

[www.agencjatop.pl](http://www.agencjatop.pl)

---

Elektroniczna wersja książki dostępna w Bibliotece Cyfrowej PL [www.bc.pollub.pl](http://www.bc.pollub.pl)

Książka udostępniona jest na licencji Creative Commons Uznanie autorstwa - na tych samych warunkach 4.0 Międzynarodowe (CC BY-SA 4.0)

Nakład: 50 egz.

# Spis treści

<b>O Konferencji.....</b>	<b>7</b>
Alicja Bilińska, Bohdan Kosharnyi, Faustyna Smarzewska	
<b>Sieci złożone w analizie bezpieczeństwa na przykładzie siatki terrorystycznej.....</b>	<b>11</b>
Dagmara Dudek	
<b>Wybrane zagadnienia związane z kwantylami.....</b>	<b>31</b>
Marcin Dziadosz	
<b>Zastosowanie teorii kolejek do poprawy jakości obsługi klientów poczty .....</b>	<b>49</b>
Jakub Kisała	
<b>Technologia i badania struktur cienkowarstwowych wykazujących zjawisko gigantycznego magnetooporu.....</b>	<b>73</b>
Katarzyna Korulczyk, Kamil Kapłon, Magdalena Paśnikowska-Łukaszuk, Aleksandra Prus	
<b>Przegląd ekopolimerów dostępnych w polskich sklepach w latach 2018–2022 .....</b>	<b>87</b>
Marek Nowosad	
<b>Naprawa czy na śmietnik? .....</b>	<b>101</b>
Emilia Popławska	
<b>Hipergrafy mieszane i ich kolorowanie.....</b>	<b>109</b>
Albert Rachwał	
<b>Wybrane metody klasyfikacji w zadaniach uczenia maszynowego .....</b>	<b>123</b>
Alicja Rachwał	
<b>Autoenkodery w nienadzorowanym uczeniu maszynowym .....</b>	<b>141</b>
Agnieszka Szczepaniak, Milena Wełna, Ewelina Stąpór, Jolanta Wójtowicz	
<b>Inspektor ochrony danych osobowych jako droga rozwoju studentów .....</b>	<b>161</b>
Joanna Szkutnik-Rogoż	
<b>Modelowanie procesu eksploatacji pojazdów wojskowych w aspekcie gotowości.....</b>	<b>169</b>
Joanna Tomasik	
<b>Analiza wpływu liczby przywiązanych ligandów na budowę włochatej nanocząstki przy wykorzystaniu symulacji komputerowych.....</b>	<b>183</b>

Konferencja tematyczna pt. *Koła naukowe etapem w rozwoju kompetencji zawodowych i naukowych*, skierowana do młodych naukowców, organizowana w ramach Jubileuszu 15-lecia Wydziału Podstaw Techniki.

### **Organizatorzy**

Dziekan Wydziału Podstaw Techniki:  
dr hab. inż. Dorota Wójcicka-Migasiuk, prof. uczelni  
Koło Naukowe Matematyków „Kwaternion”  
Koło Naukowe Inżynierii Bezpieczeństwa „A Team of Safety”

### **Komitet Naukowy**

Przewodnicząca:  
dr hab. Anna Kuczmaszewska, prof. uczelni

#### Członkowie:

dr hab. inż. Dorota Wójcicka-Migasiuk, prof. uczelni  
prof. dr hab. inż. Mychajło Paszczko  
dr hab. Yaroslav Chabanyuk, prof. uczelni  
dr hab. Mariusz Śniadkowski, prof. uczelni  
dr inż. Sebastian Gnapowski  
dr Izolda Gorgol  
dr inż. Joanna Szulżyk-Cieplak

### **Komitet Organizacyjny**

Przewodnicząca:  
dr hab. Anna Kuczmaszewska, prof. uczelni

#### Członkowie:

dr inż. Barbara Buraczyńska  
dr Paweł Właż  
mgr inż. Magdalena Paśnikowska-Łukaszuk  
mgr inż. Aleksandra Prus  
mgr inż. Dagmara Dudek  
mgr inż. Magda Włazło  
mgr inż. Magdalena Piłat-Rożek

## O Konferencji

Rok akademicki 2022/2023 jest dla Politechniki Lubelskiej rokiem jubileuszy. Dnia 13 maja 2023 roku obchodzimy 70. rocznicę utworzenia Wieczorowej Szkoły Inżynierskiej, która w kolejnych latach przekształcała się najpierw w Wyższą Szkołę Inżynierską i ostatecznie w Politechnikę Lubelską, a 28 grudnia 2022 roku minęła 15. rocznica rozpoczęcia działalności Wydziału Podstaw Techniki. Idea powstania tego Wydziału zakładała zintegrowanie nauk podstawowych w jednej jednostce celem wzmocnienia ich potencjału naukowego i dydaktycznego. Zaowocowało to już w pierwszym roku działalności uruchomieniem kształcenia na kierunku matematyka. Obecnie prowadzone są na nim cztery kierunki studiów: edukacja techniczno-informatyczna (I i II stopień kształcenia), matematyka (I i II stopień kształcenia), inżynieria bezpieczeństwa (I stopień kształcenia) oraz inżynieria i analiza danych (I stopień kształcenia, trwają przygotowania do uruchomienia II stopnia).

15-lecie powstania Wydziału Podstaw Techniki stało się okazją do zorganizowania dwóch konferencji jubileuszowych. Jedną z nich była Konferencja Młodych Naukowców pod nazwą „Koła naukowe etapem w rozwoju kompetencji zawodowych i naukowych”.

Uczestnikami Konferencji byli studenci i doktoranci zrzeszeni w kołach naukowych oraz młodzi pracownicy, których zainteresowania naukowe i pierwsze naukowe „szlify” inicjowane były właśnie w tych kołach.

Ideą Konferencji było pokazanie, w jaki sposób aktywność w studenckim ruchu naukowym wpływa na rozwój kariery zawodowej i pomaga w początkach działalności naukowej.

Koła naukowe dają możliwość rozwijania własnych zainteresowań i kompetencji poprzez realizację różnych projektów, w tym włączanie studentów w prace naukowe prowadzone w jednostkach uczelnianych, udział w konferencjach, spotkaniach i dyskusjach naukowych.

Praca w kołach naukowych pozwala pogłębiać wiedzę teoretyczną, ale jednocześnie uczy, jak skutecznie wykorzystywać ją do rozwiązywania różnych problemów. Wiele karier zawodowych i naukowych znanych przedsiębiorców, wynalazców i naukowców miało swój początek w studenckim ruchu naukowym. Nasza Konferencja, która odbyła się 18 listopada 2022 roku, była świetną okazją do prezentacji dorobku kół działających na Politechnice Lubelskiej, ale nie tylko, bo jej uczestnikami były również osoby spoza lubelskiego środowiska akademickiego.



Referat wprowadzający pt. „Koła naukowe, czyli jak skorzystać ze studiów dużo bardziej” przedstawiła absolwentka Politechniki Warszawskiej, obecnie doktorantka Politechniki Federalnej w Zurychu (Eidgenössische Technische Hochschule Zürich, ETHZ), Małgorzata Łazuka. Podzieliła się w nim swoimi doświadczeniami z okresu działalności w kole naukowym na Politechnice Warszawskiej i refleksjami o ich wpływie na jej dalsze wybory i karierę zawodową.

Organizatorem Konferencji byli Dziekan Wydziału Podstaw Techniki i dwa koła naukowe działające w tym wydziale – Studenckie Koło Naukowe „Kwaternion” i Koło Naukowe Inżynierii Bezpieczeństwa „A Team of Safety”.

Prezentowana monografia jest zbiorem 12 prac prezentowanych podczas Konferencji w postaci referatów i posterów.

Autorami rozdziału pierwszego są studenci III roku studiów I stopnia na kierunku matematyka, aktywnie działający w kole „Kwaternion”. Przedstawiają oni, na przykładzie siatki terrorystycznej, możliwości wykorzystania złożonych sieci w prewencji zagrożeń, jakie stwarza terroryzm.

Drugi rozdział poświęcony jest zagadnieniom związanym z kwantylami i ich estymatorami. Kwantyle są jednymi z ważniejszych parametrów rozkładu cech, znajdującymi szerokie zastosowanie np. w finansach czy ekonomii. Autorka jest pracownikiem Katedry Matematyki Stosowanej, w przeszłości współpracowała z kołem „Kwaternion”.

W rozdziale trzecim autor przedstawił, jak metody matematyczne (w tym przypadku teoria kolejek) mogą pomóc w poprawie jakości obsługi klientów w miejscach takich jak poczta, bank czy przychodnia zdrowia. Autor jest pracownikiem Katedry Matematyki Stosowanej, działał w kole „Kwaternion”.

Kolejny, czwarty rozdział dotyczy produkcji i badań struktur cienkościennych wykazujących zjawisko gigantycznego magnetooporu. Praca powstała w ramach działalności Koła Naukowego Elektroników i Mechatroników „SEMICON”.

W rozdziale piątym autorzy nawiązują do aktualnego problemu ograniczenia odpadów i śmieci ropopochodnych. Przedstawiają ofertę ekopolimerów dostępnych w polskich placówkach handlowych w latach 2018–2022. Autorzy tego rozdziału są absolwentami Wydziału Podstaw Techniki aktywnie działającymi w różnych kołach naukowych Wydziału w okresie swoich studiów.

Rozdział szósty monografii jest również poświęcony problemowi szeroko rozumianej ochrony środowiska. Autor, student I roku kierunku edukacja techniczno-informatyczna, rozważa dylemat „naprawiać czy wyrzucić” w kontekście szybkiego rozwoju technologii IT. Praca powstała w ramach działalności Studenckiego Koła Naukowego „InfoNaBi”.

Autorka rozdziału siódmego przez cały okres studiów była związana z kołem „Kwaternion”. W swojej pracy prezentuje wybrane zagadnienia z teorii hipergrafów mieszanych i ilustruje je przykładami zastosowań.

Rozdziały ósmy i dziewiąty poświęcone są uczeniu maszynowemu. Obaj autorzy to absolwenci, aktywnie działający w okresie studiów w kole „Kwaternion”.

Autorki rozdziału dziesiątego są członkiniami Koła Naukowego „InfoNaBi”. Ich praca dotyczy ochrony danych osobowych, historii jej powstawania, regulacji prawnych oraz zadań inspektora ochrony danych osobowych.

W rozdziale jedenastym przedstawione zostało wykorzystanie oprogramowania R do wyznaczenia współczynnika gotowości technicznej, jako miary niezawodności, pojazdów ciężarowych wykorzystywanych na terenie Polski. Autorka pracy jest doktorantką Wojskowej Akademii Technicznej i absolwentką Politechniki Lubelskiej na kierunku matematyka.

Rozdział dwunasty poświęcony jest analizie zmian finalnej struktury włóchatej nanocząstki w zależności od liczby przywiązanych do rdzenia polimerowych ligandów. Autorka jest studentką I roku kierunku edukacja techniczno-informatyczna i działa w Kole Naukowym Inżynierii Polimerów CARBON (wcześniejsza nazwa Studenckie Koło Naukowe Technologii i Przetwórstwa Tworzyw).

Przedstawione referaty są dowodem zaawansowanych kompetencji autorów w obszarach, których one dotyczą, a także wskazują, że aktywność w kołach naukowych jest dobrą inwestycją studentów w ich własny kapitał wiedzy i umiejętności na początku karier zawodowych.

*Anna Kuczmaszewska*  
Przewodnicząca komitetów  
naukowego i organizacyjnego



Alicja Bilińska<sup>1</sup>, Bohdan Kosharnyi<sup>2</sup>, Faustyna Smarzewska<sup>3</sup>

## Sieci złożone w analizie bezpieczeństwa na przykładzie siatki terrorystycznej

### Streszczenie

Praca dotyczy problematyki matematycznej analizy wybranych zagadnień związanych z sieciami złożonymi w kontekście ich zastosowań w badaniu oraz prewencji zagrożeń terrorystycznych. Pierwsza część zawiera pojęcia wstępne z zakresu teorii grafów. W dalszej części zostały przedstawione parametry sieciowe, wierzchołkowe, analiza klastrowa oraz typy struktur sieciowych organizacji terrorystycznych.

*Słowa kluczowe: sieć terrorystyczna, centralność wierzchołka, klastrowanie, supersieć, globalność, wydmowość.*

### Wstęp

Sieci złożone są stosunkowo nową koncepcją pozwalającą na analizę rozbudowanych siatek powiązań. Mogą dotyczyć one zarówno zależności między obiektami, jak i ludźmi. O ich początkach można mówić od czasu zaprezentowania w 1959 roku przez Paula Erdősa i Alfreda Rényi'ego artykułu dotyczącego tematyki grafów losowych. Natomiast samo pojęcie „sieciami złożonych” zaczerpnięte zostało od S.N. Dorogovtseva i J.F.F. Mendesa, którzy opublikowali książkę „Evolution of Networks. From Biological Nets to the Internet and WWW”. W literaturze termin ten pojawił się w latach 90. XX wieku, gdy naukowcy – m.in. matematycy, fizycy, informatycy, biologowie i socjologowie – zaczęli modelować oraz analizować rzeczywiste sieci, z którymi mieli styczność podczas badań prowadzonych w swoich dziedzinach [3].

Najbardziej znanym w dziejach atakiem terrorystycznym, w którym zginęło blisko 3 tys. osób, stał się zamach na World Trade Center i Pentagon dokonany 11 września 2001 roku przez organizację islamskich terrorystów – Al Kaidę – będącą międzynarodową siatką [9]. Atak ten ujawnił ówczesne słabości systemu

---

<sup>1</sup> Alicja Bilińska, Studenckie Koło Naukowe „Kwaterion”, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>2</sup> Bohdan Kosharnyi, Studenckie Koło Naukowe „Kwaterion”, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>3</sup> Faustyna Smarzewska, Studenckie Koło Naukowe „Kwaterion”, Wydział Podstaw Techniki, Politechnika Lubelska.

bezpieczeństwa i przyczynił się do największej od 1947 roku reformy amerykańskich służb specjalnych. Zdecydowano wówczas o konieczności stworzenia systemu, który efektywnie zapobiegałby takim zagrożeniom [6].

Jednym z podjętych działań stało się modelowanie powiązań pomiędzy terrorystami za pomocą sieci złożonych, czyli matematycznych struktur określanych jako grafy o nietrywialnych właściwościach topologicznych. Umożliwiają one badanie cech organizacji terrorystycznych oraz ich poszczególnych członków za pomocą analizy parametrów sieciowych i parametrów wierzchołkowych.

Celem przeprowadzonej analizy było odnalezienie zależności i prawidłowości występujących w przykładowej sieci terrorystycznej.

## 1. Definicje

Przedstawione definicje odnoszą się do grafu prostego (nieskierowanego, bez pętli i krawędzi równoległych).

**Definicja 1.1.** [10] Grafem nazywamy parę uporządkowaną  $G = (V; E)$ , gdzie  $V$  oznacza zbiór wierzchołków i  $V \neq \emptyset$  oraz  $E$  oznacza zbiór krawędzi. Krawędzią grafu  $e \in E$  nazywamy dwuelementowy podzbiór zbioru  $V$ , tzn.  $e = \{v_1; v_2\}$ ,  $v_1 \in V$ ,  $v_2 \in V$ .

Liczbę wierzchołków grafu będziemy oznaczali symbolem  $n$ , a liczbę krawędzi symbolem  $q$ .

Mówimy, że krawędź  $e$  jest incydentna z wierzchołkami  $v_1$  i  $v_2$  lub że łączy wierzchołki  $v_1$  i  $v_2$ . Wierzchołki  $v_1$  i  $v_2$  są wierzchołkami końcowymi krawędzi  $e$ .

**Definicja 1.2.** [10] Stopniem  $d(v)$  wierzchołka  $v \in V$  nazywamy liczbę krawędzi incydentnych z wierzchołkiem  $v$ .

**Definicja 1.3.** [10] Drogą w grafie nazywamy skończony ciąg naprzemiennie występujących wierzchołków i krawędzi, rozpoczynający się i kończący wierzchołkami, taki, że każde dwie kolejno po sobie następujące krawędzie mają wspólny wierzchołek.

**Definicja 1.4.** [10] Ścieżką  $n$ -wierzchołkową nazywamy graf, w którym są dwa wierzchołki stopnia pierwszego i  $n-2$  wierzchołków stopnia drugiego.

Jest to struktura łącząca dwa wybrane wierzchołki, gdzie między każdym kolejnym występuje krawędź. Ścieżka jest to szczególnie przypadek drogi, w której wierzchołki nie powtarzają się.

**Definicja 1.5.** [10] Graf nazywamy spójnym, jeżeli pomiędzy każdą parą wierzchołków istnieje łącząca je ścieżka.

Oznacza to, że startując z dowolnego wierzchołka, można osiągnąć każdy inny wierzchołek, poruszając się wzdłuż kolejnych krawędzi.

**Definicja 1.6.** [10] Odległością dwóch wierzchołków nazywamy długość najkrótszej ścieżki łączącej te wierzchołki. Odległość dowolnej pary wierzchołków  $u, v$  oznaczamy jako  $l(u, v)$ .

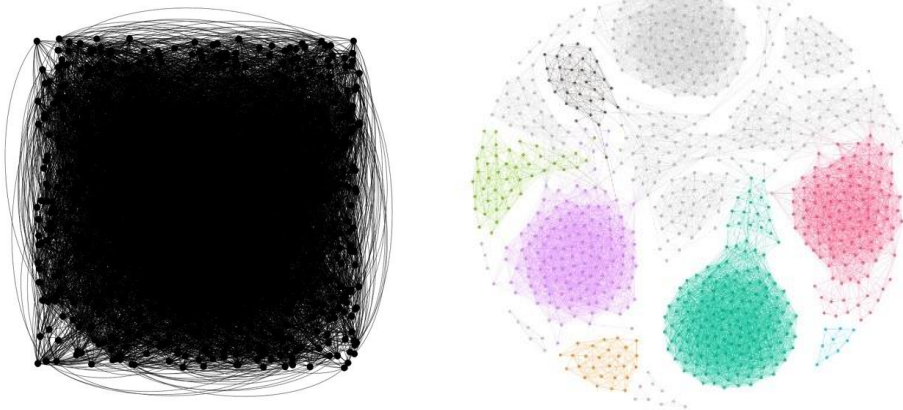
**Definicja 1.7.** [10] Średnicą grafu  $s(G)$  nazywamy największą odległość w grafie, czyli maksymalną wartość spośród największych odległości pomiędzy wierzchołkami:

$$s(G) = \max_{v \in V} \max_{u \in V} l(u, v). \quad (1)$$

## 2. Wybór i analiza sieci jako całości

W wybranej sieci wierzchołki grafu interpretujemy jako osoby fizyczne – terrorystów. Krawędziami są ogólnopojęte bezpośrednie powiązania wynikające m.in ze współpracy, znajomości, wymiany informacji lub wspólnego uczestnictwa w działaniach grupy terrorystycznej. Rozpatrywany graf jest więc nieskierowany, czyli relacja sąsiedztwa pomiędzy wierzchołkami jest symetryczna. W przypadku powiązań pośrednich nie można jednoznacznie stwierdzić istnienia relacji interpersonalnych (przykładowo dwóch terrorystów może kojarzyć się z widzenia, ale mogą się też oni zupełnie nie znać) – jednakże uczestnictwo w konkretnych wydarzeniach lub branie udziału w kluczowych formach aktywności może świadczyć o istnieniu takiej relacji [5].

Rysunek 1 przedstawia omawianą sieć w dwóch różnych układach wierzchołków na płaszczyźnie. Widać stąd, że już odpowiednie wykonanie rysunku grafu pozwala dostrzec pewne występujące w nim zależności.



(a) Graf w losowym układzie wierzchołków

(b) Graf zwizualizowany przy użyciu układu Fruchtermana-Reingolda

**Rysunek 1.** Dwie różne wizualizacje analizowanej sieci w programie Gephi

Źródło: opracowanie własne.

W analizie sieci jako całości można analizować następujące parametry: liczba wierzchołków oraz krawędzi, spójność, gęstość, odległość między wierzchołkami, średnica sieci, średnia długość ścieżki.

W rozpatrywanym grafie liczba wierzchołków, reprezentujących terrorystów, wynosi 881. Liczba krawędzi, czyli powiązań między wszystkimi terrorystami, wynosi 8592.

Rozważany graf jest niespójny, to znaczy, że nie wszystkie pary wierzchołków są połączone ścieżką, a więc nie każdy terrorysta ma powiązanie z innym. Składowymi spójności nazywamy maksymalne podgrafy spójne. Ich liczba oraz wielkość jest obiektem zainteresowania podczas przeprowadzania analizy sieci. Liczba spójnych składowych wynosi 11, co oznacza, że tyle wyróżnić można podgrafów, w których istnieją powiązania między każdym z terrorystów.

**Definicja 2.1.** [10] Gęstością grafu  $D(G)$  nazywamy stosunek liczby jego krawędzi do największej możliwej liczby krawędzi, jakie mogą być rozpięte na wierzchołkach grafu, w danej klasie grafów. Wyraża się wzorem:

$$D(G) = \frac{2q}{n(n-1)}. \quad (2)$$

Gęstość analizowanej sieci wynosi w przybliżeniu 0,022. Wartość ta jest bardzo mała, co odzwierciedla fakt, że w całej sieci nie występuje tak wiele połączeń między terrorystami, jakby mogło występować.

Odległości między wierzchołkami obrazują, ile osób pośredniczy między dwoma wybranymi terrorystami.

Średnica sieci w rozważanym grafie wynosi 14. Oznacza to, że w jednej składowej spójności pomiędzy dwoma najslabiej powiązаныmi terrorystami pośredniczy 13 osób.

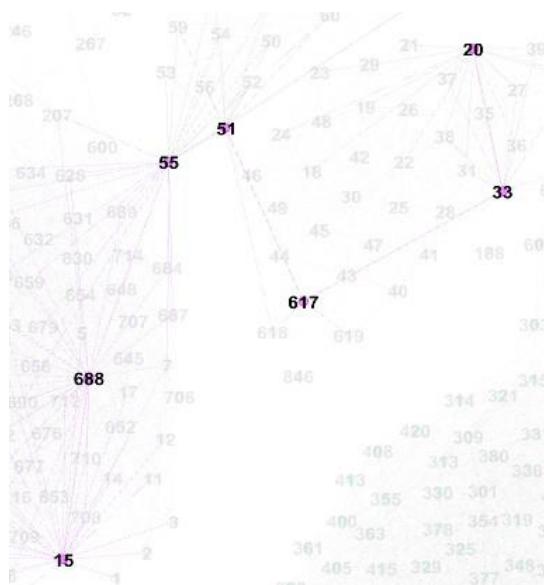
**Definicja 2.2.** [10] Średnią długością ścieżki nazywamy liczbę określoną wzorem:

$$l(G) = \frac{2}{n(n-1)} \sum_{u=1}^n \sum_{v>u}^n l(u, v), \quad (3)$$

gdzie  $l(u, v)$  jest długością najkrótszej ścieżki łączącej wierzchołki  $u$ -ty z  $v$ -tym.

Inaczej mówiąc,  $l(G)$  jest średnią arytmetyczną długości najkrótszych ścieżek między wszystkimi parami wierzchołków.

W całej sieci średnio każdy terrorysta jest połączony z innym dzięki 5 osobom.



Rysunek 2. Ścieżka o długości 6 znaleziona przy pomocy języka programowania R oraz zwizualizowana wykorzystaniem programu Gephi

Źródło: opracowanie własne.

### 3. Analiza poszczególnych parametrów wierzchołków

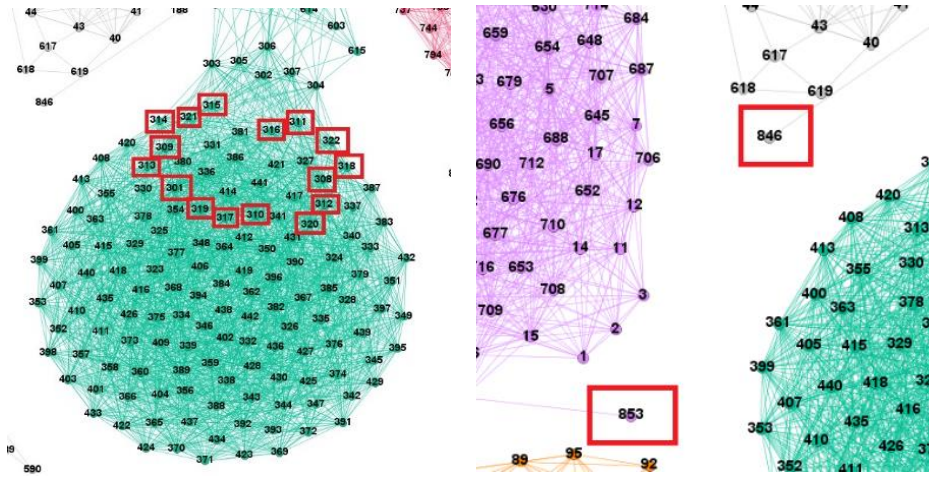
Mamy do analizy następujące parametry wierzchołkowe: stopień, pośrednictwo, bliskość, centralność wektora własnego, współczynnik klastrowania i ekscentryczność.

#### 3.1. Stopień wierzchołka

**Definicja 3.1.** [10] Średnim stopniem wierzchołka grafu nazywamy liczbę  $d_{sr} = \frac{2q}{n}$ .

Średni stopień wierzchołków wynosi 19,5, co oznacza, że średnio każdy terrorysta współpracuje z 19–20 innymi terrorystami. Maksymalny stopień wynosi 36, z czego wiemy, że terroryści nr 126, 301, 308–322 są „najpopularniejsi” – mają bezpośredni kontakt z 36 innymi terrorystami. Minimalny stopień wierzchołków to 1, a co za tym idzie, terroryści nr 845, 846, 851 i 853 mają kontakt tylko z jedną osobą im podobną. Można więc założyć, że są nowi w całej strukturze. Na pewno zaś mają w niej najmniejsze znaczenie.





(a) Wierzchołki o największym stopniu

(b) Wierzchołki o najmniejszym stopniu

**Rysunek 3. Wierzchołki o ekstremalnych stopniach w rozpatrywanej sieci**

Źródło: opracowanie własne.

### 3.2. Centralność wierzchołka

Jednym z fundamentów analizy sieci jest koncepcja centralności, czyli względnego znaczenia pojedynczego członka w grupie. Miary centralności są kluczowe dla zidentyfikowania bardziej wpływowych wierzchołków w sieci. Istnieje wiele miar centralności charakteryzujących nieco inne aspekty położenia wierzchołka w sieci [2]. Jedną z najczęściej analizowanych jest pośrednictwo.

**Definicja 3.2.** [10] Pośrednictwem (ang. *betweenness centrality*)  $b(v)$  nazywamy liczbę określoną wzorem

$$b(v) = \sum_{u \neq v \neq w} \frac{\sigma_{uw}(v)}{\sigma_{uw}}, \quad (4)$$

gdzie  $\sigma_{uw}$  jest liczbą najkrótszych ścieżek pomiędzy  $u$  oraz  $w$ , a  $\sigma_{uw}(v)$  jest liczbą tychże ścieżek, które przechodzą przez  $v$ .

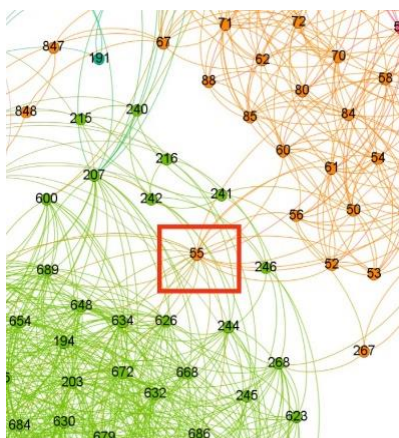
Średnie pośrednictwo w analizowanej sieci wynosi 1383,18. Maksymalne – 87384,16 – posiada terrorysta nr 55, tzn., że jest głównym pośrednikiem w tej sieci. Gdybyśmy chcieli w jednym kroku utrudnić współpracę terrorystów, to najłatwiej byłoby usunąć tego terrorystę, żeby rozbić graf. W realnym życiu oznaczałoby to, że centrum ich komunikacji zostało zrujnowane.

Kolejnym często analizowanym parametrem centralności jest bliskość.

**Definicja 3.3.** [7] Bliskość (ang. *closeness centrality*) wierzchołka  $v$  dana jest wzorem:

$$C(v) = \frac{n-1}{\sum_{u \in V} l(u,v)}. \quad (5)$$

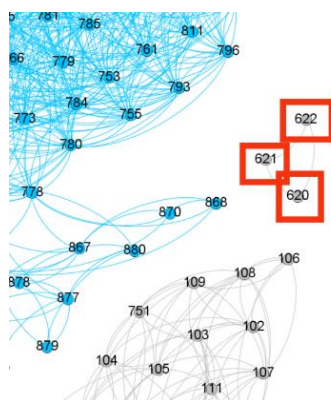
Wedle tej miary wierzchołek jest tym bardziej centralny, im jest średnio bliżej wszystkich innych wierzchołków sieci.



**Rysunek 4. Wierzchołek z największym pośrednictwem odnaleziony za pomocą programu Gephi**

Źródło: opracowanie własne.

Średnia bliskość wynosi 0,0129. Maksymalną bliskość 0,5 mają terroryści nr 589–594, 620–622, 722–724, 731–733, co znaczy, że są najbliżej wszystkich innych (w obrębie swojej społeczności), a więc mogą szybciej przekazać informacje innym.



**Rysunek 5. Wierzchołki z największą bliskością odnalezione za pomocą programu Gephi**

Źródło: opracowanie własne.

Aby lepiej zrozumieć kolejny parametr centralności, warto poznać najpierw definicję macierzy sąsiedztwa.

**Definicja 3.4.** [4] Macierzą sąsiedztwa nazywamy macierz kwadratową indeksowaną wierzchołkami grafu, której elementy przyjmują wartości 0 lub 1. Jedyneką na przecięciu  $u$ -tego wiersza i  $v$ -tej kolumny informuje o tym, że w grafie istnieje krawędź pomiędzy wierzchołkami  $u$  oraz  $v$ , zaś zero oznacza, że takiej krawędzi nie ma.

Centralność wektora własnego to miara tego, jak dobrze połączony jest węzeł z innymi dobrze połączonymi węzłami w sieci. Można go wykorzystywać do wiarygodnego porównywania różnych sieci, ze względu na to, że służy on jako miara łączności. Uwzględnia on zarówno liczbę połączeń danego wierzchołka z innymi, jak i jego znaczenie pod względem przepływu informacji.

**Definicja 3.5.** [2] Centralnością wektora własnego (ang. *eigenvector centrality*) nazywamy ważoną sumę centralności wszystkich wierzchołków, które są z nim połączone krawędzią

$$c(u) = \lambda^{-1} \sum_{v=1}^n A_{uv} c(v), \quad (6)$$

gdzie  $c(v)$  jest wektorem własnym związanym z wartością własną  $\lambda$  w macierzy sąsiedztwa grafu  $A$ .

Średnia centralność wektora własnego rozpatrywanej sieci jest równa 0,13. Maksymalną 1 – posiadają terroryści nr 308, 310, 316–319, a więc współpracują z dużą liczbą terrorystów, którzy też mają wysoką centralność wektora własnego.

**Definicja 3.6.** [10] Ekscentrycznością  $ex(v)$  wierzchołka  $v \in V$  nazywamy wartość największej odległości  $v$  od dowolnego innego wierzchołka grafu:

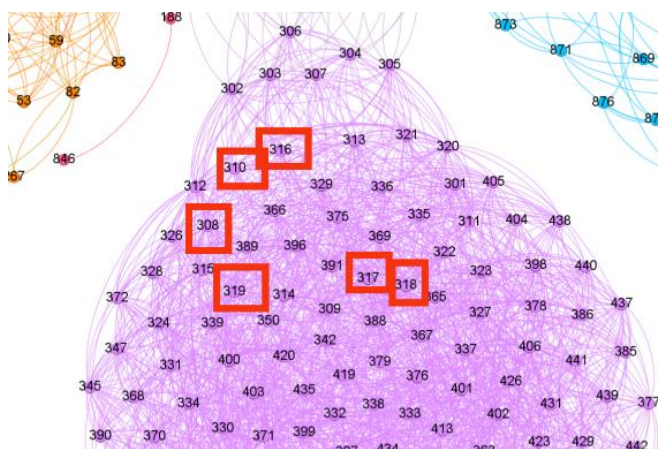
$$ex(v) = \max_{u \in V} l(u, v). \quad (7)$$

Zauważmy, że teraz można sformułować alternatywną definicję średnicy.

**Definicja 3.7.** [10] Średnicą grafu  $s(G)$  nazywamy maksymalną ekscentryczność spośród zbioru wierzchołków grafu:

$$s(G) = \max_{v \in V} ex(v). \quad (8)$$

Minimalną ekscentryczność równą 1, posiadają terroryści nr 589–594, 620–622, 722–724, 731–733, co oznacza że znajdują się w centrum badanej sieci, a maksymalną – 14 – terroryści nr 323–442, 550, 552, 555, a zatem znajdują się oni na peryferiach.



**Rysunek 6.** Wierzchołki z największą centralnością wektora własnego odnalezione za pomocą programu Gephi

Źródło: opracowanie własne.

### 3.3. Współczynnik klastrowania

Współczynnik klastrowania informuje o tym, w jakim stopniu sąsiedzi danego wierzchołka łączą się ze sobą.

**Definicja 3.8.** [1] Współczynnikiem klastrowania wierzchołka  $u$  o stopniu  $d(u)$  nazywamy liczbę

$$C(u) = \frac{2L(u)}{d(u)(d(u)-1)}, \quad (9)$$

gdzie  $L(u)$  reprezentuje liczbę połączeń między sąsiadami wężła  $u$ .  $C(u)$  przyjmuje wartości od 0 do 1.

Średnio wynosi on 0,001453. Maksymalna wartość jest równa 0,4545, tzn., że terroryści nr 4 i 692 mają najbardziej gęste sąsiedztwo, a więc mają największy wpływ na powstawanie nowych połączeń.

## 4. Analiza klastrowa

Analiza klastrowa służy do wyodrębniania społeczności w sieciach złożonych. Istnieje wiele różnych algorytmów dokonujących tego podziału – w zależności od użytego, nie jest on jednoznaczny. Aby dowiedzieć się, na jakie społeczności może zostać podzielony graf, przeprowadzona została analiza klastrowa obejmująca algorytmy: wiodącego wektora własnego, pośrednictwa krawędzi, infomap, label propagation, walktrap, szybki zachłanny, wielopoziomowy.

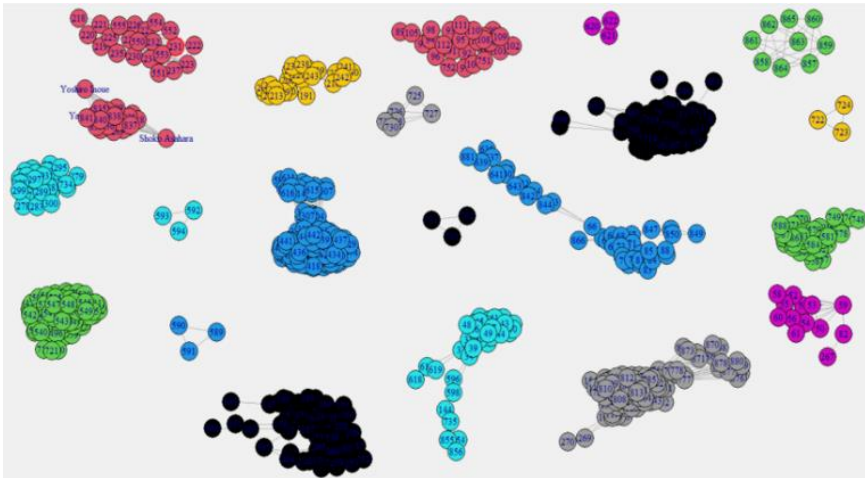
Parametrem charakteryzującym jakość podziału jest modułowość. Określa ona różnicę pomiędzy liczbą krawędzi występujących w proponowanej społeczności

a wartością oczekiwaną tej liczby, gdyby występowanie krawędzi w grafie podlegało rozkładowi jednostajnemu. Z tabeli 1 widać, że podział jest niejednoznaczny.

**Tabela 1. Liczba społeczności i modułowość w poszczególnych algorytmach**

Algorytm	Liczba społeczności	Modułowość
wiodącego wektora własnego	21	0,808083
pośrednictwa krawędzi	21	0,817694
infomap	30	0,81292
label propagation	28	0,773203
walktrap	26	0,816549
szybki zachłanny	16	0,783496
wielopoziomowy	18	0,817131

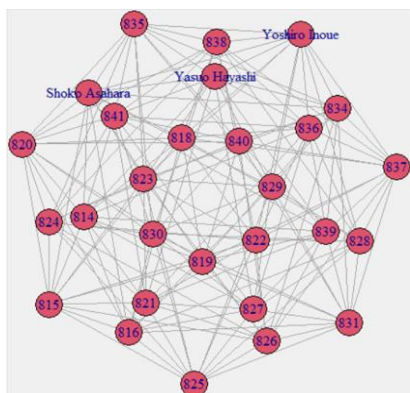
Źródło: opracowanie własne.



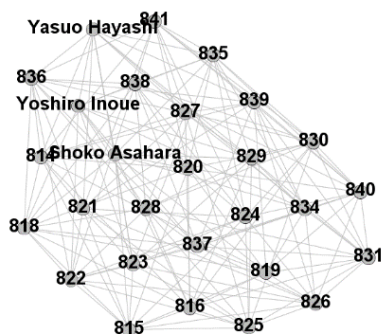
**Rysunek 7. Przykładowy podział sieci na 21 społeczności uzyskany za pomocą algorytmu wiodącego wektora własnego**

Źródło: opracowanie własne.

Wyodrębniając w grafie społeczności na podstawie algorytmów z tabeli 1, udało się zaobserwować, że pewne klastry będą pojawiały się zawsze. Przykładem jest społeczność przedstawiona na rysunku 8.



(a) Wizualizacja utworzona przy pomocy języka programowania R



(b) Wizualizacja w programie Gephi

**Rysunek 8. Klaster obrazujący strukturę sekty Aum Shinrikyō składający się z wierzchołków nr 814–841, w tym: Shoko Asahara, Yasuo Hayashi, Yoshiro Inoue**

Źródło: opracowanie własne.

## 5. Analiza struktur sieciowych organizacji terrorystycznych

Poniższy fragment został opracowany na podstawie [8]. Analiza struktur sieci terrorystycznej obejmuje 3 poziomy: lokalny, regionalny i globalny. Każdy z tych typów charakteryzuje się własnym poziomem zorganizowania i elastycznością jego zmiany (skalowania – czyli rozrostu lub degradacji), która zależy zarówno od czynników wewnętrznych, jak i zewnętrznych.

### 5.1. Sieci lokalne

Sieci lokalne nazywamy sieciami pierwszego poziomu. Są to najmniejsze struktury wchodzące w skład siatek terrorystycznych. Wyróżniamy spośród nich 3 typy, którymi są sieć liniowa, sieć odśrodkowa oraz sieć wielokanałowa. Ich ilustracje widać na rysunku 9.

Sieć liniowa charakteryzuje grupy, które oprócz udziału w akcjach terrorystycznych podejmują się też działalności przestępczej, zajmując się m.in. przemytem.

Sieć odśrodkowa dotyczy tych obszarów, gdzie występuje potrzeba koordynowania działań poprzez węzeł centralny, którym może być np. dowódca lokalnej społeczności.

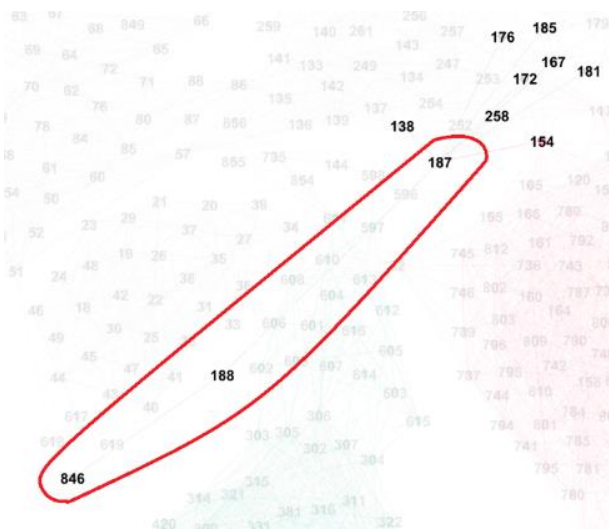
Sieć wielokanałowa występuje tam, gdzie każdy węzeł jest połączony z pozostałymi, co oznacza, że komunikacja w takim ugrupowaniu jest bardzo dobrze rozwinięta.



**Rysunek 9. Typy sieci lokalnych**

Źródło: A. Wejkszner, *Globalna sieć Al-Kaidy. Nowe państwo islamskie?*, Difin, 2017, s. 124–129, <https://depot.ceon.pl/handle/123456789/12810> (dostęp: 12.11.2022).

W analizowanym grafie występują wszystkie typy sieci lokalnych. Dla przykładu wierzchołki nr 846, 187–188, zaznaczone na rysunku 10 przed rozrostem sieci, tworzyły sieć liniową na poziomie lokalnym. Oznacza to, że mógł tu występować np. łańcuch przemytniczy, który następnie połączył siły z daną komórką terrorystyczną.

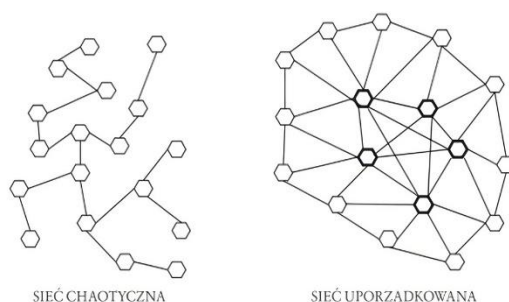


**Rysunek 10. Przykład sieci liniowej w rozpatrywanym grafie**

Źródło: opracowanie własne.

## 5.2. Sieci regionalne

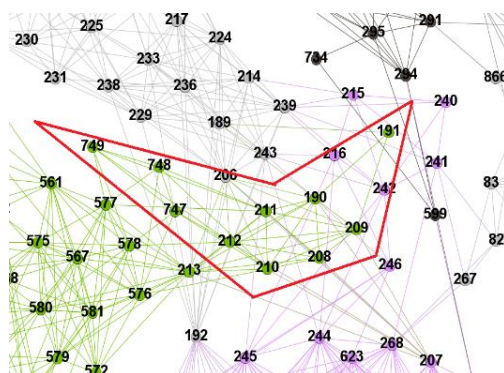
Sieci regionalne nazywamy sieciami drugiego poziomu. Posiadają one ograniczone rozmiary oraz charakteryzują się wytwarzaniem skomplikowanych powiązań pomiędzy węzłami. Przyjmują one postać chaotyczną (nieuporządkowaną) lub uporządkowaną. Sieci uporządkowane posiadają węzły o dużym znaczeniu dla grupy nazywane podstawowymi, które wytwarzają najliczniejsze i jednocześnie jedne z najważniejszych powiązań (rys. 11).



**Rysunek 11. Typy sieci regionalnych**

Źródło: A. Wejkszner, *Globalna sieć Al-Kaidy. Nowe państwo islamskie?*, Difin, 2017, s. 124–129, <https://depot.ceon.pl/handle/123456789/12810> (dostęp: 12.11.2022).

W tak wyodrębnionym klastrze analizowanej sieci, widocznym na rysunku 12, zauważyć można, jak sieć wielokanałowa, tj. sieć pierwszego poziomu, zamienia się powoli w sieć uporządkowaną, tzw. sieć drugiego poziomu. Zauważyć można, że wierzchołki numer 190, 208–212 są węzłami podstawowymi, wytwarzającymi najistotniejsze i najliczniejsze powiązania w rozpatrywanej społeczności.



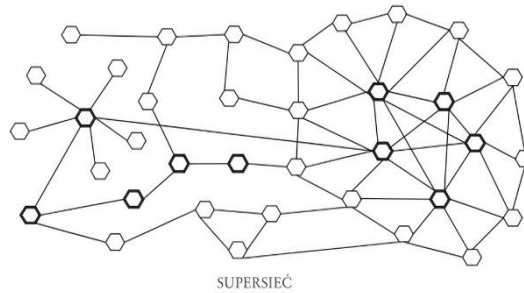
**Rysunek 12. Przykład sieci uporządkowanej w rozpatrywanym grafie**

Źródło: opracowanie własne.



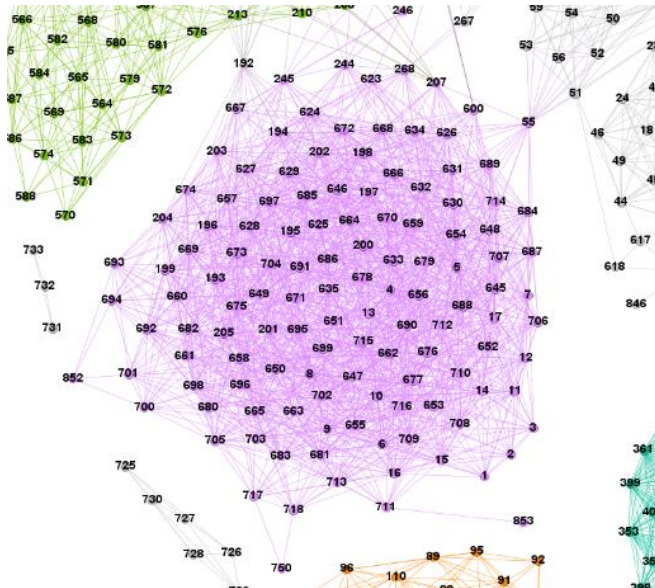
### 5.3. Sieci globalne

Sieci globalne nazywamy sieciami trzeciego poziomu. Składają się one z wielu sieci lokalnych oraz regionalnych, a więc z wielu sieci zarówno pierwszego, jak i drugiego poziomu. Charakteryzują się największym wewnętrznym zróżnicowaniem, możliwością dynamicznego modyfikowania swojej struktury w zależności od występujących potrzeb oraz zaawansowaną wewnętrzną samoorganizacją o globalnym zasięgu.



Rysunek 13. Sieć globalna

Źródło: A. Wejkszner, *Globalna sieć Al-Kaidy. Nowe państwo islamskie?*, Difin, 2017, s. 124–129, <https://depot.ceon.pl/handle/123456789/12810> (dostęp: 12.11.2022).



Rysunek 14. Przykład supersieci w rozpatrywanym grafie

Źródło: opracowanie własne.

Współczesne sieci terrorystyczne cechują się naturalnym rozrostem oraz zdolnością do dostosowywania się do panujących na danym terytorium warunków. Mogą one ewoluować i pokrywać każdy dostępny obszar rozumiany zarówno jako położenie geograficzne, jak też środowisko społeczne bądź też kulturowe. Połączenia, które wytwarzają pomiędzy węzłami, zaliczają się do najbardziej efektywnych – są to powiązania występujące w sieciach odśrodkowych i wielokanałowych.

Przez ostatnie dziesięciolecia struktura organizacji terrorystycznych uległa znacznej modyfikacji. Na przełomie lat 80. i 90. XX wieku większość istniejących wówczas grup działała w porządku hierarchicznym, natomiast obecnie struktury hierarchiczne współegzystują z sieciowymi.

#### 5.4. Glokalność

Glokalność jest pojęciem dotyczącym procesów, które występują zarówno na płaszczyźnie globalnej, jak i lokalnej. Glokalność organizacji terrorystycznych można obserwować na płaszczyźnie socjologiczno-ekonomicznej. Struktury takie wykazują pełną autonomię, działając według własnych rozwiązań oraz funkcjonują na ograniczonym terytorium, jednocześnie tworząc globalny ruch społeczny.

**Tabela 2. Kluczowe cechy glokalnej struktury organizacyjnej organizacji terrorystycznej**

Cecha	Istota
Instytucjonalizacja	Niewielki lub niezauważalny proces instytucjonalizacji struktur na poziomie lokalnym, dynamicznie zmieniający się charakter struktur globalnych.
Autonomia	Ograniczona lub pełna autonomia struktur lokalnych działających niezależnie od siebie, ale walczących o realizację wspólnych strategicznych celów.
Współwystępowanie interesów lokalnych i globalnych	Pojawienie się lokalnych struktur organizacji w wielu miejscach w tym samym czasie pozwala na łączenie wielu lokalnych interesów z interesami globalnymi danej struktury.
Łańcuch dowodzenia i komunikacja	Istnieje możliwość komunikowania się, a nawet otrzymywania zadań operacyjnych od innych węzłów w globalnej strukturze. W żaden sposób nie ogranicza to jednak możliwości swobodnej rozbudowy struktury i aktywności grupy na poziomie lokalnym.

Źródło: A. Wejkszner, *Globalna sieć Al-Kaidy. Nowe państwo islamskie?*, Difin, 2017, s. 124–129, <https://depot.ceon.pl/handle/123456789/12810> (dostęp: 12.11.2022).

### 5.5. „Wydumowość”

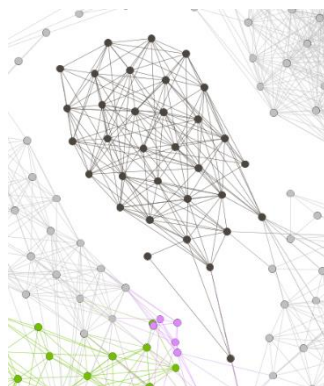
Autorzy S. Mishal i M. Rosenthal zaproponowali wyróżnienie struktur o „wydumowym” charakterze. „Wydumowość” jest to metafora obrazująca proces swobodnego przechodzenia pomiędzy stanem sprawowania kontroli nad określonym obszarem a stanem niewidocznej działalności. Ideą jest deterytorializacja panującego porządku politycznego. Praktycznie charakteryzuje to sytuację, w której poszczególne grupy terrorystyczne uaktywniają się na danym terytorium, poprzez swoje działania prowadzą do określonych zmian (np. społecznych, politycznych) na zajętej obszarze, po czym przechodzą na inny.

Struktury wydumowe wyróżnia wysoki poziom wewnętrznej zmienności. W swojej aktywności wykorzystują one tymczasowe struktury lokalne, by osiągać zamierzone cele. Najczęściej organizacja tych grup wyróżnia się sieciowością oraz ma rozbudowaną siatkę powiązań pozbawioną tylko jednego centrum dowodzenia. Aktywność struktur „wydumowych” może mieć zasięg regionalny, subregionalny lub globalny.

**Tabela 3. Kluczowe cechy „wydumowej” struktury organizacyjnej organizacji terrorystycznej**

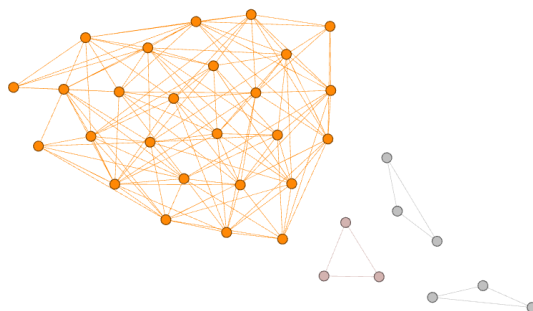
Cecha	Istota
Brak związków z państwem	Organizacje wydumowe nie posiadają swej terytorialnej manifestacji. Trudno je monitorować.
Niewielki zakres lub brak instytucjonalizacji	Brak stałych manifestacji instytucjonalnych. Organizacyjna struktura tych organizacji zmienia się w dynamiczny sposób.
Brak trwałych relacji z innymi organizacjami terrorystycznymi	Tego typu organizacje, z uwagi na zakres i istotę swojej aktywności, nie wchodzi w trwałe relacje z innymi organizacjami terrorystycznymi.
Brak jasno określonych struktur i łańcuchów dowodzenia oraz komunikacji	Organizacje wydumowe nie utrzymują rozbudowanych, zhierarchizowanych łańcuchów komunikacji i dowodzenia, te zaś, które istnieją, mogą być w każdej chwili zrekonstruowane.
Łatwość łączenia różnych interesów i partycypowania w różnych konfliktach	Pojawienie się lokalnych struktur organizacji w wielu miejscach w tym samym czasie pozwala na łączenie wielu lokalnych interesów z interesami strategicznymi.
Cel strategiczny jako substytut afiliacji terytorialnej	Istnienie wspólnego strategicznego celu (np. utworzenia kalifatu o globalnym zasięgu) może być substytutem dla afiliacji terytorialnej w wymiarze lokalnym.

Źródło: A. Wejkszner, *Globalna sieć Al-Kaidy. Nowe państwo islamskie?*, Difin, 2017, s. 124–129, <https://depot.ceon.pl/handle/123456789/12810> (dostęp: 12.11.2022).



Rysunek 15. Przykład struktury globalnej w rozpatrywanym grafie

Źródło: opracowanie własne.



Rysunek 16. Przykład struktury „wydmowej” w rozpatrywanym grafie

Źródło: opracowanie własne.

## 6. Podsumowanie

Z uwagi na możliwość powiązania struktury sieci z występującymi w niej w rzeczywistości cechami, jej analiza może służyć do poznawania istniejących zależności.

Dzięki temu – poprzez analizowanie sieci terrorystycznej pod kątem parametrów sieciowych i wierzchołkowych – w większym stopniu można zrozumieć działanie poszczególnych jej fragmentów oraz członków, co pozwala na efektywniejsze działania prewencyjne i wywiadowcze.

Przykładowo, wiedząc, które wierzchołki w sieci charakteryzują się największą centralnością wektora własnego, jesteśmy w stanie wskazać, którzy członkowie są prawdopodobnie najbardziej wpływowi – i to ich utrata mocno dezorganizowałaby działanie całej siatki. Z kolei, gdy wiemy, które wierzchołki posiadają największe pośrednictwo, umożliwia nam to wskazanie osób mogących

posiadać najwięcej interesujących nas informacji – i to kontakt z nimi byłby najbardziej wskazany.

Podsumowując, przedstawione przez nas narzędzia teorii grafów oraz własności różnych typów sieci terrorystycznych mogą być wykorzystywane do efektywniejszego opisu, analizy oraz zrozumienia działających struktur w celu minimalizowania zagrożeń płynących z działalności terrorystycznej.

## Literatura

- [1] Barabási A. L., *Network Science*, Cambridge University Press, 2016, <http://networksciencebook.com/> (dostęp: 18.11.2022).
- [2] Negre C. F. A., Morzan U. N., Hendrickson H. P., Pal R., Lisi G. P., Loria J. P., Rivalta I., Ho J., Batista V. S., *Eigenvector centrality for characterization of protein allosteric pathways*, “Proceedings of the National Academy of Sciences” 2018, vol. 115, no. 52, pp. E12201–E12208, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6310864/> (dostęp: 18.11.2022).
- [3] Pawełczyk M., Skrobacki Z., *Sieci złożone jako narzędzie modelowania systemów transportowych i logistycznych*, <https://www.logistyka.net.pl/bank-wiedzy/item/87709> (dostęp: 20.11.2022).
- [4] Rachelski M., *Macierz sąsiedztwa*, <http://smurf.mimuw.edu.pl/node/375> (dostęp: 26.11.2022).
- [5] Rossi R., Ahmed N., *The Network Data Repository with Interactive Graph Analytics and Visualization*, <https://networkrepository.com/TerroristRel.php> (dostęp: 26.04.2022).
- [6] Szymańska M., *Eksplozja terroryzmu w USA a przekształcenia w amerykańskich służbach specjalnych po 11 września 2001 r.*, „Przegląd Bezpieczeństwa Wewnętrznego” 2014, t. 6, nr 10, ss. 186–210, <https://www.abw.gov.pl/pl/pbw/publikacje/przegladbezpieczenstwa-1/1008,Przeglad-Bezpieczenstwa-Wewnetrznego-nr-10-6-2014.html> (dostęp: 12.11.2022).
- [7] Tarapata Z., *Czy sieci rządzą światem? Od Eulera do Barabasięgo*, „Biuletyn Instytutu Systemów Informacyjnych” 2012, t. 10, ss. 31–51, [https://isi.wat.edu.pl/sites/default/files/sekretariat/Biul\\_ISI/biul\\_isi\\_10\\_20\\_12.pdf](https://isi.wat.edu.pl/sites/default/files/sekretariat/Biul_ISI/biul_isi_10_20_12.pdf) (dostęp: 12.11.2022).
- [8] Wejkszner A., *Globalna sieć Al-Kaidy. Nowe państwo islamskie?*, Difin, 2017, <https://depot.ceon.pl/handle/123456789/12810> (dostęp: 12.11.2022).
- [9] Kean T. H., Hamilton L., *The 9/11 Commission report: final report of the National Commission on Terrorist Attacks upon the United States*, National Commission on Terrorist Attacks upon the United States, Washington 2004.
- [10] Wojciechowski J., Piękosz K., *Grafy i sieci*, Wydawnictwo Naukowe PWN, Warszawa 2013.

## **Complex networks in analysis of security on the example of a terrorist network**

### **Abstract**

The work concerns the issues of mathematical analysis of selected topics related to complex networks in the context of their applications in the research on terrorist threats and prevention from them. The first part contains introductory terms of graph theory. In the following part there are presented such concepts as: network parameters, vertex parameters, cluster analysis and terrorist organizations' types of network structures.

*Keywords: terrorist network, vertex centrality, clustering, supernetwork, glocality, duneiness.*



## **Wybrane zagadnienia związane z kwantylami**

### **Streszczenie**

Kwantyle są jednym z ważniejszych i często wykorzystywanych parametrów rozkładów badanej cechy czy zmiennej losowej. W rozdziale przedstawione zostaną podstawowe pojęcia i informacje związane z kwantylami. Wprowadzony zostanie model miary Value at Risk, która jest bezpośrednim przykładem zastosowania kwantyli w finansach i ekonomii. Następnie zaprezentowana zostanie specjalna reprezentacja kwantyla z próby (estymatora kwantyla rozkładu rzeczywistego), tzw. reprezentacja Bahadura, która jest bardzo istotna ze względu na badanie granicznych własności kwantyla z próby. Dodatkowo przedstawiona zostanie symulacja weryfikująca pewne graniczne wnioski związane z kwantylem z próby.

*Słowa kluczowe: kwantyl, zależne zmienne losowe, reprezentacja Bahadura kwantyla z próby.*

### **Wstęp**

We współczesnych badaniach stale rośnie zapotrzebowanie na analizy statystyczne. Odgrywają one bardzo ważną rolę w procesach gromadzenia, analizy i interpretacji danych. Często spotyka się sytuację, w których analizy statystyczne oparte na klasycznych parametrach i ich estymatorach nie pozwalają na wyciągnięcie poprawnych wniosków. W takich sytuacjach pomocne okazują się statystyki pozycyjne. Jest to grupa statystyk wyznaczanych na podstawie uporządkowanych prób losowych, wykorzystywane są one w konstrukcji parametrycznych oraz nieparametrycznych estymatorów.

Do podstawowych statystyk pozycyjnych zalicza się kwantyle z próby. Jednym z często spotykanych kwantyli jest mediana. Mediana z próby okazuje się pomocna przy szacowaniu wartości średniej, gdy rozkład populacji jest np. asymetryczny. Wykazuje ona większą stabilność niż średnia arytmetyczna, na którą mają mocny wpływ wartości ekstremalne.

Dodatkowo, w badaniach statystycznych często rozważanym problemem jest brak znajomości rozkładu badanej cechy mierzalnej. W efekcie zachodzi potrzeba oszacowania danych parametrów tego rozkładu. Do szacowania wykorzystuje się wspomniane już estymatory, czyli parametry obliczone dla próby badawczej i na ich podstawie szacuje się prawdziwą wartość parametru w populacji. Kwantyle

---

<sup>1</sup> Mgr inż. Dagmara Dudek, Katedra Matematyki Stosowanej, Wydział Podstaw Techniki, Politechnika Lubelska, dagmara.dudek@pollub.pl.



rozkładu empirycznego są bardzo istotnym parametrem wykorzystywanym w badaniach statystycznych, kiedy nie znamy rzeczywistego rozkładu badanej cechy. Wykorzystywane do pomiarów ryzyka rynkowego, finansowego i operacyjnego. Okazują się również świetnym narzędziem w analizach dochodów oraz zjawisk, które występują rzadko i których pojawienie się generuje duże straty finansowe.

Celem niniejszej pracy jest:

- przedstawienie zagadnień teoretycznych związanych z kwantylami,
- przedstawienie modelu miary VaR, gdzie zastosowanie znajdują kwantyle,
- omówienie specjalnej, asymptotycznej reprezentacji kwantyla z próby, tzw. reprezentacji Bahadura, która jest istotna, m.in. ze względu na badanie granicznych właściwości kwantyli z próby,
- zaprezentowanie symulacji, która zilustruje pewne graniczne właściwości kwantyla z próby.

## 1. Kwantyl jako ważny parametr rozkładu

### 1.1. Pojęcia ogólne

Jak wspomniano wcześniej, kwantyle rozkładu cechy mierzalnej (lub zmiennej losowej) są podstawowymi pozycyjnymi miarami położenia rozkładu. Kwantyle są to wartości cechy zmiennej, które dzielą badaną zbiorowość na określone części pod względem liczby jednostek.

**Definicja 1.1.** Kwantyl rzędu  $p \in (0,1)$  zmiennej losowej  $X$  przyjmującej wartości rzeczywiste definiujemy jako

$$Q(p) = \inf\{x: P(X \leq x) \geq p\}. \quad (1)$$

Definicja 1.1 określa kwantyl w sposób jednoznaczny.

Należy zauważyć, że kwantyl  $Q(p)$  ma następującą własność:

$$P(X \leq Q(p)) \geq p, P(X \geq Q(p)) \geq 1 - p, \quad (2)$$

mimo że liczb spełniających nierówności (1) może być nieskończenie wiele.

Do najczęściej stosowanych kwantyli należą:

- mediana,
- kwartyle,
- decyle,
- centyle.

Jednym z popularniejszych kwantyli jest wspomniana wyżej mediana. Mediana, to wartość cechy zmiennej, która dzieli badaną zbiorowość na dwie części w taki sposób, że połowa jednostek zbiorowości charakteryzuje się wartościami nie wyższymi, a połowa nie niższymi od mediany. Medianą zmiennej losowej  $X$  nazywamy liczbę  $x$  spełniającą nierówności:

$$P(X \leq x) \geq \frac{1}{2}, \quad P(X \geq x) \geq \frac{1}{2}. \quad (3)$$

**Przykład 1.1.** Zmienna losowa dyskretna ma rozkład przedstawiony w tabeli 1.

**Tabela 1. Rozkład zmiennej losowej dyskretnej**

$x_i$	1	2	3	4
$P(X = x_i)$	0,1	0,3	0,5	0,1

Źródło: opracowanie własne.

Medianą w tym przykładzie jest liczba 3, ponieważ spełnia ona łącznie obie nierówności (3):

$$P(X \leq 3) = \frac{9}{10} > \frac{1}{2} \text{ oraz } P(X \geq 3) = \frac{3}{5} > \frac{1}{2}. \quad (4)$$

Kwartyle dzielimy na kwartył pierwszy (dolny), kwartył drugi (mediana) oraz kwartył trzeci (górnny). Każdy z kwartyli dzieli zbiorowość uporządkowaną na dwie części, przy czym:

- kwartył pierwszy dzieli zbiorowość na dwie części w ten sposób, że 25% jednostek zbiorowości ma wartości zmiennej mniejsze lub równe kwartyłowi pierwszemu, a 75% – równe lub większe od tego kwartyła,
- kwartył drugi (mediana) dzieli zbiorowość na dwie części w ten sposób, że 50% jednostek ma wartości mniejsze lub równe medianie oraz 50% – równe lub większe od mediany,
- kwartył trzeci dzieli zbiorowość na dwie części w ten sposób, że 75% jednostek ma wartości zmiennej mniejsze lub równe kwartyłowi trzeciemu, a 25% – równe lub większe od kwartyła trzeciego.

Decyle i centyle wyznacza się podobnie jak kwartyle. Decyle dzielą zbiorowość uporządkowaną na 10 części pod względem liczebności. Np. decyl czwarty to taka wartość cechy, że 40% wszystkich jednostek zbiorowości ma wartości mniejsze lub równe w stosunku do niej, a 60% równe lub wyższe. Decyli jest 9, a piąty decyl jest medianą.

Centyle (percentyle) dzielą zbiorowość uporządkowaną na 100 części pod względem liczebności. Na przykład czterdziesty dziewiąty centyl jest taką wartością, że 49% wszystkich jednostek badanej zbiorowości ma wartości od niej niższe lub równe, a 51% jednostek – wartości równe lub wyższe. Centyli jest 99, a pięćdziesiąty centyl jest równy medianie.

## 1.2. Zastosowanie kwantyli

Kwantyle są bardzo ważnym narzędziem w wielu dziedzinach nauki, w tym w ekonomii oraz finansach. W 1994 roku instytucja finansowa JP Morgan opublikowała metodę zarządzania ryzykiem służącą m.in. do wyznaczania miary

*VaR*, czyli *Value at Risk*. Obecnie *VaR* jest najpopularniejszym rozwiązaniem w przypadku pomiaru ryzyka rynkowego, ale jest również stosowana do innych rodzajów ryzyka, np. kredytowego i operacyjnego. Dodatkowo, miara *VaR* jest podstawą dla wielu innych miar, które służą do analizowania ryzyka przedsiębiorstwa. Miara ta jest zdefiniowana na podstawie definicji kwantyla rozkładu.

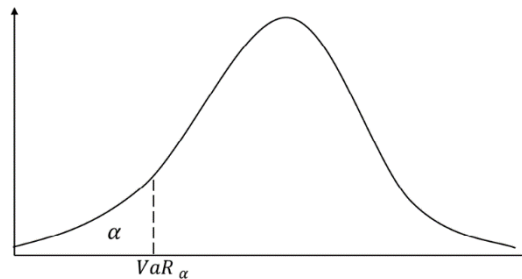
Niech  $\{X_i, 1 \leq i \leq n\}$  będzie wartością rynkową aktywów notowanych w wybranej jednostce czasu i obserwowanych w  $n$  okresach. Niech

$$Y_i = \log \left( \frac{X_i}{X_{i-1}} \right), \quad (5)$$

będzie logarytmiczną stopą zwrotu. Przyjmijmy, że  $\{Y_i, 1 \leq i \leq n\}$  jest ściśle stacjonarnym zależnym procesem z rozkładem określonym poprzez funkcję dystrybuanty  $F$ . Niech  $p$  będzie dodatnią, bliską zera liczbą. Wtedy poziom  $1-p$  miary *VaR* wynosi

$$VaR_p = \inf\{x: F(x) \geq p\}. \quad (6)$$

Miarę *Value at Risk* ilustruje rysunek 1.



**Rysunek 1. Graficzna interpretacja miary *VaR***

Źródło: M. Chlebus, *Pomiar ryzyka rynkowego za pomocą miary Value at Risk – podejście dwuetapowe*, Uniwersytet Warszawski, Warszawa 2014.

Kluczowym, praktycznym zagadnieniem jest wybór metody szacowania *VaR*. Istnieje kilka metod wyznaczania tej miary. Do metod tych zaliczamy m.in. symulacje historyczne, symulacje z wykorzystaniem metody Monte Carlo oraz metodę szacowania miary *VaR* polegającą na wyznaczeniu kwantyla rozkładu.

W praktyce najczęściej nie jest znany rozkład danej cechy i konieczne jest wtedy wyznaczenie kwantyla z próby, czyli estymatora kwantyla rozkładu rzeczywistego.

## 2. Reprezentacja Bahadura kwantyla z próby

Jak wynika z powyższych rozważań kwantyl z próby jest bardzo ważnym estymatorem ze względu na zastosowania w innych dziedzinach nauki. Poniżej

zostanie przedstawiona reprezentacja Bahadura kwantyla z próby, która jest bardzo użytecznym narzędziem w dalszych badaniach, np. przy badaniu pewnych granicznych właściwości kwantyla z próby.

## 2.1. Niezależne zmienne losowe

Niech  $(\Omega, F, P)$  będzie przestrzenią probabilistyczną,  $\{X_n, n \geq 1\}$  ciągiem zmiennych losowych zdefiniowanych na tej przestrzeni. Załóżmy, że zmienne losowe  $X_n, n \geq 1$  mają ten sam rozkład zdefiniowany za pomocą dystrybuanty  $F$ . Ciąg ten możemy traktować jako nieskończoną próbę losową pochodzącą z populacji generalnej.

Niech  $Q(p) \in \mathbb{R}$  będzie ustalonym punktem, który spełnia warunek  $F(Q(p)) = p$ . O dystrybuancie  $F$  zakładamy ponadto, że:

- $F'(x) = f(x) > 0$ ,
- posiada drugą pochodną w pewnym otoczeniu punktu  $Q(p)$ ,
- $F''(x)$  jest w tym otoczeniu ograniczona.

Powyższe założenia zapewniają nam, że dla  $0 < p < 1$ ,  $Q(p)$  jest jedynym kwantylem rzędu  $p$  rozkładu  $F$ .

Dla próby  $X_1, X_2, \dots, X_n, n \geq 1$  empiryczna funkcja dystrybuanty zdefiniowana jest jako

$$F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x), \quad x \in \mathbb{R}, \quad (7)$$

gdzie

$$I(a) = \begin{cases} 1, & a \in A \\ 0, & a \notin A \end{cases} \quad (8)$$

Ponadto, dla  $n = 1, 2, \dots$  niech  $Q_n(p)$  oznacza kwantyl rzędu  $p$  wyznaczony z próby  $(X_1, X_2, \dots, X_n)$  o liczebności  $n$ .

Bahadur [2] jako pierwszy przedstawił reprezentację dla kwantyli z próby w zależności od empirycznej funkcji dystrybuanty dla niezależnych zmiennych losowych o tym samym rozkładzie. Reprezentacja Bahadura pozwala przybliżać asymptotycznie kwantyl z próby za pomocą średnich z próby pewnych ograniczonych zmiennych losowych (dystrybuant empirycznych) i w ten sposób badać pewne graniczne własności kwantyla z próby.

**Twierdzenie 2.1.** [2] Niech  $0 < p < 1$  oraz  $\{X_n, n \geq 1\}$  będzie ciągiem niezależnych zmiennych losowych o funkcji dystrybuanty  $F$ . Załóżmy, że  $F$  ma pochodne co najmniej drugiego rzędu w pewnych sąsiedztwie  $Q(p)$  oraz  $F'(Q(p)) = f(Q(p)) > 0$ . Wtedy

$$Q_n(p) = Q(p) - \frac{F_n(Q(p)) - p}{f(Q(p))} + O\left(n^{-\frac{3}{4}} \log(n)\right). \quad (9)$$

Bahadur [2] w swojej pracy otrzymał rząd zbieżności  $R_n = O\left(n^{\frac{-3}{4}} \log(n)\right)$  i pozostawił otwarty problem wyznaczenia dokładnego rzędu zbieżności. Kiefer [20] wyznaczył dokładny rząd zbieżności wynoszący  $O\left(n^{\frac{-3}{4}} \log(\log(n))^{\frac{3}{4}}\right)$ . Z kolei Ghosh [17] przedstawił znacznie prostszy dowód i wymagający mniej założeń w porównaniu do dowodu przedstawionego przez Bahadura [2].

W kolejnych latach wielu badaczy weryfikowało, czy reprezentację zaproponowaną przez Bahadura [2] można przenieść na przypadek, gdy między zmiennymi losowymi występuje pewna, określona struktura zależności.

## 2.2. Zależne zmienne losowe

Jedną z rozważanych i ważnych struktur zależności są zmienne ujemnie stowarzyszone.

**Definicja 2.1.** Skończona rodzina zmiennych losowych  $\{X_i, 1 \leq i \leq n\}$  jest ujemnie stowarzyszona (NA), jeśli dla każdej pary rozłącznych podzbiorów  $A$  oraz  $B$  zbioru  $\{1, 2, \dots, n\}$

$$\text{Cov}\left(f_1(X_i, i \in A), f_2(X_j, j \in B)\right) \leq 0, \quad (10)$$

gdzie  $f_1$  oraz  $f_2$  są dowolnymi rzeczywistymi niemalejącymi współrzędnościowo funkcjami.

Nieskończona rodzina zmiennych losowych jest ujemnie stowarzyszona, jeśli każda jej skończona podrodzina jest ujemnie stowarzyszona.

Zależność ujemnego stowarzyszenia została dobrze zbadana. Joag-Dev i Proschan [19] opisali najważniejsze własności tego rodzaju zależności. Newman [23] przedstawił centralne twierdzenie graniczne dla zmiennych losowych NA, Shao [26] dowiódł nierówności Rosenthala i twierdzenie o kompletnej zbieżności w prawie wielkich liczb.

Ling [22] przedstawił reprezentację analogiczną do tej zaproponowanej przez Bahadura dla zmiennych ujemnie stowarzyszonych przy pewnych ogólnych założeniach. W pracy otrzymano rząd zbieżności wynoszący  $O(\tau_n)$ , gdzie  $\{\tau_n, n \geq 1\}$  jest ciągiem liczb rzeczywistych z założeniem, że  $\tau_n \rightarrow 0$  oraz  $\frac{\sqrt{n}\tau_n^2}{\log n} \rightarrow \infty$ .

Xing i Yang [34] przedstawili reprezentację Bahadura dla zmiennych NA i otrzymali szybszą zbieżność modelu niż zbieżność zaprezentowana przez Linga [22]. Otrzymali rząd zbieżności wynoszący  $O\left(t_n n^{\frac{-1}{4}}\right)$ , gdzie  $\{t_n, n \geq 1\}$  jest ciągiem dodatnich liczb rzeczywistych spełniający warunki:  $t_n \rightarrow 0$  oraz  $\frac{t_n^2 n^{\frac{1}{2}}}{72} \geq (1+r) \log n$  dla pewnego  $r > 0$ .

Xu wraz ze współautorami [36] również przedstawili reprezentację Bahadura dla zmiennych  $NA$  i poprawili wyniki zaprezentowane przez Xinga i Yanga [34]. Otrzymali rząd zbieżności  $O(r_n)$ , gdzie  $\{r_n, n \geq 1\}$  jest ciągiem dodatnich liczb rzeczywistych spełniających warunki:  $r_n \rightarrow 0$  oraz  $nr_n^2 \geq 2\log n + 2(1 + \varepsilon)\log\log n$  dla pewnego  $\varepsilon > 0$ .

Kolejną strukturą zależności jest dodatnie stowarzyszenie. Pojęcie to zostało wprowadzone przez Esary'ego i innych [14].

**Definicja 2.2.** Skończona rodzina zmiennych losowych  $\{X_i, 1 \leq i \leq n\}$  jest dodatnio stowarzyszona (PA), jeśli dla każdej pary rozłącznych podzbiorów  $A$  oraz  $B$  zbioru  $\{1, 2, \dots, n\}$

$$\text{Cov}\left(f_1(X_i, i \in A), f_2(X_j, j \in B)\right) \geq 0, \quad (11)$$

gdzie  $f_1$  oraz  $f_2$  są dowolnymi rzeczywistymi niemalejącymi współrzędnościowo funkcjami.

Nieskończona rodzina zmiennych losowych jest ujemnie stowarzyszona, jeśli każda jej skończona podrodzina jest ujemnie stowarzyszona.

Więcej informacji na temat tego rodzaju zależności można znaleźć w pracy Bulinskiego i Shashkina [8].

Przy pewnych ogólnych założeniach, Douge [12] przedstawił reprezentację Bahadura dla zmiennych losowych dodatnio stowarzyszonych i otrzymał rząd zbieżności  $O\left(n^{\frac{1}{2} - \frac{\beta_b}{4}} (\log n)^\gamma\right)$ , gdzie  $b > \frac{5 + \sqrt{17}}{2}$ ,  $\beta_b = \frac{b-3}{b-1}$  oraz  $\gamma > \frac{3\beta_b}{4b} + \frac{1}{b}$ .

Kolejny rodzaj zależności zmiennych losowych, to zmienne losowe  $NUOD$ ,  $NLOD$  oraz  $NOD$ .

**Definicja 2.3.** Skończony zbiór zmiennych losowych  $X_1, X_2, \dots, X_n$  nazywamy *negatively upper orthant dependent (NUOD)*, jeśli dla każdej liczby rzeczywistej  $x_1, x_2, \dots, x_n$ ,

$$P(X_i > x_i, i = 1, 2, \dots, n) \leq \prod_{i=1}^n P(X_i > x_i), \quad (12)$$

oraz *negatively lower orthant dependent (NLOD)*, jeśli dla każdej liczby rzeczywistej  $x_1, x_2, \dots, x_n$

$$P(X_i \leq x_i, i = 1, 2, \dots, n) \leq \prod_{i=1}^n P(X_i \leq x_i). \quad (13)$$

Skończony zbiór zmiennych losowych  $X_1, X_2, \dots, X_n$  jest  $NOD$ , jeśli jest jednocześnie  $NUOD$  oraz  $NLOD$ .

Nieskończony ciąg  $\{X_n, n \geq 1\}$  jest  $NOD$  ( $NUOD$  lub  $NLOD$ ), jeśli każdy skończony podzbiór jest  $NOD$  ( $NUOD$  lub  $NLOD$ ).

Koncepcja ciągów  $NOD$  została wprowadzona przez Joag-Deva oraz Proschana [19]. W swojej pracy zauważyli, że zmienne losowe  $NA$  są również zmiennymi losowymi  $NOD$ , ale ani zmienne  $NUOD$ , ani  $NLOD$  nie są zmiennymi  $NA$ .

Li wraz ze współautorami [21] w swojej pracy rozważali reprezentację Bahadura dla ciągów zmiennych losowych  $NOD$ . Przy pewnych ogólnych założeniach otrzymali rząd zbieżności wynoszący  $O\left(\frac{(\log)^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right)$ .

W kolejnej części zostanie omówiona inna grupa zależności zmiennych losowych, tzw. zależność typu „mieszania”. Ten typ zależności opisuje asymptotyczną niezależność zmiennych losowych, gdy różnica ich indeksów dąży do nieskończoności. Ten rodzaj zależności opisują odpowiednio zdefiniowane współczynniki, których wartość zależy od  $\sigma$ -ciał generowanych przez te zmienne losowe.

Pierwsza rozważana struktura zależności typu „mieszania”, to zmienne losowe  $\varphi$ -mieszające. Pojęcie  $\varphi$ -mieszających zmiennych losowych zostało zaproponowane przez Dobrushina [11].

**Definicja 2.4.** Niech dane będzie  $\sigma$ -ciał

$$F_n^m = \sigma(X_i, n \leq i \leq m, i \in \mathbb{N}) \quad (14)$$

generowane przez zmienne losowe  $X_n, \dots, X_m, 1 \leq n \leq m \leq \infty$ .

Dla  $n \geq 1$  definiujemy

$$\varphi(n) = \sup_{m \geq 1, A \in F_1^m} \sup_{B \in F_{m+n}^\infty, P(A) > 0} |P(B|A) - P(B)|. \quad (15)$$

Jeśli  $\varphi \rightarrow 0$ , gdy  $n \rightarrow \infty$ , wtedy  $\{X_n, n \geq 1\}$  nazywamy ciągiem  $\varphi$ -mieszającym.

Od momentu wprowadzenia pojęcia zmiennych  $\varphi$ -mieszających, wielu badaczy analizowało właściwości tych ciągów. Utev [30] badał centralne twierdzenie graniczne, Chen wraz ze współautorami [9] badali całkowitą zbieżność ciągów, Yang i współautorzy [39] badali ograniczenie Berry'ego–Essena. Więcej na temat własności ciągów  $\varphi$ -mieszających można znaleźć w rozdziale czwartym książki Billingsley'a [3]. W kolejnych latach powstało wiele prac rozważających reprezentację Bahadura dla ciągów z powyższą strukturą zależności.

Sen [25] w pracy przedstawił reprezentację dla kwantyla z próby analogiczną do reprezentacji Bahadura, ale dla zmiennych losowych  $\varphi$ -mieszających. Przy pewnych ogólnych założeniach oraz założeniu, że współczynniki mieszające spełniają warunek  $\sum_{n=1}^{\infty} (\varphi(n))^{\frac{1}{2}} < \infty$ , Sen [25] otrzymał rząd zbieżności  $O\left(n^{-\frac{5}{8}} \log n\right)$ . Następnie zmieniając założenie na współczynniki mieszające i zakładając, że spełniają one warunek  $\sum_{n=1}^{\infty} n^k (\varphi(n))^{\frac{1}{2}} < \infty$ , dla pewnego  $k \leq 1$ , otrzymał w tej samej pracy rząd zbieżności wynoszący  $O\left(n^{-\frac{3}{4}} \log n\right)$ .

Babu i Singh [1] rozważali ciągi  $\varphi$ -mieszających zmiennych losowych o rozkładzie jednostajnym. Przy założeniu  $\sum_{n=1}^{\infty} (\varphi(n))^{\frac{1}{2}} < \infty$ , pokazali, że rząd zbieżności wynosi  $O\left(n^{\frac{-3}{4}}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{4}}\right)$ .

Yoshihara [40] rozważał reprezentację Bahadura dla przypadku zmiennych losowych  $\varphi$ -mieszających i wyniki jego pracy są rozszerzeniem wyników przedstawionych przez Sena [25]. W pracy założono, że współczynniki  $\varphi$ -mieszające spełniają warunek  $\varphi(n) = O(n^{-2})$ . Przy takim warunku otrzymano rząd zbieżności  $O\left(n^{\frac{-3}{4}} \log n\right)$ .

Również w pracy Yanga wraz ze współautorami [38] rozważano reprezentację Bahadura dla zmiennych losowych  $\varphi$ -mieszających. Przy warunku na współczynniki mieszające  $\varphi(n) = O(n^{-3})$  otrzymano rząd zbieżności  $O\left(n^{\frac{-3}{4}} \log n\right)$ . Następnie, gdy osłabiono założenie na współczynniki mieszające, tzn. założono, że spełniają warunek  $\sum_{n=1}^{\infty} (\varphi(n))^{\frac{1}{2}} < \infty$ , otrzymano rząd zbieżności  $O\left(n^{\frac{-1}{2}}(\log n)^{\frac{1}{2}}\right)$ .

Wu wraz ze współautorami [33] również przedstawili reprezentację Bahadura dla zmiennych losowych  $\varphi$ -mieszających. Dużą zaletą pracy był brak założeń na tempo zaniku współczynników mieszających. Otrzymano rząd zbieżności  $O\left(n^{\frac{-3}{4}+\delta}\right)$  dla pewnego  $0 < \delta < \frac{1}{4}$ . Dodatkowo w pracy przedstawiono zastosowanie reprezentacji Bahadura. W pracy zbadano asymptotyczną normalność kwantyla z próby i ograniczenie Berry'ego–Esseena.

Kolejnym rodzajem zależności zmiennych losowych, które są często rozważane w licznych pracach, są  $\alpha$ -mieszające zmienne losowe, tzw. silnie mieszające zmienne losowe.

**Definicja 2.5.** Niech dane będzie  $\sigma$ -ciało

$$F_n^m = \sigma(X_i, n \leq i \leq m, i \in \mathbb{N}) \quad (16)$$

generowane przez zmienne losowe  $X_n, \dots, X_m, 1 \leq n \leq m \leq \infty$ .

Dla  $n \geq 1$  definiujemy

$$\alpha(n) = \sup_{n \in \mathbb{N}, A \in F_1^n} \sup_{B \in F_{m+n}^{\infty}} |P(A \cap B) - P(A)P(B)|. \quad (17)$$

Jeśli  $\alpha \rightarrow 0$ , gdy  $n \rightarrow \infty$ , wtedy  $\{X_n, n \geq 1\}$  nazywamy ciągiem  $\alpha$ -mieszającym.

Zauważmy, że  $\alpha(n) \leq \varphi(n)$ . Aby uzyskać więcej informacji na temat własności i innych wyników badań nad ciągami  $\alpha$ -mieszającymi warto przeczytać prace Bosq [4], Fana i Yao [15], Hosseinioun wraz ze współautorami [18] oraz Fana wraz ze współautorami [16].

$\alpha$ -mieszające zmienne losowe pojawiają się w wielu pracach związanych z reprezentacją Bahadura. Wspomniani już wcześniej Babu i Singh [1], w tej samej pracy poza analizą zmiennych  $\varphi$ -mieszających badali również zmienne



$\alpha$ -mieszające. Przy założeniu, że współczynniki mieszające spełniają warunek  $\alpha(n) = O(e^{-\theta n})$  dla pewnego  $\theta > 0$  otrzymali reprezentację Bahadura i rząd zbieżności wynoszący  $O\left(n^{\frac{-3}{4}}(\log n)(\log \log n)^{\frac{1}{4}}\right)$ .

We wspomnianej wcześniej pracy Yoshihary [40], również poza wyznaczeniem reprezentacji Bahadura dla zmiennych  $\varphi$ -mieszających, wyznaczono analogiczną reprezentację dla zmiennych silnie mieszających. Nałożono dodatkowy warunek na współczynniki  $\alpha$ -mieszające, tzn. współczynniki muszą spełniać warunek  $\alpha(n) = O\left(n^{\frac{-5}{2}-\epsilon}\right)$  dla  $\epsilon > 0$ . Przy takim założeniu pokazano, że rząd zbieżności wynosi  $O\left(n^{\frac{-3}{4}} \log n\right)$ .

Sun [27] wyprowadził reprezentację Bahadura dla kwantyla z próby silnie mieszających zmiennych losowych przy pewnych ogólnych założeniach i dodatkowym założeniu na współczynniki mieszające. Przy założeniu, że współczynniki mieszające spełniają warunek  $\alpha(n) = O(n^{-\beta})$  dla  $\beta > 10$  otrzymał rząd zbieżności wynoszący  $O\left(n^{\frac{-3}{4}+\delta} \log n\right)$  dla  $\delta \in \left(\frac{11}{4(\beta+1)}, \frac{1}{4}\right)$ . Wyniki te rozszerzyły twierdzenie podane przez Yoshiharę [40].

Dalsze badania nad reprezentacją Bahadura przeprowadził Wang ze współautorami [32]. W pracy skupiono się na zmiennych  $\alpha$ -mieszających i otrzymano wyniki lepsze od tych przedstawionych przez Suna [27]. Przy analogicznym założeniu, że współczynniki mieszające spełniają warunek  $\alpha(n) = O(n^{-\beta})$ , gdzie  $\beta > 11$ , przedstawiono reprezentację Bahadura dla zmiennych silnie mieszających i otrzymano rząd zbieżności  $O\left(n^{\frac{-3}{4}+\delta}(\log n)^{\frac{3}{4}}\right)$ .

Kolejną pracą, w której badano reprezentację Bahadura dla zmiennych silnie mieszających, jest praca Zhanga wraz ze współautorami [41]. Autorzy założyli, że  $\alpha(n) = O(n^{-\beta})$  oraz osłabili założenie na  $\beta$  zakładając, że  $\beta > 3$ . Przy takim warunku otrzymano reprezentację Bahadura i rząd zbieżności rzędu  $O\left(n^{\frac{-1}{2}}(\log \log n * \log n)^{\frac{1}{2}}\right)$ . Następnie przy założeniu, że  $\alpha(n) = O(n^{-\beta})$ , dla  $\beta > \max\left\{3 + \frac{5}{1+\delta}, 1 + \frac{2}{\delta}\right\}$  oraz  $\delta > 0$  otrzymano rząd zbieżności  $O\left(n^{\frac{-3}{4} + \frac{\delta}{4(2+\delta)}}(\log \log n * \log n)^{\frac{1}{2}}\right)$ .

Kolejnym rodzajem zależności typu „mieszania” są  $\psi$ -mieszające zmienne losowe.

**Definicja 2.6.** Niech dane będzie  $\sigma$ -ciało

$$F_n^m = \sigma(X_i, n \leq i \leq m, i \in \mathbb{N}) \quad (18)$$

generowane przez zmienne losowe  $X_n, \dots, X_m, 1 \leq n \leq m \leq \infty$ .

Dla  $n \geq 1$  definiujemy

$$\psi(n) = \sup_{m \geq 1, A \in F_1^m} \sup_{B \in F_{m+n}^\infty, P(A)P(B) > 0} \frac{|P(AB) - P(A)P(B)|}{P(A)P(B)}. \quad (19)$$

Jeśli  $\psi \rightarrow 0$ , gdy  $n \rightarrow \infty$ , wtedy  $\{X_n, n \geq 1\}$  nazywamy ciągiem  $\psi$ -mieszającym.

Koncept  $\psi$ -mieszających zmiennych losowych został wprowadzony przez Bluma wraz ze współautorami [5] i dodatkowo w tej pracy przedstawione zostało silne prawo wielkich liczb, Yang [37] analizował prawie pewną zbieżność sum ważonych oraz Wang ze współautorami [31] podali nierówności związane z momentami dla przypadku  $\psi$ -mieszających zmiennych losowych.

Reprezentacja Bahadura dla ciągu zmiennych losowych  $\psi$ -mieszających była analizowana w pracy Xinga i Yanga [35]. W pracy przedstawiono reprezentację Bahadura przy założeniu, że współczynniki mieszające spełniają warunek  $\sum_{j=1}^\infty \psi^{\frac{1}{r}}(2^j) < \infty$ , gdzie  $r > \frac{1}{\delta}$  oraz  $0 < \delta < \frac{1}{4}$ . Przy takim ograniczeniu otrzymano rząd zbieżności wynoszący  $O\left(n^{-\frac{3}{4}}(\log n)^{\frac{1}{2}}\right)$ .

Ostatnim omawianym w tej pracy rodzajem zależności są  $\rho^*$ -mieszające zmienne losowe.

**Definicja 2.7.** Ciąg zmiennych losowych  $\{X_n, n \geq 1\}$  nazywamy  $\rho^*$ -mieszającym, jeśli

$$\rho^*(n) = \sup\{\rho(S, T) : S, T \subset \mathbb{N}, \text{dist}(S, T) \geq n\} \rightarrow 0, \quad (20)$$

gdy  $n \rightarrow \infty$  oraz gdzie

$$\rho(S, T) = \sup\left\{\frac{|Cov(X, Y)|}{\sqrt{Var(X)Var(Y)}} : X \in L_2(\sigma(S)), Y \in L_2(\sigma(T))\right\}, \quad (21)$$

$$\text{dist}(S, T) = \min_{i \in S, j \in T} |j - i| \quad (22)$$

oraz  $\sigma(S)$  i  $\sigma(T)$  są  $\sigma$ -ciałami generowanymi odpowiednio przez  $\{X_i, i \in S\}$  oraz  $\{X_j, j \in T\}$ .

Różne własności takich ciągów były badane w wielu pracach, m.in. Bradley [6] analizował centralne twierdzenie graniczne, Bryc i Smolenski [7], Peligrad i Gut [24] oraz Utev i Peligrad [29] otrzymali nierówności związane z momentami, Sung [28] badał kompletną zbieżność sum ważonych dla  $\rho^*$ -mieszających ciągów zmiennych losowych.

W pracy Dudek i Kuczmaszewskiej [13] przedstawiono reprezentację Bahadura dla zmiennych losowych  $\rho^*$ -mieszających. Przy pewnych ogólnych założeniach, ale bez nakładania ograniczeń na współczynniki mieszające otrzymano rząd zbieżności wynoszący  $O\left(n^{-\frac{3}{4} + \delta}\right)$ , gdzie  $0 < \delta < \frac{1}{4}$ . Dodatkowo, w pracy wyznaczono ograniczenie Berry'ego Esseena dla ciągu zmiennych losowych  $\rho^*$ -mieszających.

Z powyższych rozważań widać, że reprezentacja Bahadura kwantyla z próby była rozważana w wielu pracach. Reprezentacja ta jest bardzo istotna, ponieważ pozwala badać graniczne właściwości ciągów zmiennych losowych zarówno niezależnych, jak i wykazujących określone struktury zależności.

### 3. Symulacje

W różnych pracach, np. Dudek i Kuczmaszewskiej [13] i Wu ze współautorami [33], poza wyznaczeniem reprezentacji Bahadura, były badane graniczne właściwości ciągów zmiennych losowych odpowiednio  $\rho^*$  i  $\varphi$ -mieszających. Zgodnie z wnioskami z tych prac statystyka

$\sqrt{n}(Q_n(p) - Q(p))$ , dąży według rozkładu do rozkładu normalnego. W tym rozdziale zostanie przeprowadzona symulacja, aby sprawdzić słuszność tego wniosku.

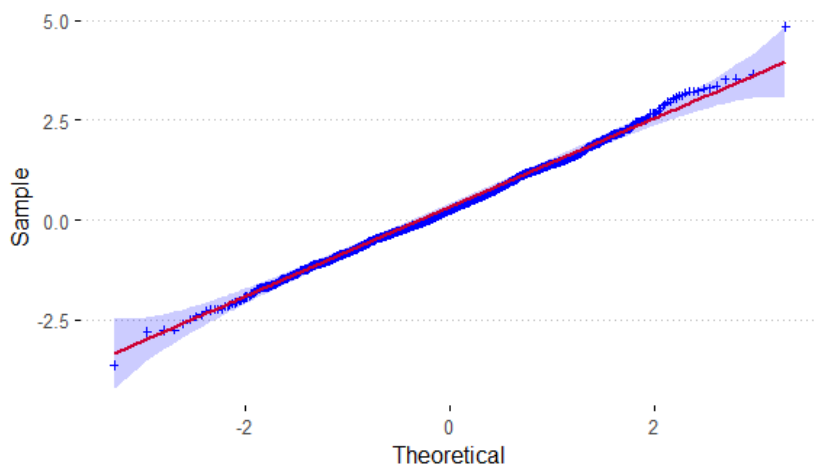
Niech  $\{\varepsilon_n\}$  będzie ciągiem zmiennych losowych o jednakowym rozkładzie, zerowej wartości oczekiwanej oraz skończonej wariancji. Niech

$$X_n = \sum_{k=0}^m a_k \varepsilon_{n+k}, \quad (23)$$

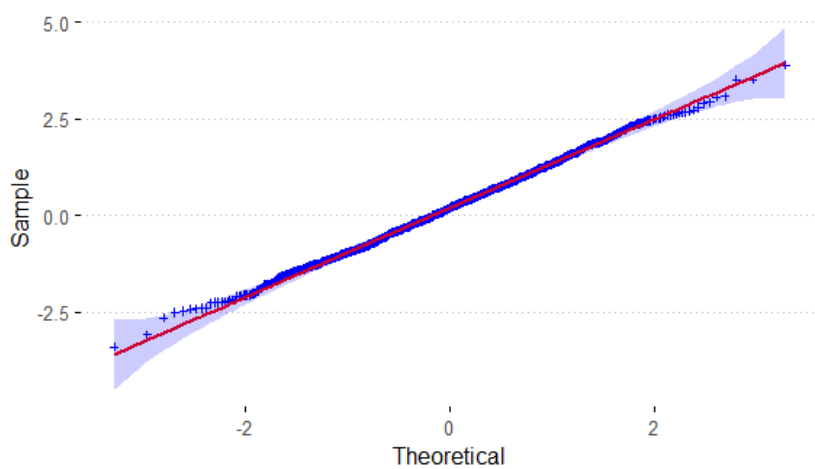
gdzie  $m$  to ustalona dodatnia stała oraz  $a_k, k = 0, 1, \dots, m$  to ustalone stałe. Wtedy  $\{X_n\}$  nazywamy procesem średniej ruchomej rzędu  $m$ . Możemy łatwo sprawdzić, że proces średniej ruchomej jest procesem  $\rho^*$ -mieszającym oraz  $\varphi$ -mieszającym.

Niech oraz  $X_n = \sum_{k=0}^m \varepsilon_{n+k}$ . Można łatwo pokazać, że ciąg  $\{X_n\}$  jest ciągiem zmiennych losowych wykazujący własność  $\rho^*$ -mieszania oraz  $\varphi$ -mieszania.

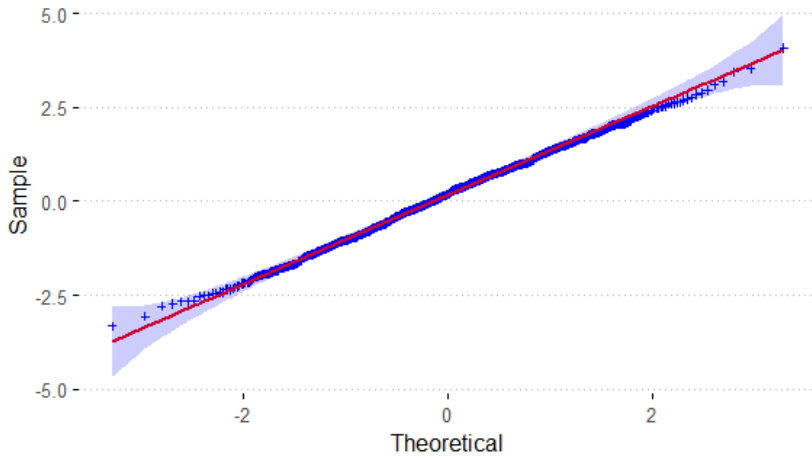
Wykorzystując program R studio wyznaczamy 1000 razy statystykę  $S_n = \sqrt{n}(Q_n(p) - Q(p))$ , gdzie  $n$  jest rozmiarem pobranej próby. W symulacji obieramy  $m = 10$  i  $p = 0,05$ . Zgodnie z literaturą [13, 33] statystyka ta powinna dążyć według rozkładu do rozkładu normalnego ze średnią równą zero i wariancją równą  $\left(\frac{\sigma_p}{f(Q(p))}\right)^2$ . Aby to zilustrować w programie R studio wygenerowano wykresy kwantyl–kwantyl. Ten typ wykresu służy do sprawdzania, czy dana zmienna losowa ma dany rozkład (np. rozkład normalny). Na jednej osi wykresu umieszczone są kwantyle rozkładu badanej zmiennej, a na drugiej osi kwantyle porównywanego rozkładu (przy estymowanych jego parametrach). Jeśli zmienna ma idealnie zadany rozkład, wykres ten przedstawia dokładnie prostą. Rysunki 2–5 przedstawiają wykresy kwantyl–kwantyl dla prób rozmiaru  $n = \{200, 500, 1000, 2000\}$  odpowiednio.

**Rysunek 2. Próba  $n = 200$** 

Źródło: opracowanie własne.

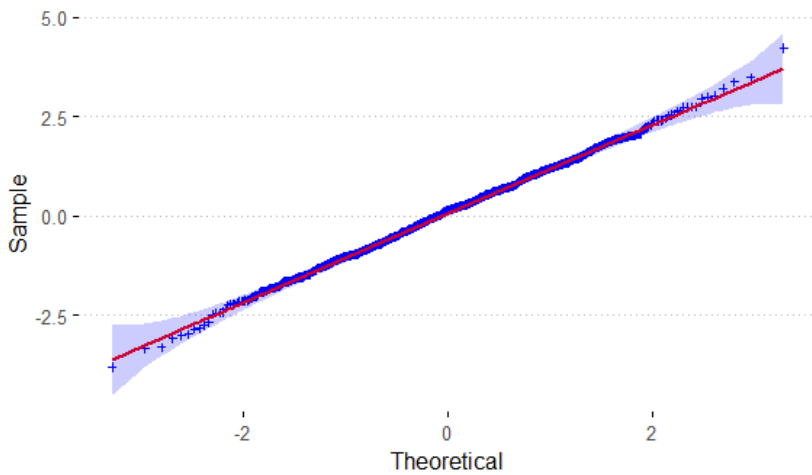
**Rysunek 3. Próba  $n = 500$** 

Źródło: opracowanie własne.



**Rysunek 4. Próba  $n = 1000$**

Źródło: opracowanie własne.



**Rysunek 5. Próba  $n = 2000$**

Źródło: opracowanie własne.

Z rysunków 2–5 widać, że wraz ze wzrostem próby losowej rozkład statystyki  $S_n$  zbiega do rozkładu normalnego, co jest w zgodzie z wnioskami zamieszczonymi w wyżej wspomnianej literaturze.

## Literatura

- [1] Babu G. J., Singh K., *On deviations between empirical and quantile processes for mixing random variables*, “Journal of Multivariate Analysis” 1978, vol. 8, no. 4, pp. 532–549.
- [2] Bahadur R. R., *A note on quantiles in large samples*, “The Annals of Mathematical Statistics” 1966, vol. 37, no. 3, pp. 577–580.
- [3] Billingsley P., *Convergence of probability measures*, Wiley, New York 1968.
- [4] Bosq D., *Nonparametric statistics for stochastic processes*, Springer, New York, 1998.
- [5] Blum J. R., Hanson D. L., Koopmans L., *On the strong law of large numbers for a class of stochastic process*, “Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete” 1963, vol. 2, pp. 1–11.
- [6] Bradley R. C., *On the spectral density and asymptotic normality of weakly dependent random fields*, “Journal of Theoretical Probability” 1992, vol. 5, pp. 355–373.
- [7] Bryc W., Smolenski W., *Moment conditions for almost sure convergence of weakly correlated random variables*, “Proceedings of the American Mathematical Society” 1993, vol. 119, no. 2, pp. 629–635.
- [8] Bulinski A., Shashkin A., *Limit Theorems For Associated Random Fields and Related Systems*, “Advances Series on Statistical Science and Applied Probability” 2007, vol. 10, pp. 169–228.
- [9] Chen P. Y., Hu T. C., Volodin A., *Limiting behavior of moving average process under  $\phi$ -mixing assumption*, “Statistics & Probability Letters” 2009, vol. 79 (1), pp. 105–111.
- [10] Chlebus M., *Pomiar ryzyka rynkowego za pomocą miary Value at Risk – podejście dwuetapowe*, Uniwersytet Warszawski, rozprawa doktorska, 2014.
- [11] Dobrushin R. L., *The central limit theorem for non-stationary Markov chain*, “Theory of Probability & Its Applications” 1990, vol. 35, pp. 72–88.
- [12] Douge L., *The Bahadur representation of sample quantiles for associated sequences*, “Statistics Theory” 2022, pp. 1–11.
- [13] Dudek D., Kuczmaszewska A., *On the Bahadur Representation of Quantiles for a Sample from  $\rho^*$ -Mixing Structure Population*, “Advances in Science and Technology Research Journal” 2022, vol. 16, no. 3, pp. 316–330.
- [14] Esary J. D., Proschan F., Walkup D. W., *Association of Random Variables, with Applications*, “The Annals of Mathematical Statistics” 1967, vol. 38, no. 5, pp. 1466–1474.
- [15] Fan J. Q., Yao Q. W., *Nonlinear time series: nonparametric and parametric methods*, Springer, New York, 2005.

- 
- [16] Fan G. L., Liang H. Y., Wang J. F., *Empirical likelihood for heteroscedastic partially linear errors in variables model with  $\alpha$ -mixing errors*, “Statistical Papers” 2011, vol. 54, pp. 85–112.
- [17] Ghosh J. K., *A New Proof of the Bahadur Representation of Quantiles and Application*, “The Annal of Mathematical Statistics” 1971, vol. 42, no. 6, pp. 1957–1961.
- [18] Hosseinioun N., Doosti H., Nirumand H. A., *Nonparametric estimation of the derivatives of a density by the method of wavelet for mixing sequences*, “Statistical Papers” 2012, vol. 53, no. 1, pp. 195–203.
- [19] Joag-Dev K., Proschan F., *Negative association of random variables with applications*, “The Annals of Statistics” 1999, vol. 11, no. 1, pp. 963–991.
- [20] Kiefer J., *On Bahadur's representation of sample quantiles*, “The Annal of Mathematical Statistics” 1967, vol. 38, no. 5, pp. 1323–1342.
- [21] Li X., Yang W., Hu S., *The Bahadur representation for sample quantile under NOD sequences*, “Journal of Nonparametric Statistics” 2011, vol. 23, no. 1, pp. 59–65.
- [22] Ling N. X., *The Bahadur representation for sample quantiles under negatively associated sequences*, “Statistics and Probability Letters” 2008, vol. 78, no. 16, pp. 2660–2663.
- [23] Newman C. M., *Asymptotic independence and limit theorems for positively and negatively dependent random variables*, “Inequalities in Statistics and Probability” 1984, vol. 5, pp. 127–140.
- [24] Peligrad M., Gut A., *Almostsure results for a class of dependent random variables*, “Journal of Theoretical Probability” 1999, vol. 12, pp. 87–104.
- [25] Sen P. K., *On the Bahadur representation of sample quantiles for sequences of  $\varphi$ -mixing random variables*, “Journal of Multivariate Analysis” 1972, vol. 2, no. 1, pp. 77–95.
- [26] Shao Q. M., *A comparison theorem on moment inequalities between negatively associated and independent random variables*, “Journal of Theoretical Probability” 2000, vol. 13, pp. 343–356.
- [27] Sun S. X., *The Bahadur representation for sample quantiles under weak dependence*, “Statistics and Probability Letters” 2006, vol. 76, no. 12, pp. 1238–1244.
- [28] Sung S. H., *Complete convergence for weighted sums of  $\rho^*$ -mixing random variables*, “Discrete Dynamics in Nature and Society” 2010, vol. 2010, pp. 1–13.
- [29] Utev S., Peligrad M., *Maximal inequalities and an invariance principle for a class of weakly dependent random variables*, “Journal of Theoretical Probability” 2003, vol. 16, pp. 101–115.
- [30] Utev S. A., *On the central limit theorem for  $\varphi$ -mixing arrays of random variables*, “Theory of Probability & Its Applications” 1990, vol. 35, no. 1, pp. 131–139.

- 
- [31] Wang X. J., Hu S. H., Shen Y., Yang W. Z., *Maximal inequality for  $\psi$ -mixing sequences and its applications*, “Applied Mathematics Letters”, 2010, vol. 23, pp. 1156–1161.
- [32] Wang X. J., Hu S. H., Yang W. Z., *The Bahadur representation for sample quantiles under strongly mixing sequence*, “Journal of Statistical Planning and Inference” 2011, vol. 141, no. 2, pp. 655–662.
- [33] Wu Y., Yu W., Wang X., *The Bahadur representation of sample quantiles for  $\phi$ -mixing random variables and its application*, “Statistics” 2021, vol. 55, no. 2, pp. 426–444.
- [34] Xing G. D., Yang S. C., *A remar on the Bahadur representation of sample quantiles for negatively associated sequences*, “Journal of the Korean Statistical Society” 2011, vol. 40, no. 3, pp. 277–280.
- [35] Xing G., Yang S., *On the Bahadur representation of sample quantiles for  $\psi$ -mixing sequences and its applications*, “Communications in Statistics Theory and Methods” 2018, vol. 48, no. 5, pp. 1060–1072.
- [36] Xu S., Ge L., Miao Y., *On the Bahadur representation of sample quantiles and order statistics for NA sequences*, “Journal of the Korean Statistical Society” 2013, vol. 42, no. 1, pp. 1–7.
- [37] Yang S. C., *Almost sure convergence of weighted sums of mixing sequences*, “Journal of System Science and Mathematical Science Chinese Series” 1995, vol. 15, no. 3, pp. 254–265.
- [38] Yang W., Hu S., Wang X., *The Bahadur representation for sample quantiles under dependent sequence*, “Acta Mathematicae Applicatae Sinica” 2019, vol. 35, pp. 521–531,
- [39] Yang W. Z., X. H.Wang, X. Q.Li, Hu S., *Berry-Esseen bound of sample quantiles for  $\phi$ -mixing random variables*, “Journal of Mathematical Analysis and Applications” 2012, vol. 388, pp. 451–462.
- [40] Yoshihara K., *The Bahadur representation of sample quantile for sequences of strongly mixing random variables*, “Statistic and Probability Letters” 1995, vol. 24, no. 4, pp. 299–305.
- [41] Zhang Q., Yang W., Hu S., *On Bahadur representation for sample quantiles under  $\alpha$ -mixing sequence*, “Statistical Papers” 2014, vol. 55, pp. 285–299.



## Selected issues related to quantiles

### Abstract

Quantiles are one of the most important and often used parameters of distributions of the feature or random variable. The paper presents basic concepts and information related to quantiles. The model of the Value at Risk measure will be introduced, which is a direct example of the use of quantiles in finance and economics. Next, a special representation of the sample quantile will be presented, the so-called Bahadur representation, which is very important due to the study of limit properties of the sample quantile. In addition, a simulation will be presented to verify some boundary conclusions related to the sample quantile.

*Keywords: quantile, dependent random variables, Bahadur representation of sample quantile.*

**Marcin Dziadosz<sup>1</sup>**

## **Zastosowanie teorii kolejek do poprawy jakości obsługi klientów poczty<sup>2</sup>**

### **Streszczenie**

W pracy przedstawione zostały wybrane modele kolejkowe wraz z ich praktycznymi zastosowaniami. Praca zawiera najważniejsze elementy teorii obsługi masowej. Ponadto, zaprojektowany i rozwiązany został problem praktyczny, do którego dopasowano odpowiedni model kolejkowy i obliczono jego najważniejsze charakterystyki.

*Słowa kluczowe: modele kolejkowe, prawo Little'a, teoria kolejek.*

### **Wstęp**

Wieloletnia działalność w Studenckim Kole Naukowym „Kwaterion” Politechniki Lubelskiej była dla mnie, jak i dla pozostałych członków koła, okresem pełnym osobistego rozwoju, poszukiwań naukowych oraz zgłębiania wiedzy na temat różnych zastosowań „królowej nauk” – matematyki. Jednym z obszarów, który szczególnie zwrócił moją uwagę, jest tzw. teoria kolejek.

Teoria obsługi masowej zajmuje się analizowaniem systemów kolejkowych, na przykład w sklepie, czy też w banku. Modele kolejkowe są konstruowane w taki sposób, aby na ich podstawie dało się przewidzieć pewne wielkości, między innymi długość kolejki, czy też czas oczekiwania klienta na obsługę. Teoria masowej obsługi jest wykorzystywana przy podejmowaniu decyzji biznesowych, dzięki niej możliwa jest też optymalizacja systemu kolejkowego w danym przedsiębiorstwie [1, 4, 8, 9].

Celem pracy jest ukazanie zastosowania teorii kolejek do poprawy jakości obsługi klientów poczty. Nazwę każdego modelu da się zapisać za pomocą notacji Kendalla  $A/B/c/L/N$ , w której każdy symbol opisuje w pewien sposób model kolejkowy. Zaprezentowane zostały modele  $M/M/1/0$ ,  $M/M/1/L$ ,  $M/M/c/0$  oraz  $M/M/c/L$ . W przypadku każdego modelu przedstawione zostały najważniejsze wielkości charakteryzujące dany model kolejkowy, takie jak średnia liczba klien-

---

<sup>1</sup> Mgr inż. Marcin Dziadosz, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>2</sup> Niniejszy tekst stanowi fragment pracy dyplomowej pod tytułem *Wybrane modele kolejkowe w teorii obsługi masowej*, obronionej w dn. 7 lipca 2022 na Wydziale Podstaw Techniki Politechniki Lubelskiej.

tów w całym systemie kolejkowym, czy też średni czas spędzony przez klienta w poczekalni.

System kolejkowy składa się zazwyczaj z poczekalni, w której klienci czekają oraz z punktów obsługi. Każdy taki punkt możemy scharakteryzować parametrem obciążenia obsługi.

## 1. Obciążenie punktów obsługi

Obciążeniem punktów obsługi nazywamy następujący iloraz:

$$\rho = \frac{\lambda}{\mu}, \quad (1)$$

gdzie  $\lambda$  jest intensywnością napływu klientów (zgłoszeń), natomiast  $\mu$  jest intensywnością obsługi klientów (obsługą) w punktach obsługi [5].

## 2. Ogólna charakterystyka modelu kolejkowego

Każdy model kolejkowy można opisać poprzez tzw. notację Kendalla:  $A/B/c/L/N$ .

Każda litera powyższej notacji charakteryzuje w pewien sposób rozważany model kolejkowy [2, 17]:

- $A$  oznacza rozkład prawdopodobieństwa czasu pomiędzy kolejnymi pojawiającymi się klientami,
- $B$  oznacza rozkład prawdopodobieństwa czasu obsługi,
- $c$  oznacza liczbę stanowisk obsługi,
- $L$  oznacza liczbę miejsc w poczekalni (ile klientów może przebywać w kolejce),
- $N$  oznacza liczbę wszystkich klientów.

Warto zwrócić uwagę na to, że jeżeli w notacji brakuje  $L$  lub  $N$ , to oznacza to, że są wielkościami nieskończenie wielkimi. W praktyce oznacza to, że nie są one ograniczone.

Ponadto, aby oznaczyć rozkład wykładniczy w notacji, wykorzystuje się literę  $M$ , natomiast do oznaczenia dowolnego rozkładu wykorzystywana jest litera  $G$ .

## 3. Prawo Little'a

Rozważamy pewien system kolejkowy, w którym klienci pojawiają się, przebywają w systemie przez pewien czas, a potem go opuszczają. Przez  $\lambda$  oznaczamy intensywność napływu klientów, przez  $W$  średni czas przebywania klienta w systemie, natomiast przez  $L$  średnią liczbę klientów przebywających w systemie. Wartości te są ze sobą powiązane w następujący sposób:

$$L = \lambda W. \quad (2)$$

Jest to tzw. prawo Little’a, które jest jednym z najważniejszych zależności w teorii masowej obsługi [1, 3, 13].

## 4. Wybrane modele kolejkowe

### 4.1. Model kolejkowy M/M/1/0

Pierwszym rozważanym modelem kolejkowym w pracy będzie model M/M/1/0 [5, 16]. W modelu tym zakładamy, że:

- rozkład prawdopodobieństwa czasu pomiędzy kolejnymi pojawiającymi się klientami jest wykładniczy o parametrze  $\lambda$ ,
- rozkład prawdopodobieństwa czasu obsługi również jest wykładniczy o parametrze  $\mu$ ,
- jest tylko jeden punkt obsługi,
- nie ma miejsc w poczekalni,
- liczba wszystkich klientów jest nieograniczona.

Zauważmy, że w modelu kolejkowym M/M/1/0 możliwy jest tylko stan, w którym mamy 0 klientów (wtedy punkt obsługi jest wolny) oraz stan, gdy mamy 1 klienta (wtedy punkt obsługi jest zajęty, a inne zgłoszenia w tym czasie nie są uwzględniane).

Tabela 1. Interpretacja stanów w modelu kolejkowym M/M/1/0

Stan	Interpretacja
$S_0$	0 klientów, wolny punkt obsługi
$S_1$	1 klient, zajęty punkt obsługi, inne zgłoszenia nie są uwzględniane

Źródło: opracowanie własne.

Przyjmujemy  $\lambda_0 = \lambda$  oraz  $\mu_1 = \mu$ . Wtedy wyrażenie  $\frac{1}{\lambda}$  jest równe wartości oczekiwanej czasu pomiędzy kolejnymi pojawiającymi się klientami, natomiast wyrażenie  $\frac{1}{\mu}$  jest równe wartości oczekiwanej czasu obsługi.

W teorii obsługi masowej przyjmuje się, że:

$$Q_s = \frac{\lambda_0 \lambda_1 \lambda_2 \dots \lambda_{s-1}}{\mu_1 \mu_2 \mu_3 \dots \mu_s} = \prod_{i=1}^s \frac{\lambda_{i-1}}{\mu_i} \tag{3}$$

oraz

$$Q_0 = 1. \tag{4}$$

Dla tego modelu mamy więc:

$$Q_1 = \frac{\lambda_0}{\mu_1} = \frac{\lambda}{\mu} = \rho, \tag{5}$$

gdzie  $\rho$  oznacza obciążenie punktów obsługi. Aby obliczyć prawdopodobieństwa poszczególnych stanów, skorzystamy z następującej zależności:

$$p_i = \frac{Q_i}{\sum_{k=0}^n Q_k}, \text{ dla } 0 \leq i \leq n \quad (6)$$

Otrzymujemy zatem prawdopodobieństwo stanu kiedy mamy 0 klientów:

$$p_0 = \frac{Q_0}{Q_0 + Q_1} = \frac{1}{1 + \rho}, \quad (7)$$

a także prawdopodobieństwo stanu gdy mamy 1 klienta:

$$p_1 = \frac{Q_1}{Q_0 + Q_1} = \frac{\rho}{1 + \rho}. \quad (8)$$

W przypadku tego modelu możemy również obliczyć prawdopodobieństwo tego, że klient zostanie obsłużony (oznaczamy jako  $p_{yes}$ ):

$$p_{yes} = p_0 = \frac{Q_0}{Q_0 + Q_1} = \frac{1}{1 + \rho} \quad (9)$$

oraz prawdopodobieństwo tego, że klient nie zostanie obsłużony (oznaczamy jako  $p_{no}$ ):

$$p_{no} = p_1 = \frac{Q_1}{Q_0 + Q_1} = \frac{\rho}{1 + \rho}. \quad (10)$$

## 4.2. Model kolejkowy M/M/1/L

Rozważmy teraz model M/M/1/L [5, 14]. W modelu tym zakładamy, że:

- rozkład prawdopodobieństwa czasu pomiędzy kolejnymi pojawiającymi się klientami jest wykładniczy o parametrze  $\lambda$ ,
- rozkład prawdopodobieństwa czasu obsługi również jest wykładniczy o parametrze  $\mu$ ,
- jest tylko jeden punkt obsługi,
- jest  $L$  miejsc w poczekalni,
- liczba wszystkich klientów jest nieograniczona.

Zauważmy, że w modelu kolejkowym M/M/1/L możliwe są następujące stany przedstawione w tabeli 2.

Tabela 2. Interpretacja stanów w modelu kolejkowym M/M/1/L

Stan	Interpretacja
$S_0$	0 klientów, wolny punkt obsługi, wolne $L$ miejsc w poczekalni
$S_1$	1 klient, zajęty punkt obsługi, wolne $L$ miejsc w poczekalni
$S_2$	2 klientów, zajęty punkt obsługi, wolne $L - 1$ miejsc w poczekalni
...	...
$S_{L+1}$	$L + 1$ klientów, zajęty punkt obsługi, zajęte $L$ miejsc w poczekalni, inne zgłoszenia nie są uwzględniane

Źródło: opracowanie własne.

Jeśli przyjmiemy:

$$\lambda_0 = \lambda_1 = \dots = \lambda_L = \lambda, \mu_1 = \mu_2 = \dots = \mu_{L+1} = \mu, \quad (11)$$

to mamy wtedy:

$$Q_0 = 1, Q_i = \frac{\lambda^i}{\mu^i} = \rho^i, \text{ gdzie } i \in \{1, 2, \dots, L + 1\}. \quad (12)$$

Możemy więc wyznaczyć prawdopodobieństwa poszczególnych stanów:

$$p_i = \frac{\rho^i}{\sum_{j=0}^{L+1} \rho^j} = \frac{\rho^i}{\frac{1-\rho^{L+2}}{1-\rho}} = \frac{\rho^i(1-\rho)}{1-\rho^{L+2}}, \text{ gdzie } i \in \{0, 1, \dots, L + 1\}. \quad (13)$$

Otrzymujemy zatem prawdopodobieństwo stanu  $S_0$ :

$$p_0 = \frac{1-\rho}{1-\rho^{L+2}}, \quad (14)$$

prawdopodobieństwo stanu  $S_1$ :

$$p_1 = \frac{\rho(1-\rho)}{1-\rho^{L+2}}, \quad (15)$$

czy też prawdopodobieństwo stanu  $S_{L+1}$ :

$$p_{L+1} = \frac{\rho^{L+1}(1-\rho)}{1-\rho^{L+2}}. \quad (16)$$

Prawdopodobieństwo, że klient nie zostanie obsłużony, jest równe (stan  $S_{L+1}$ ):

$$p_{no} = p_{L+1} = \frac{\rho^{L+1}(1-\rho)}{1-\rho^{L+2}}. \quad (17)$$

Natomiast prawdopodobieństwo, że klient zostanie obsłużony, wyraża się następującą zależnością (stany  $S_0, \dots, S_L$ ):

$$p_{yes} = 1 - p_{no} = 1 - \frac{\rho^{L+1}(1-\rho)}{1-\rho^{L+2}} = \frac{1-\rho^{L+1}}{1-\rho^{L+2}}. \quad (18)$$

Średnia liczba klientów przebywających w poczekalni jest równa:

$$\begin{aligned}
 \hat{v} &= 1 \cdot p_2 + 2 \cdot p_3 + \dots + L \cdot p_{L+1} = \sum_{v=1}^L v p_{v+1} = \sum_{v=1}^L v \frac{\rho^{v+1}(1-\rho)}{1-\rho^{L+2}} \\
 &= \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \sum_{v=1}^L v \rho^{v-1} = \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} (\sum_{v=1}^L \rho^v)' = \\
 &= \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} (\rho^1 + \rho^2 + \rho^3 + \dots + \rho^L)' = \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} [\rho(1 + \rho^1 + \rho^2 + \dots + \rho^{L-1})]' = \\
 &= \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \left[ \rho \left( \frac{1-\rho^L}{1-\rho} \right) \right]' = \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \left[ \left( \frac{\rho-\rho^{L+1}}{1-\rho} \right) \right]' = \\
 &= \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \cdot \frac{[1-(L+1)\rho^L](1-\rho) - (\rho-\rho^{L+1})(-1)}{(1-\rho)^2} = \\
 &= \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \cdot \frac{1-\rho-(L+1)\rho^L+(L+1)\rho^{L+1}+\rho-\rho^{L+1}}{(1-\rho)^2} = \\
 &= \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \cdot \frac{1-L\rho^L+L\rho^{L+1}-\rho^L}{(1-\rho)^2} = \frac{\rho^2(1-\rho)}{1-\rho^{L+2}} \cdot \frac{1+\rho^L(L\rho-L-1)}{(1-\rho)^2}. \tag{19}
 \end{aligned}$$

Ostatecznie otrzymujemy:

$$\hat{v} = \frac{(1+\rho^L(L\rho-L-1))\rho^2}{(1-\rho^{L+2})(1-\rho)}. \tag{20}$$

Wyznamy teraz średnią liczbę klientów w punkcie obsługi:

$$\hat{l} = 1 \cdot p_1 + 1 \cdot p_2 + 1 \cdot p_3 + \dots + 1 \cdot p_{L+1} = 1 \cdot (1 - p_0). \tag{21}$$

A więc:

$$\hat{l} = 1 \cdot \left( 1 - \frac{1-\rho}{1-\rho^{L+2}} \right) = \frac{\rho-\rho^{L+2}}{1-\rho^{L+2}}. \tag{22}$$

Aby wyznaczyć średnią liczbę klientów przebywających w całym systemie kolejkowym możemy skorzystać z następującego wzoru (korzystamy ze wzoru na wartość oczekiwaną zmiennej losowej typu dyskretnego):

$$\begin{aligned}
 \hat{n} &= 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + \dots + (L+1) \cdot p_{L+1} = \sum_{v=1}^{L+1} v p_v = \sum_{v=1}^{L+1} v p_0 \rho^v = \\
 &= p_0 \rho \sum_{v=1}^{L+1} v \rho^{v-1} = p_0 \rho (\sum_{v=1}^{L+1} \rho^v)' = p_0 \rho \left( \frac{\rho-\rho^{L+2}}{1-\rho} \right)' = \\
 &= p_0 \rho \frac{(1-(L+2)\rho^{L+1})(1-\rho) - (\rho-\rho^{L+2})(-1)}{(1-\rho)^2} = \\
 &= \rho \cdot \frac{1-\rho}{1-\rho^{L+2}} \cdot \frac{1-\rho-(L+2)\rho^{L+1}+(L+2)\rho^{L+2}+\rho-\rho^{L+2}}{(1-\rho)^2} = \\
 &= \rho \cdot \frac{1-\rho}{1-\rho^{L+2}} \cdot \frac{1-(L+2)\rho^{L+1}+(L+1)\rho^{L+2}}{(1-\rho)^2} = \frac{\rho(1-(L+2)\rho^{L+1}+(L+1)\rho^{L+2})}{(1-\rho)(1-\rho^{L+2})}. \tag{23}
 \end{aligned}$$

Ostatecznie otrzymujemy:

$$\hat{n} = \frac{\rho(1-(L+2)\rho^{L+1}+(L+1)\rho^{L+2})}{(1-\rho)(1-\rho^{L+2})}. \tag{24}$$

Alternatywnym sposobem na wyznaczenie średniej liczby klientów przebywających w całym systemie kolejkowym jest zsumowanie wielkości  $\acute{v}$  oraz wielkości  $\acute{l}$ . Jest to suma średniej liczby klientów w punkcie obsługi oraz średniej liczby klientów przebywających w poczekalni:

$$\acute{n} = \acute{v} + \acute{l}. \quad (25)$$

Kolejną ważną charakterystyką rozważanego modelu kolejkowego jest średni czas przebywania w poczekalni. Możemy ją wyliczyć w oparciu o prawo Little'a [1, 7]:

$$\acute{w} = \frac{\acute{v}}{\lambda}. \quad (26)$$

Analogicznie wyznaczamy średni czas przebywania klienta w całym systemie kolejkowym:

$$\acute{q} = \frac{\acute{n}}{\lambda}. \quad (27)$$

Ważnym parametrem systemów kolejkowych jest też średni czas przebywania klienta w punkcie obsługi:

$$\acute{s} = p_{yes} \cdot \frac{1}{\mu} = \frac{1-\rho^{L+1}}{1-\rho^{L+2}} \cdot \frac{1}{\mu}. \quad (28)$$

### 4.3. Model kolejkowy M/M/c/0

Kolejnym rozważanym modelem kolejkowym w pracy jest model M/M/c/0 [5, 6]. W modelu tym zakładamy, że:

- rozkład prawdopodobieństwa czasu pomiędzy kolejnymi pojawiającymi się klientami jest wykładniczy o parametrze  $\lambda$ ,
- rozkład prawdopodobieństwa czasu obsługi również jest wykładniczy o parametrze  $\mu$ ,
- jest  $c$  punktów obsługi,
- nie ma miejsc w poczekalni,
- liczba wszystkich klientów jest nieograniczona.

W tabeli 3 przedstawiono interpretację stanów w modelu kolejkowym M/M/c/0.



Tabela 3. Interpretacja stanów w modelu kolejkowym M/M/c/0

Stan	Interpretacja
$S_0$	0 klientów, wolne $c$ punktów obsługi
$S_1$	1 klient, wolne $c - 1$ punktów obsługi
$S_2$	2 klientów, wolne $c - 2$ punktów obsługi
...	...
$S_c$	$c$ klientów, zajęte $c$ punktów obsługi, inne zgłoszenia nie są uwzględniane

Źródło: opracowanie własne.

W modelu tym przyjmujemy, że:

$$\lambda = \lambda_0 = \lambda_1 = \lambda_2 = \dots = \lambda_{c-2} = \lambda_{c-1}, \quad (29)$$

a także, z racji tego, że w tym modelu mamy  $c$  punktów obsługi:

$$\mu_1 = \mu, \mu_2 = 2\mu, \dots, \mu_c = c\mu. \quad (30)$$

W kolejnym kroku wyznaczamy współczynniki  $Q_i$ :

$$Q_0 = 1, \quad Q_i = \frac{\lambda_0 \lambda_1 \lambda_2 \dots \lambda_{i-2} \lambda_{i-1}}{\mu_1 \mu_2 \mu_3 \dots \mu_{i-1} \mu_i}. \quad (31)$$

Korzystając z poprzednich równości, otrzymujemy:

$$Q_i = \frac{\lambda^i}{i! \mu^i} = \frac{1}{i!} \cdot \left(\frac{\lambda}{\mu}\right)^i, \quad (32)$$

a zatem:

$$Q_i = \frac{\rho^i}{i!}. \quad (33)$$

Wyznamy teraz prawdopodobieństwa  $p_i$ . W przypadku skończonej liczby stanów mamy:

$$p_i = \frac{Q_i}{\sum_{j=0}^c Q_j}, \text{ oraz } \sum_{i=0}^c p_i = 1, \quad (34)$$

a więc:

$$p_i = \frac{\frac{\rho^i}{i!}}{\sum_{j=0}^c \frac{\rho^j}{j!}}. \quad (35)$$

Otrzymujemy zatem prawdopodobieństwo stanu  $S_0$ :

$$p_0 = \frac{1}{\sum_{j=0}^c \frac{\rho^j}{j!}}. \quad (36)$$

prawdopodobieństwo stanu  $S_1$ :

$$p_1 = \frac{\rho}{\sum_{j=0}^c \frac{\rho^j}{j!}}, \quad (37)$$

oraz prawdopodobieństwo stanu  $S_c$ :

$$p_c = \frac{\frac{\rho^c}{c!}}{\sum_{j=0}^c \frac{\rho^j}{j!}}. \quad (38)$$

Prawdopodobieństwo, że klient nie zostanie obsłużony, jest równe (stan  $S_c$ ):

$$p_{no} = \frac{\frac{\rho^c}{c!}}{\sum_{j=0}^c \frac{\rho^j}{j!}} = p_c. \quad (39)$$

Natomiast prawdopodobieństwo, że klient zostanie obsłużony, wyraża się następującą zależnością (stany  $S_0, \dots, S_{c-1}$ ):

$$p_{yes} = 1 - p_{no} = 1 - \frac{\frac{\rho^c}{c!}}{\sum_{j=0}^c \frac{\rho^j}{j!}}. \quad (40)$$

W przypadku tego modelu warto również wyznaczyć średnią liczbę klientów w punktach obsługi:

$$\bar{l} = \sum_{i=0}^c i \cdot p_i = \sum_{i=0}^c i \cdot \frac{\frac{\rho^i}{i!}}{\sum_{j=0}^c \frac{\rho^j}{j!}}. \quad (41)$$

#### **4.4. Model kolejkowy M/M/c/L**

Rozważmy teraz model kolejkowy M/M/c/L [5, 7]. W modelu tym zakładamy, że:

- rozkład prawdopodobieństwa czasu pomiędzy kolejnymi pojawiającymi się klientami jest wykładniczy o parametrze  $\lambda$ ,
- rozkład prawdopodobieństwa czasu obsługi również jest wykładniczy o parametrze  $\mu$ ,
- jest  $c$  punktów obsługi,
- jest  $L$  miejsc w poczekalni,
- liczba wszystkich klientów jest nieograniczona.

W tabeli 4 przedstawiono interpretację stanów w modelu kolejkowym M/M/c/L.

Tabela 4. Interpretacja stanów w modelu kolejkowym M/M/c/L

Stan	Interpretacja
$S_0$	0 klientów, wolne $c$ punktów obsługi, wolne $L$ miejsc w poczekalni
$S_1$	1 klient, wolne $c - 1$ punktów obsługi, wolne $L$ miejsc w poczekalni
$S_2$	2 klientów, wolne $c - 2$ punktów obsługi, wolne $L$ miejsc w poczekalni
...	...
$S_c$	$c$ klientów, zajęte $c$ punktów obsługi, wolne $L$ miejsc w poczekalni
$S_{c+1}$	$c + 1$ klientów, zajęte $c$ punktów obsługi, wolne $L - 1$ miejsc w poczekalni
...	...
$S_{c+L}$	$c + L$ klientów, zajęte $c$ punktów obsługi, zajęte $L$ miejsc w poczekalni, inne zgłoszenia nie są uwzględniane

Źródło: opracowanie własne.

W modelu tym przyjmujemy, że:

$$\lambda = \lambda_0 = \lambda_1 = \dots = \lambda_{c+L-2} = \lambda_{c+L-1}, \quad (42)$$

a także:

$$\mu_1 = \mu, \mu_2 = 2\mu, \dots, \mu_c = c\mu, \mu_{c+1} = c\mu, \dots, \mu_{c+L} = c\mu. \quad (43)$$

Wyznamy teraz współczynniki  $Q_i$  w zależności od liczby klientów  $i$ . W przypadku gdy  $0 \leq i \leq c$ , mamy:

$$Q_0 = 1, \quad (44)$$

$$Q_i = \frac{\lambda_0 \lambda_1 \lambda_2 \dots \lambda_{i-2} \lambda_{i-1}}{\mu_1 \mu_2 \mu_3 \dots \mu_{i-1} \mu_i}, \quad (45)$$

$$Q_i = \frac{\lambda^i}{i! \mu^i} = \frac{1}{i!} \cdot \left(\frac{\lambda}{\mu}\right)^i, \quad (46)$$

$$Q_i = \frac{\rho^i}{i!}. \quad (47)$$

W przypadku gdy  $c + 1 \leq i \leq c + L$  mamy:

$$Q_i = \frac{\lambda^i}{(\mu_1 \mu_2 \dots \mu_{c-1} \mu_c)(\mu_{c+1} \mu_{c+2} \dots \mu_{i-1} \mu_i)}, \quad (48)$$

$$Q_i = \frac{\lambda^i}{c! \mu^c \cdot (c\mu)^{i-c}}, \quad (49)$$

$$Q_i = \frac{\lambda^i}{c! \mu^i \frac{c^i}{c^i}}, \quad (50)$$

$$Q_i = \frac{c^c}{c!} \cdot \left(\frac{\rho}{c}\right)^i, \frac{\rho}{c} < 1. \quad (51)$$

Prawdopodobieństwa poszczególnych stanów  $p_i$  są określone następującą formułą:

$$p_i = Q_i p_0, \sum_{k=0}^{L+c} p_k = 1, \quad (52)$$

a więc mamy:

$$p_0 = \frac{1}{\sum_{k=0}^{L+c} Q_k}, \quad (53)$$

$$p_i = \frac{Q_i}{\sum_{k=0}^{L+c} Q_k}. \quad (54)$$

Prawdopodobieństwo, że klient nie zostanie obsłużony jest równe (stan  $S_{L+c}$ ):

$$p_{no} = \frac{Q_{L+c}}{\sum_{k=0}^{L+c} Q_k} = p_{L+c}. \quad (55)$$

Natomiast prawdopodobieństwo, że klient zostanie obsłużony wyraża się następującą zależnością:

$$p_{yes} = 1 - p_{no} = \sum_{i=0}^{L+c-1} p_i. \quad (56)$$

Wyznamy teraz średnią liczbę klientów w punktach obsługi:

$$\acute{l} = 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + \dots + (c-1) \cdot p_{c-1} + c \cdot p_c + c \cdot p_{c+1} + \dots + c \cdot p_{c+L} \quad (57)$$

Możemy zatem napisać:

$$\acute{l} = \sum_{k=1}^c k p_k + \sum_{k=c+1}^{c+L} c p_k. \quad (58)$$

Następnie:

$$\acute{l} = \sum_{k=1}^c k \frac{\rho^k}{k!} p_0 + \sum_{k=c+1}^{c+L} c \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^k p_0. \quad (59)$$

Ostatecznie otrzymujemy:

$$\acute{l} = p_0 \left[ \sum_{k=1}^c \frac{\rho^k}{(k-1)!} + \sum_{k=c+1}^{c+L} \frac{c^{c+1}}{c!} \left(\frac{\rho}{c}\right)^k \right]. \quad (60)$$

Średnia liczba klientów przebywających w poczekalni jest równa:

$$\acute{v} = 1 \cdot p_{c+1} + 2 \cdot p_{c+2} + 3 \cdot p_{c+3} + \dots + L \cdot p_{c+L}. \quad (61)$$

Możemy więc napisać:

$$\acute{v} = \sum_{k=1}^L k \cdot p_{c+k}. \quad (62)$$

Następnie:

$$\acute{v} = \sum_{k=1}^L k \cdot p_0 \cdot \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^{c+k}, \quad (63)$$

$$\dot{v} = p_0 \cdot \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^{c+1} \sum_{k=1}^L k \cdot \left(\frac{\rho}{c}\right)^{k-1}, \quad (64)$$

$$\dot{v} = p_0 \cdot \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^{c+1} \left(\sum_{k=1}^L \left(\frac{\rho}{c}\right)^k\right)'. \quad (65)$$

Ostatecznie:

$$\dot{v} = p_0 \cdot \frac{c^c}{c!} \cdot \left(\frac{\rho}{c}\right)^{c+1} \cdot \frac{1 + \left(\frac{\rho}{c}\right)^L \left(L \cdot \frac{\rho}{c} - L - 1\right)}{\left(1 - \frac{\rho}{c}\right)^2}. \quad (66)$$

Aby wyznaczyć średnią liczbę klientów przebywających w całym systemie kolejkowym możemy skorzystać z formuły:

$$\dot{n} = 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + \dots + (c-1) \cdot p_{c-1} + c \cdot p_c + (c+1) \cdot p_{c+1} + \dots + (c+L) \cdot p_{c+L}. \quad (67)$$

Możemy zatem napisać:

$$\dot{n} = \sum_{k=1}^c k p_k + \sum_{k=c+1}^{c+L} k p_k. \quad (68)$$

Następnie:

$$\dot{n} = \sum_{k=1}^c k \frac{\rho^k}{k!} p_0 + \sum_{k=c+1}^{c+L} k \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^k p_0. \quad (69)$$

Ostatecznie otrzymujemy:

$$\dot{n} = p_0 \left[ \sum_{k=1}^c \frac{\rho^k}{(k-1)!} + \sum_{k=c+1}^{c+L} k \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^k \right]. \quad (70)$$

Alternatywnym sposobem na wyznaczenie średniej liczby klientów przebywających w całym systemie kolejkowym jest zsumowanie wielkości  $\dot{v}$  oraz  $\dot{l}$ :

$$\dot{n} = \dot{v} + \dot{l}. \quad (71)$$

Kolejną ważną charakterystyką rozważanego modelu kolejkowego jest średni czas przebywania klienta w poczekalni. Możemy ją wyliczyć w oparciu o prawo Little'a [1,7]:

$$\dot{w} = \frac{\dot{v}}{\lambda}. \quad (72)$$

Analogicznie wyznaczamy średni czas przebywania klienta w całym systemie kolejkowym:

$$\dot{q} = \frac{\dot{n}}{\lambda}. \quad (73)$$

Ważnym parametrem systemów kolejkowych jest też średni czas przebywania klienta w punkcie obsługi:

$$\acute{s} = p_{yes} \cdot \frac{1}{\mu}. \quad (74)$$

Mamy więc:

$$\acute{s} = (1 - p_{L+c}) \cdot \frac{1}{\mu}, \quad (75)$$

czyli:

$$\acute{s} = \frac{1}{\mu} \cdot \left(1 - p_0 \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^{L+c}\right). \quad (76)$$

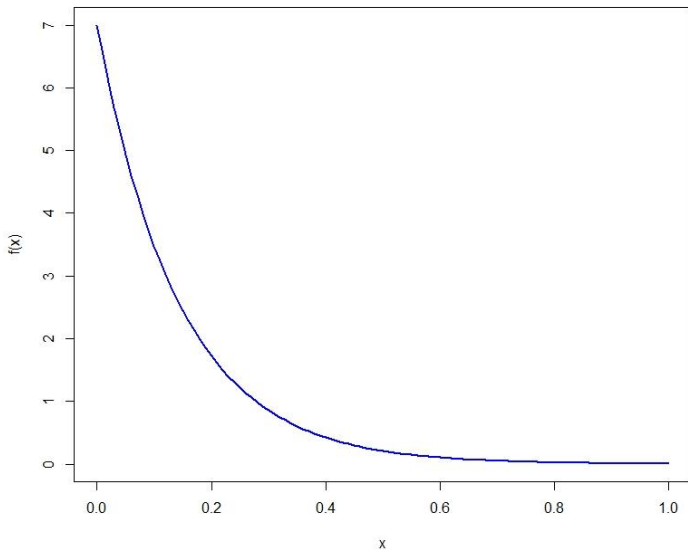
Alternatywny sposób na wyznaczenie średniego czasu przebywania klienta w punkcie obsługi opiera się na prawie Little'a:

$$\acute{s} = \frac{i}{\lambda}. \quad (77)$$

## 5. Zastosowanie teorii kolejek do poprawy jakości obsługi klientów poczty

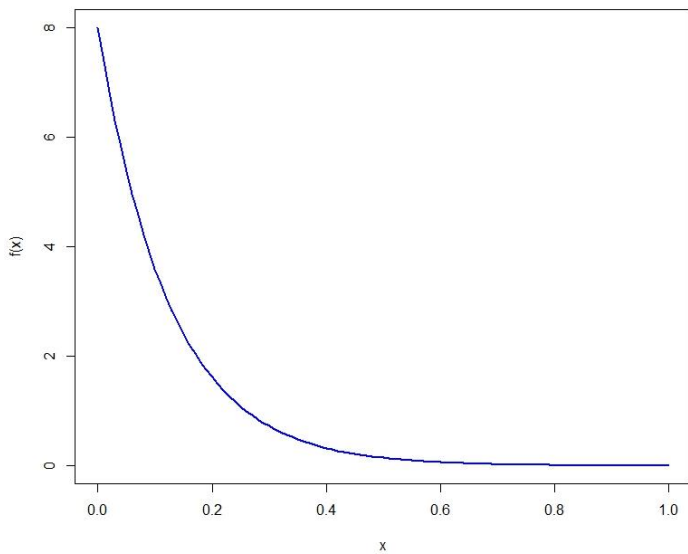
W celu praktycznego wykorzystania przedstawionej teorii, przeanalizujemy teraz problem dotyczący omawianych modeli kolejkowych. Aby usprawnić proces obliczeń, wszystkie potrzebne wzory i formuły zostały zaprogramowane w środowisku RStudio [10, 15]. Wyniki zostały podane w zaokrągleniu. Ponadto, wszystkie wykresy zamieszczone w pracy również zostały wygenerowane za pomocą RStudio [11, 12].

Rozważmy pewną pocztę, w której jest tylko jeden punkt obsługi klienta oraz  $L = 15$  miejsc w poczekalni. Rozkład prawdopodobieństwa czasu pomiędzy kolejnymi pojawiającymi się klientami jest wykładniczy o parametrze  $\lambda = 7$  (rysunek 1), a rozkład prawdopodobieństwa czasu obsługi jest wykładniczy o parametrze  $\mu = 8$  (rysunek 2).



**Rysunek 1. Wykres funkcji gęstości rozkładu czasu pomiędzy kolejnymi pojawiającymi się klientami – poczta**

Źródło: opracowanie własne.



**Rysunek 2. Wykres funkcji gęstości rozkładu czasu obsługi – poczta**

Źródło: opracowanie własne.

Na poczcie planowane jest wdrożenie pewnych zmian, aby średni czas pobytu klienta w całym systemie kolejkowym spadł poniżej 9 minut.

Rozważane są jednak dwie drogi – wysłanie jedyne go pracownika na płatne szkolenia (każde szkolenie poprawia parametr intensywności obsługi o 1 i kosztuje 1000 zł) lub utworzenie nowych punktów obsługi klienta (jeden nowy punkt obsługi wiąże się jednak z kosztem 3000 zł).

Którą opcję należy wybrać, aby średni czas pobytu klienta w całym systemie kolejkowym spadł poniżej 9 minut (około 0,15 h) możliwie najniższym kosztem?

Przyjrzyjmy się najpierw, jak wyglądają obecne charakterystyki systemu kolejkowego na poczcie. Do analizy podanej sytuacji wykorzystamy model kolejkowy M/M/c/L.

Obciążenie punktu obsługi:

$$\rho = \frac{\lambda}{\mu} = 0,875. \quad (78)$$

Prawdopodobieństwo, że klient nie zostanie obsłużony:

$$p_{no} = \frac{Q_{L+c}}{\sum_{k=0}^{L+c} Q_k} = p_{L+c} \approx 0,016. \quad (79)$$

Prawdopodobieństwo, że klient zostanie obsłużony:

$$p_{yes} = 1 - p_{no} = \sum_{i=0}^{L+c-1} p_i \approx 0,984. \quad (80)$$

Średnia liczba klientów w punkcie obsługi:

$$\acute{l} = p_0 \left[ \sum_{k=1}^c \frac{\rho^k}{(k-1)!} + \sum_{k=c+1}^{c+L} \frac{c^{c+1}}{c!} \left(\frac{\rho}{c}\right)^k \right] \approx 0,86. \quad (81)$$

Średnia liczba klientów przebywających w poczekalni:

$$\acute{v} = p_0 \cdot \frac{c^c}{c!} \cdot \left(\frac{\rho}{c}\right)^{c+1} \cdot \frac{1 + \left(\frac{\rho}{c}\right)^L \left(L \frac{\rho}{c} - L - 1\right)}{\left(1 - \frac{\rho}{c}\right)^2} \approx 4,18. \quad (82)$$

Średnia liczba klientów przebywających w całym systemie kolejkowym:

$$\acute{n} = p_0 \left[ \sum_{k=1}^c \frac{\rho^k}{(k-1)!} + \sum_{k=c+1}^{c+L} k \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^k \right] \approx 5,04. \quad (83)$$

$$\acute{n} = \acute{v} + \acute{l} \approx 5,04. \quad (84)$$

Średni czas przebywania klienta w poczekalni (w oparciu o prawo Little'a):

$$\acute{w} = \frac{\acute{v}}{\lambda} \approx 0,597. \quad (85)$$

Średni czas przebywania klienta w całym systemie kolejkowym:

$$\acute{q} = \frac{\acute{n}}{\lambda} \approx 0,720. \quad (86)$$



Średni czas przebywania klienta w punkcie obsługi:

$$\hat{s} = p_{yes} \cdot \frac{1}{\mu} = \frac{1}{\mu} \cdot \left(1 - p_0 \frac{c^c}{c!} \left(\frac{\rho}{c}\right)^{L+c}\right) \approx 0,123. \quad (87)$$

Z prawa Little'a:

$$\hat{s} = \frac{\hat{l}}{\lambda} \approx 0,123. \quad (88)$$

### 5.1. Droga I – wysłanie pracownika na płatne szkolenia

Wartości poszczególnych charakterystyk modelu kolejkowego M/M/c/L w zależności od liczby szkoleń odbytych przez pracownika poczty, przy  $\lambda = 7$ ,  $L = 15$  oraz  $c = 1$ , zostały przedstawione w tabelach 5 oraz 6.

**Tabela 5. Wartości poszczególnych charakterystyk modelu kolejkowego M/M/c/L w zależności od liczby szkoleń odbytych przez pracownika poczty, przy  $\lambda = 7$ ,  $L = 15$  oraz  $c = 1$**

Liczba szkoleń	Koszt	$\mu$	$\hat{l}$	$\hat{v}$	$\hat{n}$
0	0 zł	8	0,8605986	4,180814	5,041413
1	1000 zł	9	0,7746341	2,484878	3,259512
2	2000 zł	10	0,6993005	1,594393	2,293694
3	3000 zł	11	0,6361962	1,105976	1,742172
4	4000 zł	12	0,5832896	0,8149277	1,398217
5	5000 zł	13	0,5384491	0,6277604	1,166209
6	6000 zł	14	0,4999962	0,4998741	0,9998703

Źródło: opracowanie własne.

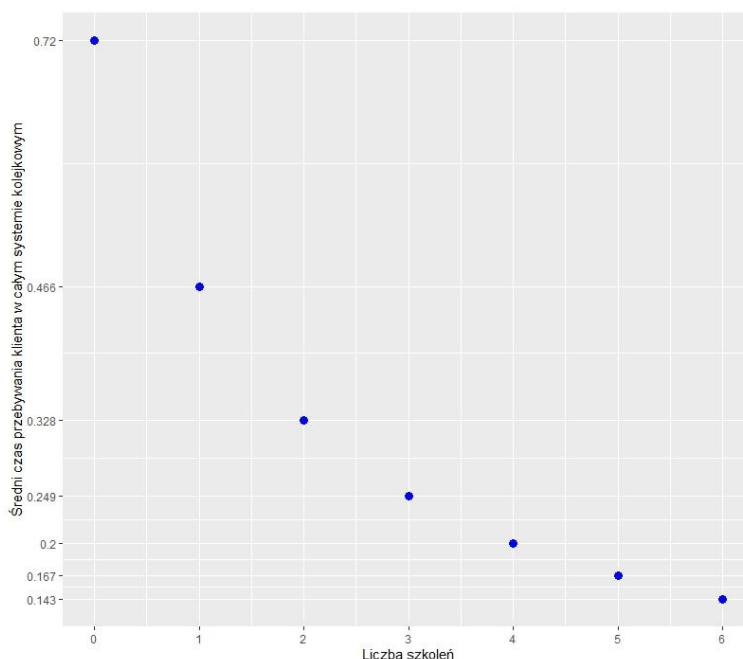
**Tabela 6. Wartości poszczególnych charakterystyk modelu kolejkowego M/M/c/L w zależności od liczby szkoleń odbytych przez pracownika poczty, przy  $\lambda = 7$ ,  $L = 15$  oraz  $c = 1$**

Liczba szkoleń	Koszt	$\mu$	$\hat{w}$	$\hat{q}$	$\hat{s}$
0	0 zł	8	0,5972592 h	0,7202019 h	0,1229427 h
1	1000 zł	9	0,3549825 h	0,4656446 h	0,110662 h
2	2000 zł	10	0,2277705 h	0,3276706 h	0,09990007 h
3	3000 zł	11	0,1579966 h	0,2488817 h	0,09088517 h
4	4000 zł	12	0,1164182 h	0,1997453 h	0,08332709 h
5	5000 zł	13	0,08968005 h	0,1666014 h	0,0769213 h
6	6000 zł	14	0,07141059 h	0,1428386 h	0,07142803 h

Źródło: opracowanie własne.

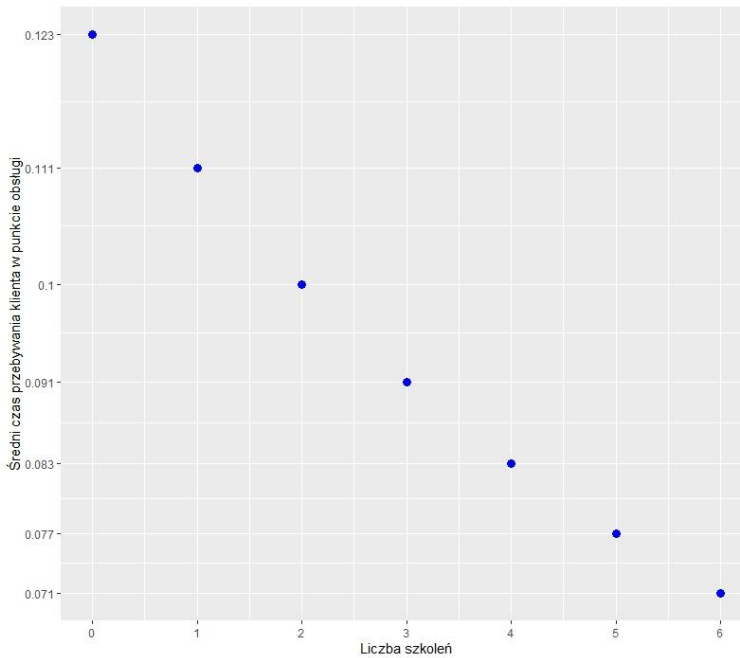
Optymalnym rozwiązaniem opisywanego problemu w przypadku wyboru drogi I jest więc wysłanie pracownika poczty na 6 szkoleń przy łącznym koszcie równym 6000 zł. Czas pobytu klienta w całym systemie kolejkowym ( $\bar{q}$ ) wynosi wtedy 0,1428386 h, a więc spada poniżej 0,15 h.

Zależność pomiędzy liczbą płatnych szkoleń odbytych przez pracownika poczty a średnim czasem pobytu klienta w całym systemie kolejkowym została przedstawiona na rysunku 3, natomiast zależność pomiędzy liczbą odbytych szkoleń a średnim czasem pobytu klienta w punkcie obsługi – na rysunku 4.



**Rysunek 3. Wykres zależności pomiędzy liczbą płatnych szkoleń odbytych przez pracownika a średnim czasem pobytu klienta w całym systemie kolejkowym**

Źródło: opracowanie własne.



Rysunek 4. Wykres zależności pomiędzy liczbą płatnych szkoleń odbytych przez pracownika poczty a średnim czasem pobytu klienta w punkcie obsługi

Źródło: opracowanie własne.

## 5.2. Droga II – otwarcie nowych punktów obsługi klienta

Wartości poszczególnych charakterystyk modelu kolejkowego M/M/c/L w zależności od liczby punktów obsługi klienta, przy  $\lambda = 7$ ,  $L = 15$  oraz  $\mu = 8$ , zostały przedstawione w tabelach 7 oraz 8.

Tabela 7. Wartości poszczególnych charakterystyk modelu kolejkowego M/M/c/L w zależności od liczby punktów obsługi klienta, przy  $\lambda = 7$ ,  $L = 15$  oraz  $\mu = 8$

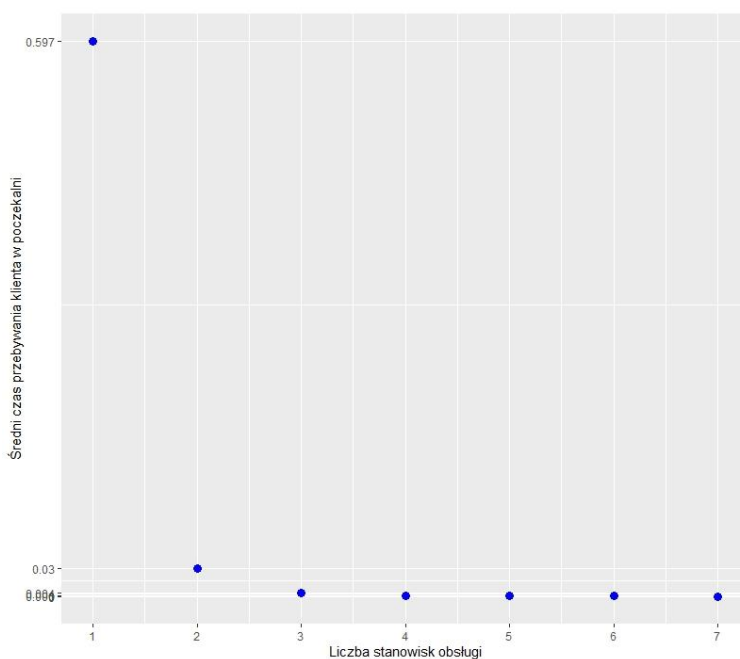
$c$	Łączny koszt	$\hat{l}$	$\hat{v}$	$\hat{n}$
1	0 zł	0,8605986	4,180814	5,041413
2	3000 zł	0,8749995	0,2071177	1,082117
3	6000 zł	0,875	0,02687125	0,9018712
4	9000 zł	0,875	0,003646245	0,8786462
5	12000 zł	0,875	0,0004580849	0,8754581
6	15000 zł	0,875	0,0000519	0,8750519
7	18000 zł	0,875	0,0000053	0,8750053

Źródło: opracowanie własne.

**Tabela 8. Wartości poszczególnych charakterystyk modelu kolejkowego M/M/c/L w zależności od liczby punktów obsługi klienta, przy  $\lambda = 7$ ,  $L = 15$  oraz  $\mu = 8$**

<b>c</b>	<b>Łączny koszt</b>	<b><math>\acute{w}</math></b>	<b><math>\acute{q}</math></b>	<b><math>\acute{s}</math></b>
1	0 zł	0,5972592 h	0,7202019 h	0,1229427 h
2	3000 zł	0,02958824 h	0,1545882 h	0,1249999 h
3	6000 zł	0,00383875 h	0,1288387 h	0,125 h
4	9000 zł	0,0005208921 h	0,1255209 h	0,125 h
5	12000 zł	0,0000654 h	0,1250654 h	0,125 h
6	15000 zł	0,0000074 h	0,1250074 h	0,125 h
7	18000 zł	0,0000008 h	0,1250008 h	0,125 h

Źródło: opracowanie własne.



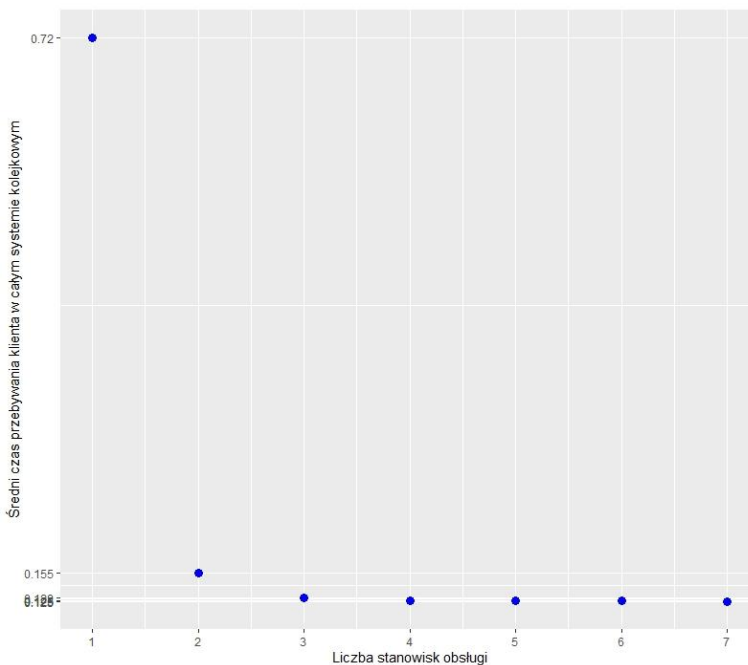
**Rysunek 5. Wykres zależności pomiędzy liczbą punktów obsługi klienta a średnim czasem pobytu klienta w poczekalni**

Źródło: opracowanie własne.

Optymalnym rozwiązaniem opisywanego problemu w przypadku wyboru drogi II jest więc praca trzech równoległych punktów obsługi klienta. Jeden już funkcjonuje, należy zatem otworzyć dwa nowe stanowiska, przy łącznym koszcie równym 6000 zł.

Czas pobytu klienta w całym systemie kolejkowym ( $\bar{q}$ ) wynosi wtedy 0,1288387 h, a więc spada poniżej 0,15 h.

Zależność pomiędzy liczbą punktów obsługi klienta a średnim czasem pobytu klienta w poczekalni została przedstawiona na rysunku 5, natomiast zależność pomiędzy liczbą punktów obsługi klienta a średnim czasem pobytu klienta w całym systemie kolejkowym na rysunku 6.



**Rysunek 6. Wykres zależności pomiędzy liczbą punktów obsługi klienta a średnim czasem pobytu klienta w całym systemie kolejkowym**

Źródło: opracowanie własne.

### 5.3. Wybór optymalnego rozwiązania

W celu zdecydowania, którą opcję należy wybrać, tak aby średni czas pobytu klienta w całym systemie kolejkowym spadł poniżej 9 minut (około 0,15 h), możliwie najniższym kosztem, porównamy teraz optymalne rozwiązania dla poszczególnych wariantów zmian.

Tabela 9. Wybór optymalnego rozwiązania

Droga	Koszt	$\acute{w}$	$\acute{q}$	$\acute{s}$
I – 6 szkoleń pracownika	6000 zł	0,0714106 h	0,1428386 h	0,071428 h
II – 2 nowe punkty obsługi	6000 zł	0,0038388 h	0,1288387 h	0,125 h

Źródło: opracowanie własne.

Jak widzimy, koszty w przypadku obu możliwości są takie same. Średni czas pobytu klienta w poczekalni jest znacznie krótszy w przypadku wariantu II, a średni czas pobytu klienta w punkcie obsługi jest krótszy w przypadku wariantu I. Decydujący jest jednak średni czas pobytu klienta w całym systemie kolejkowym ( $\acute{q}$ ).

Wybór drogi II, czyli utworzenie 2 nowych punktów obsługi, zapewnia nam krótszy średni czas pobytu klienta w całym systemie kolejkowym, a więc jest to optymalne rozwiązanie rozważanego problemu.

## 6. Podsumowanie

Teoria obsługi masowej zajmuje się analizowaniem systemów kolejkowych. Jest wykorzystywana przy podejmowaniu decyzji biznesowych. W artykule przedstawiono najważniejsze elementy teorii kolejek, na przykład prawo Little'a.

W pracy zaprezentowane zostało zastosowanie teorii kolejek do poprawy jakości obsługi klientów poczty. Eksperyment polegał na wyborze odpowiednich zmian na poczcie, mających na celu skrócenie czasu pobytu klienta w całym systemie kolejkowym do poniżej 9 minut, przy możliwie najniższym koszcie. Rozważane były dwie drogi – wysłanie pracownika na płatne szkolenia, aby podnieść parametr intensywności obsługi lub otwarcie nowych punktów obsługi klienta.

Dopasowany został odpowiedni model kolejkowy. Policzone zostały również najważniejsze charakterystyki rozważanych systemów kolejkowych. Aby usprawnić proces obliczeń, wszystkie potrzebne formuły i zależności zostały zaprogramowane w środowisku RStudio. Wszystkie wykresy zamieszczone w pracy także zostały wygenerowane za pomocą tego programu.

Podsumowując, modele kolejkowe charakteryzują się szerokim zakresem praktycznych możliwości w teorii obsługi masowej. Na ich podstawie da się przewidzieć pewne wielkości, między innymi długość kolejki czy też czas oczekiwania klienta na obsługę. To kluczowe narzędzie podczas podejmowania ważnych decyzji biznesowych. Dzięki nim możliwa jest optymalizacja systemu kolejkowego w danym przedsiębiorstwie, która zapewnia wzrost efektywności obsługi klientów oraz wzrost zysków.

## Literatura

- [1] Cooper R. B., *Introduction to queueing theory*, Boca Raton, Florida, 1981.
- [2] Filipowicz B., *Modelowanie i analiza sieci kolejkowych*, AGH, Kraków, 1997.
- [3] Little J. D. C., *Little's law as viewed on its 50th anniversary*, "Operations Research" 2011, vol. 59, no. 3, pp. 536–549.
- [4] Obretenow A., Dimitrow B., *Teoria masowej obsługi: poradnik*, Państwowe Wydawnictwo Naukowe, Warszawa, 1989.
- [5] Oniszczyk W., *Metody modelowania*, Politechnika Białostocka, Białystok 1995.
- [6] Rolski T., *Queues and communication networks. An outline of continuous time theory*, Mathematical Institute, Wrocław 2008.
- [7] Shortle J. F., Thompson J. M., Gross D., Harris C. M., *Fundamentals of queueing theory*, Hoboken, New Jersey 2018.
- [8] Sztrik J., *Basic Queueing Theory*, Debrecen, 2012.
- [9] Zukerman M., *Introduction to queueing theory and stochastic teletraffic models*, Hong Kong, 2021.
- [10] Gunther N., *Simulating a Queue in R*, <https://www.r-bloggers.com/2010/05/simulating-a-queue-in-r/> (dostęp: 22.06.2022).
- [11] Salazar R., *Queueing Models with R. Exploring the "queueing" R package*, <https://towardsdatascience.com/queueing-models-with-r-a794c78e6820> (dostęp: 22.06.2022).
- [12] O'Connor M., *Introduction to Queuing Theory*, <https://rpubs.com/moconnor13/QueuingTheoryIntro> (dostęp: 22.06.2022)
- [13] Marshall G., *Proof: Little's Law (why to Limit WIP)*, <https://fuza.co.uk/2016/10/17/proof-littles-law-why-to-limit-wip/> (dostęp: 22.06.2022).
- [14] *System M/M/1/L*, Politechnika Białostocka, [http://aragorn.pb.bialystok.pl/~walenty/modelowanie/Pliki\\_Pdf/SystemMM1L.pdf](http://aragorn.pb.bialystok.pl/~walenty/modelowanie/Pliki_Pdf/SystemMM1L.pdf) (dostęp: 22.06.2022).
- [15] Canadilla P., *Analysis of Queueing Networks and Models*, <https://cran.r-project.org/web/packages/queueing/queueing.pdf> (dostęp: 22.06.2022).
- [16] *Queueing Models – M/M/1 Queuing System*, Universal Teacher Publications, <http://www.universalteacherpublications.com/univ/ebooks/or/Ch10/mm1ex.htm> (dostęp: 22.06.2022).
- [17] Wolny M., *Teoria Masowej Obsługi (Teoria Kolejek)*, [http://dydaktyka.polsl.pl/roz6/mwolny/Shared%20Documents/Badania%20operacyjnej/8\\_Teoria\\_Kolejek.pdf](http://dydaktyka.polsl.pl/roz6/mwolny/Shared%20Documents/Badania%20operacyjnej/8_Teoria_Kolejek.pdf) (dostęp: 22.06.2022).

## **Queueing theory application for improving the quality of mail customer service**

### **Abstract**

The paper presents selected queueing models along with their practical applications. The work contains the most important elements of mass service theory. In addition, a practical problem was designed and solved, to which the appropriate queueing model was matched and its most important characteristics were calculated.

*Keywords: queueing models, Little's law, queueing theory.*





**Jakub Kisała<sup>1</sup>**

## **Technologia i badania struktur cienkowarstwowych wykazujących zjawisko gigantycznego magnetooporu**

### **Streszczenie**

Praca dotyczy produkcji oraz badań struktur cienkowarstwowych wykazujących zjawisko gigantycznego magnetooporu. W ramach działań Koła Naukowego Elektroników i Mechatroników SEMICON Politechniki Lubelskiej produkowane są struktury warstwowe, których wartość rezystancji stałoprądowej zależna jest od natężenia zewnętrznego pola magnetycznego. W tym celu wykorzystywana jest technologia rozpylania magnetonowego. Aby możliwe było zaobserwowanie zjawiska gigantycznego magnetooporu ważne jest odpowiednie zaprojektowanie struktury, gdyż grubości poszczególnych warstw sięgają rzędu kilku nanometrów. Struktury tego typu mają potencjonalne zastosowanie jako czujniki pola magnetycznego. W ramach działań koła projektowane są także urządzenia mające dokonywać pomiaru rezystancji badanych struktur lub służące jako stanowisko pomiarowe umożliwiając uzyskanie zmiennego co do wartości stałego pola magnetycznego za pomocą magnesów neodymowych lub elektromagnesów trzymających.

*Słowa kluczowe: rozpylanie magnetonowe, stałe pole magnetyczne, cienkie warstwy, magnetoopór.*

### **Wstęp**

Koło naukowe Elektroników i Mechatroników SEMICON działa na Wydziale Elektrotechniki i Informatyki Politechniki Lubelskiej od 2012 roku. Działalność koła zawsze polegała na wspieraniu studentów wykazujących zainteresowanie dodatkowymi projektami, chcących zaprezentować je większemu gronu. W ostatnim czasie prace koła, oprócz pomocy studentom w rozwijaniu ich pomysłów w osobnych projektach, skupiły się na otrzymanych grantach naukowych dotyczących zjawiska gigantycznego magnetooporu występującego w strukturach cienkowarstwowych.

### **1. Gigantyczny magnetoopór**

Gigantyczny magnetoopór (ang. *Giant Magnetoresistance*, GMR), zwany również gigantyczną magnetorezystancją, należy do grupy zjawisk ogólnie okre-

---

<sup>1</sup> Mgr inż. Jakub Kisała, Katedra Elektroniki i Technik Informacyjnych, Wydział Elektrotechniki i Informatyki, Politechnika Lubelska, j.kisala@pollub.pl.

ślanych jako magnetorezystancyjne. Charakteryzują je zmiany wartości rezystancji stałoprądowej materiału lub struktury pod wpływem działania zewnętrznego pola magnetycznego. Wyróżnia się kilka rodzajów magnetooporu, m.in. magnetoopór anizotropowy, gigantyczny, tunelowy, kolosalny. Ich podstawy działania mogą być diametralnie różne, są także możliwe do zaobserwowania w innych materiałach lub strukturach.

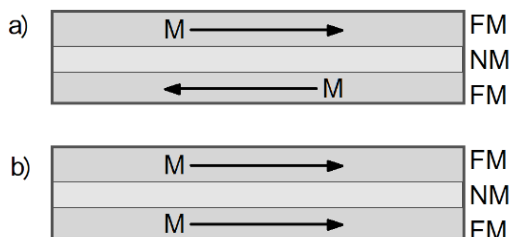
Nazwa gigantyczny magnetoopór została wykorzystana przez jednego z dwóch naukowców prowadzących równoległe badania w dwóch niezależnych zespołach w Niemczech oraz we Francji. W 1988 roku obydwaj naukowcy, Alberta Fert oraz Petera Grunberg, opublikowali stworzone przy pomocy swoich zespołów prace, dzięki którym w 2007 otrzymali Nagrodę Nobla z dziedziny fizyki. Określenie „gigamagnetyczny” nawiązuje do znacznych zmian rezystancji stałoprądowych badanych struktur warstwowych w stosunku do zmian rezystancji w ramach znanego do tej pory zjawiska anizotropowego magnetooporu. O ile anizotropowy magnetoopór obserwowany jest w materiałach ferromagnetycznych, o tyle gigantyczny magnetoopór zaobserwowany został w konkretnie w tym celu zaprojektowanych strukturach cienkowarstwowych.

Struktury cienkowarstwowe, wykazujące zjawisko gigantycznej magnetorezystancji, to struktury składające się z warstw ferromagnetycznych przedzielonych metalicznymi warstwami niemagnetycznymi, a więc warstwami materiałów nie wykazujących silnych właściwości magnetycznych tj. paramagnetyki i diamagnetyki. Jednakże oba materiały, z których wykonane są warstwy, muszą charakteryzować się wysoką przewodnością elektryczną. Najprostszą strukturę, w której możliwe jest do zaobserwowania zjawisko gigantycznego magnetooporu, stanowi trójwarstwowa struktura FM/NM/FM, gdzie FM oznacza materiał ferromagnetyczny, a NM materiał niemagnetyczny.

Źródłem zmian rezystancji stałoprądowej struktury warstwowej w ramach zjawiska gigantycznego magnetooporu jest zależność rozpraszania elektronów przewodzenia w warstwach ferromagnetycznych i na interfejsach FM/NM od zgodności ich spinu z kierunkiem namagnesowania warstwy ferromagnetycznej. Wyróżnić można dwie graniczne konfiguracje ustawienia domen tego typu podstawowej struktury, przedstawione na rysunku 1.

W konfiguracji antyrównoległej (rys. 1a) elektrony, w zależności od spinu, rozpraszane są częściej w co drugiej warstwie. Konfiguracja ta ma miejsce, gdy kąt między kierunkiem namagnesowania naprzemiennych warstw ferromagnetycznych wynosi  $180^\circ$ . Jest to konfiguracja domyślna, w idealnej strukturze zachodzi ona dla zerowego zewnętrznego pola magnetycznego. Konfiguracja równoległa (rys. 1b) charakteryzuje się najmniejszym oporem elektrycznym dla danej struktury, ponieważ elektrony o spinie zgodnym z kierunkiem namagnesowania obu warstw ferromagnetycznych rozpraszane są rzadziej od elektronów z odwrotnym spinem. Działanie zewnętrznego pola magnetycznego na taką strukturę warstwową doprowadza do zmiany namagnesowań poszczególnych warstw,

zgodnie z kierunkiem działania pola, co przekłada się na obserwowane zmiany rezystancji struktury.



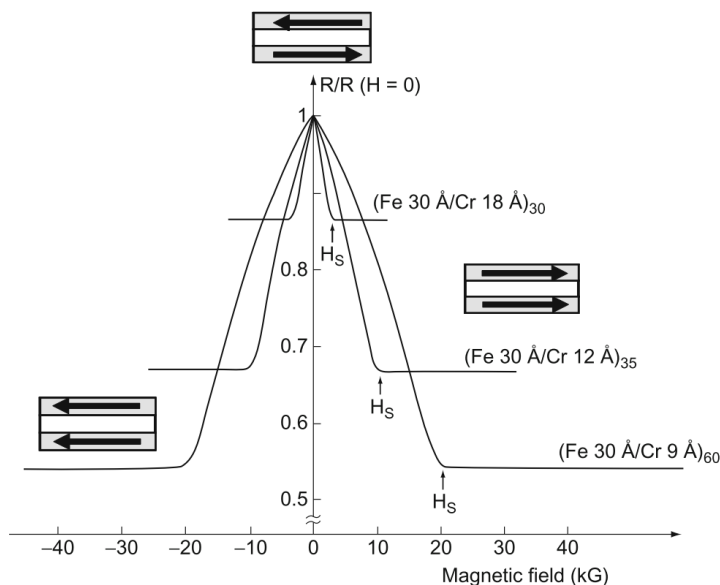
**Rysunek 1. Pojedyncza struktura warstwowa FM/NM/FM w dwóch konfiguracjach namagnesowań warstw ferromagnetycznych: a) antyrównoległa, b) równoległa (FM – materiał ferromagnetyczny, NM – materiał niemagnetyczny, M – zwrot wypadkowego namagnesowania warstwy)**

Źródło: opracowanie własne.

Kolejność i grubość warstw struktury GMR jest kluczowa dla wystąpienia tego efektu magnetorezystancyjnego. Szczególnie ważna jest grubość warstwy niemagnetycznej. Wyróżniane są dwie główne konfiguracje, CPP i CIP, różniące się kierunkiem natężenia prądu przepływającego przez strukturę. W zależności od konfiguracji struktury, chcąc osiągnąć największe możliwe zmiany rezystancji, wartość ta powinna być mniejsza od średniej drogi swobodnej lub długości dyfuzji spinu w danym materiale niemagnetycznym [1]. Elektrony rozpraszane są głównie w warstwach ferromagnetycznych, a przedzielająca warstwa niemagnetyczna ma mieć na nie jak najmniejszy wpływ [13]. Jej działaniem jest magnetyczne odizolowanie dwóch sąsiednich warstw. Jedną z ról warstwy niemagnetycznej jest także doprowadzenie do sprzężenia antyferromagnetycznego między przedzielonymi warstwami. Grubość tej warstwy przekłada się bezpośrednio na odległość pomiędzy warstwami ferromagnetycznymi, a więc na siłę ich wzajemnego oddziaływania. Wiele z parametrów tego typu struktur wpływa na amplitudę zmian rezystancyjnych w ramach zjawiska GMR, są to m.in. zastosowane materiały, grubości poszczególnych warstw, kształt struktury, liczba warstw, zastosowanie dodatkowych warstw, sieć krystalograficzna, temperatura.

Przedstawiony na rysunku 2 wykres jest uznawany za wzorcowy dla gigantycznego magnetooporu. Zespół francuskich badaczy przygotował i zbadał szereg struktur wielowarstwowych naprzemiennych warstw żelaza jako ferromagnetyku oraz chromu jako materiału metalicznego nieprzejawiającego własności ferromagnetycznych. Kształt zmian rezystancji uzależniony jest od indukcji pola magnetycznego. Charakterystyka ta przypominać może kształtem krzywą Gaussa. Obserwowany jest punkt graniczny  $H_s$ , powyżej którego zwiększanie wartości indukcji pola magnetycznego nie wywołuje dalszych zmian rezystancji. Odpowiada on indukcji nasycenia, a więc polu, powyżej którego nie następuje już

dalsza zmiana konfiguracji magnetycznej warstw ferromagnetycznych. Zmiany rezystancji struktury, w przypadku gigantycznego magnetooporu, rozpatrywać można w zależności od kąta między kierunkami namagnesowań naprzemiennych warstw ferromagnetycznych.



**Rysunek 2. Zjawisko gigantycznego magnetooporu – zmiany rezystancji struktury warstwowym Fe/Cr w zależności od wartości indukcja zewnętrznego pola magnetycznego**

Źródło: J. Inoue, *GMR, TMR and BMR*, [in:] ed. T. Shinjo, *Nanomagnetism and Spintronics*, Elsevier Science, 2009, pp. 15–92.

## 2. Technologia struktur cienkowarstwowych

Struktury wykazujące zjawisko gigantycznego magnetooporu do wyprodukowania, w porównaniu z wytwarzaniem elementów półprzewodnikowych, nie wymagają procesów technologicznych zachodzących w wysokich temperaturach, tj. dyfuzji lub implantacji. Produkcja struktur GMR jest stosunkowo prosta: już kilka procesów litograficznych wystarczy, aby uzyskać funkcjonalne urządzenie [3]. W celu uzyskania założonych efektów produkcji struktur GMR wymagane jest dobranie odpowiedniej metody zapewniającej kontrolę takich parametrów, jak czystość warstw i ich grubość. Metody wytwarzania cienkowarstwowych struktur GMR podzielić można na trzy główne grupy: elektrodepozycję, chemiczne (ang. *chemical vapour deposition*, CVD) i fizyczne osadzanie z fazy lotnej (ang. *physical vapour deposition*, PVD).

Elektrodepozycja, zwana także elektroosadzaniem, to technika polegająca na osadzaniu materiału z elektrolitu za pomocą umieszczonych w nim dwóch elektrod.

Metoda ta jest szeroko stosowana zarówno w przemyśle, jak i w badaniach naukowych. Jest ona przystępna i ekonomiczna, przy jednoczesnej możliwości otrzymania warstw wysokiej jakości [4, 6].

Chemicznym osadzaniem z fazy lotnej określa się procesy polegające na wytworzeniu warstwy materiału z reagentów reagujących ze sobą w fazie lotnej [9]. Metody tej grupy mają szerokie zastosowanie, gdyż w ten sposób wytwarzane mogą być warstwy metaliczne, półprzewodnikowe, a także warstwy izolacyjne, tj. tlenek krzemu lub azotek krzemu. Zaletą tych procesów jest duża szybkość osadzania warstw. Niestety, metoda ta wymaga użycia względnie wysokich temperatur (ok. 300°C), co przekłada się na trudności tworzenia wielowarstw magnetycznych.

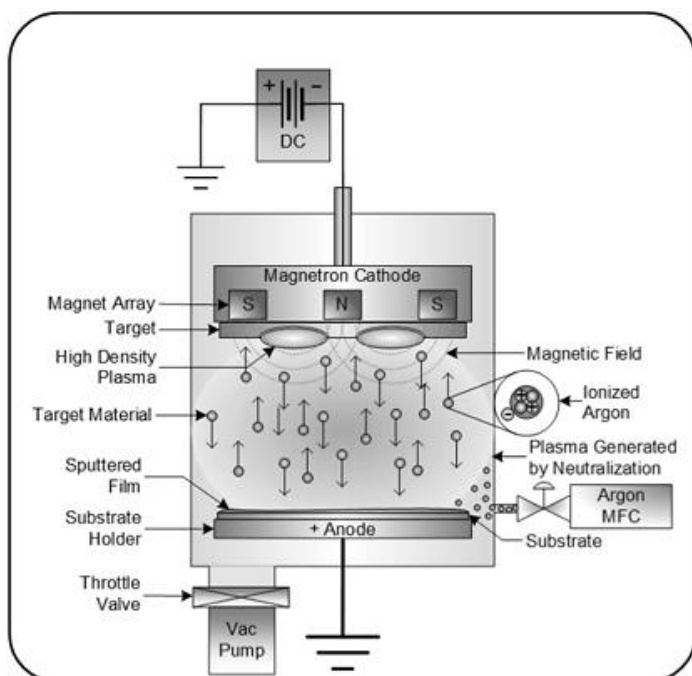
Fizyczne osadzanie z fazy gazowej to grupa zjawisk, w których nowe warstwy powstają z fazy gazowej bez udziału procesów chemicznych. W metodach tych wyróżnić można trzy główne procesy: (1) odseparowanie atomów ze źródła materiału (ang. *target*), (2) transport wybitych cząstek w kierunku podłoża oraz (3) kondensację wybitych cząstek na podłożu – co tworzy tym samym nowo powstałą warstwę. Wśród metod PVD wyróżnia się dwie główne grupy, różniące się przede wszystkim sposobem wybijania atomów materiału ze źródła, są to: naporowywanie (ang. *evaporation*) oraz rozpylanie (ang. *sputtering*) [2].

W przypadku naporowywania atomy z targetu doprowadzane są do stanu gazowego poprzez dostateczne ogrzanie materiału. Proces ten zachodzi w wysokiej próżni. Odparowywanie z wiązki molekularnej (ang. *molecular beam epitaxy*, MBE) jest jedną z metod naporowywania, dzięki której otrzymywane są warstwy monokrystaliczne wzrastające epitaksjalnie, a więc o kierunku krystalograficznym zgodnym z kierunkiem podłoża [14].

Rozpylanie, określane także napyłaniem, charakteryzuje się wykorzystaniem energii kinetycznej zewnętrznych cząstek gazów szlachetnych w celu wybicia atomów pożądanego materiału z targetu. W komorze napyłania między katodą a anodą atomy gazu szlachetnego jonizują się, a ich dodatni ładunek powoduje przyciąganie do katody, którą stanowi źródło materiału (lub umieszczona jest ona za targetem). Skutkuje to kontaktem rozprędzonych cząstek gazu szlachetnego z materiałem targetu. Cząsteczki żądanego materiału są wybijane i osadzają się w komorze napyłania, m.in. na przygotowanym podłożu znajdującym się w odpowiedniej odległości kilkudziesięciu centymetrów od targetu. Rozpylanie wykorzystujące gazy szlachetne określa się rozpyłaniem jonowym. Jest to grupa metod, którą najczęściej używa się w produkcji urządzeń elektroniki spinowej. Wśród rozpyłań jonowych wyróżnić można metodę stałoprądową (ang. *DC sputtering*) oraz zmiennoprądową (ang. *RF sputtering*) – ze względu na rodzaj zastosowanego napięcia między elektrodami w komorze próżniowej. Szczególnym rodzajem rozpyłania jonowego jest rozpylanie magnetronowe (ang. *magnetron sputtering*) [2, 14].

Rozpylanie magnetronowe charakteryzuje się zastosowaniem pod targetem dodatkowego pola magnetycznego. Pole magnetyczne przy źródle materiału

powoduje uwięzienie elektronów ponad powierzchnią targetu, doprowadzając do zwiększenia gęstości jonów. Przekłada się to na zwiększenie efektywności całego procesu rozpylania. Już za pomocą zasilania stałoprądowego magnetronu napyłać można materiały metaliczne, a więc zarówno warstwy ferromagnetyczne, jak i przedzielające je warstwy niemagnetyczne [16]. Komora napyłania magnetronowego przedstawiona została na rysunku 3.



Rysunek 3. Schemat komory rozpylania magnetronowego z zasilaniem stałoprądowym DC

Źródło: M. Hughes, *What is DC Sputtering?*, <http://www.semicon.com/news/94-what-is-dc-sputtering>, (dostęp: 16.10.2023).

Koło Naukowe SEMICON dysponuje systemem rozpylania magnetronowego NANO 36™ Kurt J. Lesker®. Studenci mają możliwość zapoznać się ze sposobem działania napyłarki oraz poszczególnymi procesami technologicznymi struktur cienkowarstwowych wykorzystywanymi również w przemyśle. W ramach prac koła wykonywane są struktury warstwowe m.in. z wykorzystaniem żelaza, permaloju, miedzi, tytanu i chromu. Struktury te domyślnie posłużą jako czujniki przetwarzające natężenie pola magnetycznego na sygnał elektryczny poprzez zmianę rezystancji elektrycznej mierzonej struktury.

### 3. Czujniki pola magnetycznego

Rozwój robotyzacji oraz internetu rzeczy wymaga także ciągłego rozwoju czujników różnych wielkości, w tym pola magnetycznego. Wykorzystanie dodatkowych magnesów lub elementów metalowych maszyn i konstrukcji rozszerza zastosowanie czujników pola magnetycznego jako czujników odległości, obrotów, natężenia prądu elektrycznego itp.

Obecnie najbardziej znanymi czujnikami do pomiaru pola magnetycznego są te bazujące na efekcie Halla, odkrytym w 1879 roku. Efekt obserwowany jest w materiałach przewodzących prąd elektryczny i polega na powstaniu różnicy potencjałów elektrycznych na dwóch przeciwległych krawędziach danego materiału pod wpływem działania zewnętrznego pola magnetycznego. Spowodowane jest to działaniem siły Lorentza na elektrony znajdujące się w polu magnetycznym, która wpływa na kumulację na jednej z krawędzi materiału. Jednak już w 1856 przez Lorda Kelvina zanotowane zostało zjawisko zmiany rezystancji w żelazie oraz w niklu, będącymi metalami ferromagnetycznymi, w obecności zewnętrznego pola magnetycznego. Zjawisko to zostało określone magnetorezystancją. Po pewnym czasie stały się one alternatywą dla czujników bazujących na efekcie Halla. Wykorzystywane są głównie zjawiska AMR, GMR i TMR. Ostatni z nich, tunelowy magnetoopór, odkryty został już w roku 1975, jednak nie wzbudził zainteresowania naukowców i dopiero od roku 1995 pojawiły się możliwości wykorzystania go na szerszą skalę.

Czujniki Halla charakteryzując się najmniejszą czułością i rozdzielczością, jednak nadrabiają to, oferując wysoki zakres dynamiczny przy niskich kosztach produkcji. Duża wydajność i dostępność przy niskich kosztach odpowiadają za popularnością tych czujników na rynku. Czujniki GMR osiągają lepsze parametry od czujników AMR. Czujniki TMR, mogące osiągać większe zmiany rezystancji od czujników GMR zdają się być droższą alternatywą [5]. Porównanie parametrów czujników Halla, AMR i GMR przedstawione zostało w tabeli 1.

Tabela 1. Porównanie parametrów czujników magnetycznych

	Czujniki GMR	Czujniki AMR	Czujniki Halla
<b>Wymiary</b>	małe	duże	małe
<b>Poziom sygnału</b>	wysoki	średni	niski
<b>Czułość</b>	duża	duża	niska
<b>Stabilność temperaturowa</b>	duża	średnia	niska
<b>Pobór mocy</b>	niski	wysoki	niski
<b>Cena</b>	niska	wysoka	niska

Źródło: D. Mitra, R. Mitra, *Development of Sensors Based on Giant Magnetoresistance Material*, „Procedia Engineering” 2012, vol. 32, pp. 60–68.



Czujniki magnetyczne działające na podstawie zjawiska GMR są atrakcyjne do stosowania we wszelkich bezkontaktowych rejestratorach pozycji, np. jako czujniki odległości, kąta, prędkości oraz prędkości obrotowej; w formach tych najczęściej znajdują zastosowanie w branży automotive [11]. Rezystory cienkowarstwowe GMR mogą być także stosowane jako czujniki wielu innych wielkości, tj. naprężenia mechaniczne, czujniki wibracyjne oraz czujniki do pomiaru różnicy ciśnień. Czujniki wykorzystujące zjawisko gigantycznego magnetooporu są bardzo szeroko stosowane w inżynierii medycznej. Znajdują zastosowanie zarówno w postaci układów mikroprzepływowych, jak i biosensorów np. w celu wykrywania protein lub bakterii [10]. Służą one również konstrukcji diagnostycznych sond igłowych oraz w skaningowej mikroskopii [7]. Pomiar prądów wirowych pozwalają na wykorzystanie struktur GMR w systemach diagnostyki nieniszczącej NDT (ang. *non-destructive testing*) [12]. Badania nieniszczące to zestaw metod stosowanych w celu oceny właściwości materiału. Powodem tego typu badań jest chęć uniknięcia potencjalnych awarii. Możliwe jest badanie m.in. defektów kadłubów konstrukcji lotniczych lub metalowych zbiorników służących przechowywaniu niebezpiecznych substancji [15]. Zmniejszanie wymiarów struktur GMR pozwala wykorzystywać je w formie matryc.

#### 4. Zewnętrzne pole magnetyczne

W badaniach struktur GMR niezbędne jest poddanie struktur działaniu zewnętrznego pola magnetycznego. W przypadku aplikacji nawigacyjnych, np. w kompasach, mierzone jest słabe pole magnetyczne ziemi. W czujnikach natężenia prądu mierzone jest natężenie pola magnetycznego powstałego w wyniku przepływu prądu przez przewód elektryczny. Zastosowania struktur GMR jako czujników położenia lub tym podobnych wymagają wykorzystania dodatkowego źródła pola magnetycznego jako odniesienia. W grę wchodzi magnesy trwałe oraz elektromagnesy. Można je również wykorzystać w badaniach rezystancyjnych struktur cienkowarstwowych.

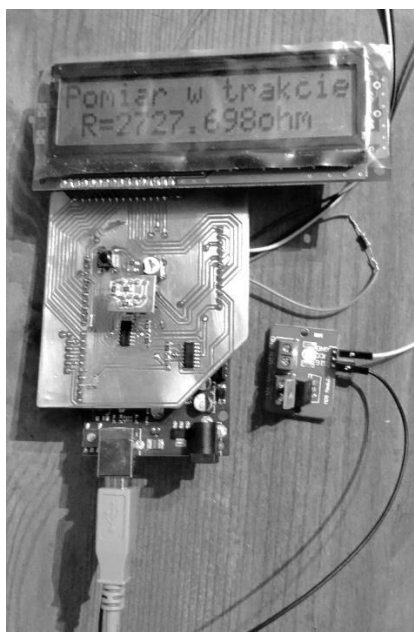
Magnesy neodymowe występują w różnych wielkościach, jednak magnesy wytwarzające silniejsze pole magnetyczne, rzędu kilku dziesiątych tesli, osiągają większe rozmiary. Elektromagnesy za to są elementami, które są źródłem pola magnetycznego w przypadku ich zasilenia. Najprostszy elektromagnes stanowi cewka zasilona stałym napięciem. W jej wnętrzu powstaje silniejsze pole magnetyczne, które najczęściej wzmacniane jest dodatkowym rdzeniem w postaci materiału ferromagnetycznego.

Szczególnym rodzajem elektromagnesu jest układ cewek Helmholtza. Składa się on z dwóch cewek, o promieniu znacznie większym od długości każdej z cewek. Pole magnetyczne powstaje w przestrzeni między cewkami. Zaletą tego układu jest uzyskanie zbliżonego do jednorodnego pola magnetycznego. Cewki takie wymagają jednak użycia dużej ilości przewodu, co przekłada się na duży koszt tego typu stanowiska.

## 5. Projekty koła naukowego SEMICON

W ramach prac prowadzonych w Studenckim Kole Naukowym Elektroników i Mechatroników SEMICON opracowywane są urządzenia umożliwiające potencjalne wykorzystanie wyprodukowanych struktur cienkowarstwowych, wykazujących zjawisko gigantycznego magnetooporu w roli czujników pola magnetycznego oraz stanowiska pomiarowe mogące posłużyć badaniom tychże struktur.

Urządzenie mające wykorzystywać wyprodukowaną strukturę jako czujnik natężenia pola magnetycznego składa się z modułu Arduino UNO oraz stworzonej płytki obwodu drukowanego (rys. 4). Urządzenie to dokonuje pomiaru rezystancji dzięki wykorzystaniu mostka Wheatstone'a. Na etapie rozwoju urządzenia jako zewnętrzne pole magnetyczne wykorzystywane były elektromagnesy, których wartość natężenia pola magnetycznego sterowana była wygładzonym sygnałem PWM na bramce tranzystora MOSFET. Wykorzystanie zewnętrznego teslomierza pozwala zmierzyć powstałe pole magnetyczne i zestawić je ze wskazaniami pomiaru rezystancji na wyświetlaczu urządzenia.



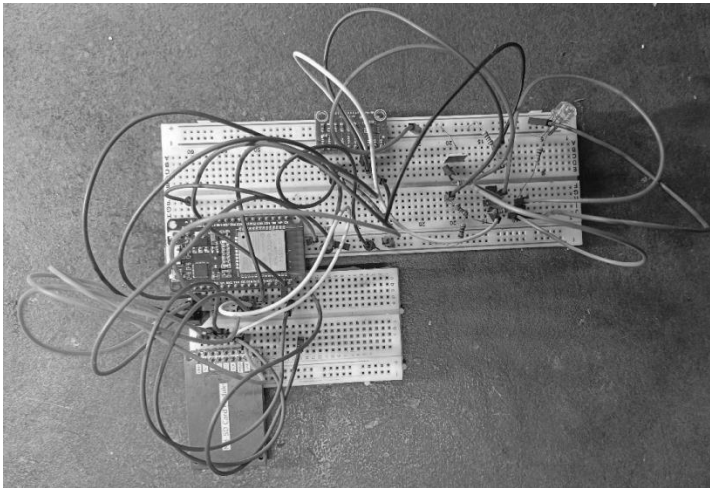
Rysunek 4. Urządzenie pomiarowe nr 1

Źródło: opracowanie własne.

Rozwój stworzonego urządzenia zakłada rozbudowanie modułu pomiarowego w celu zwiększenia możliwości regulacji zakresu pomiarowego, ze względu na

zastosowanie mostka Wheatstone'a. Możliwe jest także zastosowanie zewnętrznego przetwornika ADC, aby zwiększyć rozdzielczość dokonywanych pomiarów.

Drugi z projektów również zakłada pomiar rezystancji struktury cienkowarstwowej w celu określenia zmian rezystancyjnych w ramach zjawiska GMR w polu magnetycznym elektromagnesu (rys. 5). Wykorzystuje ono mikrokontroler ESP32 oraz zewnętrzny konwerter analogowo-cyfrowy ADS1115. Pomiar rezystancji wykonywany jest metodą techniczną z poprawnie mierzonym prądem.



**Rysunek 5. Urządzenie pomiarowe nr 2**

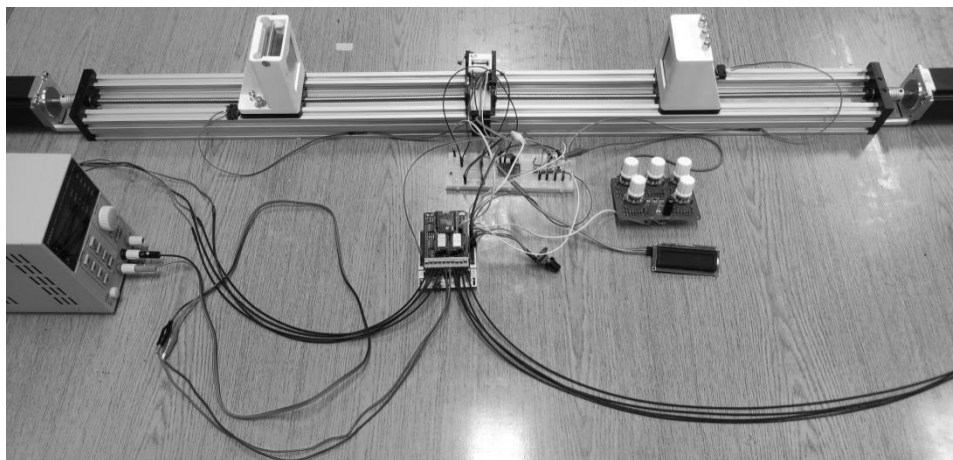
Źródło: opracowanie własne.

Urządzenie umożliwia przeprowadzenie pomiarów w dwóch trybach pracy: w pomiarze pojedynczym oraz pomiarze seryjnym, polegającym na wykonaniu serii kolejnych pomiarów struktury dla określonych programowo wartości indukcji pola magnetycznego wytwarzanego przez sterowany elektromagnes. W celu pomiaru indukcji pola wykorzystywany jest dodatkowy czujnik Halla.

Trzeci z układów stanowić ma stanowisko pomiarowe. Składa się ono z dwóch przewodnic liniowych, dwóch silników krokowych oraz dwóch silnych magnesów neodymowych (rys. 6). Stworzony układ, bazujący na mikroprocesorze STM32F411RE i wykorzystujący sterownik silników krokowych DFRobot DRI0023, odpowiada za sterowanie wzajemnym położeniem magnesów neodymowych poprzez dostosowanie odpowiedniej liczby kroków silników krokowych. Pomiar rezystancji również zrealizowany jest za pomocą mostka Wheatstone'a, w którym zastosowany został rezystor dekadowy.

Skonstruowane urządzenie pozwala mierzyć rezystancję struktury w zależności od pola magnetycznego między dwoma silnymi magnesami neodymowymi. Mierzona struktura znajduje się dokładnie pośrodku okładów magnesów neody-

mowych. Konstrukcja stanowiska umożliwia wykorzystanie go w połączeniu z zewnętrznym miernikiem i przeprowadzeniem półautomatycznego pomiaru dla zadanych wartości natężenia pola magnetycznego. Maksymalna osiągalna wartość indukcji pola magnetycznego wynosi ponad 0,3 T.



Rysunek 6. Stanowisko pomiarowe z magnesami neodymowymi

Źródło: opracowanie własne.

## 6. Podsumowanie

Działalność Koła Naukowego Elektroników i Mechatroników SEMICON, funkcjonującego na Wydziale Elektrotechniki i Informatyki Politechniki Lubelskiej, umożliwia chętnym studentom tak rozwój naukowy, jak i zdobycie umiejętności praktycznych w konstrukcji fizycznych projektów. Tego typu działalność jest możliwa dzięki uzyskanemu wsparciu finansowemu. W ramach programu *Studenckie koła naukowe tworzą innowacje*, organizowanego przez Ministerstwo Edukacji i Nauki, kołu naukowemu został przyznany grant o numerze SKN/SP/495923/2021. Tytuł grantu brzmiał „Opracowanie technologii czujników pola magnetycznego bazujących na rezystorach cienkowarstwowych”. Dzięki uzyskanym środkom w kwocie 70 tys. złotych możliwe było opracowanie metody wytwarzania i wykonanie prototypu czujnika indukcji pola magnetycznego na bazie struktury cienkowarstwowej. Dodatkowo, w ramach programu *Projekty studenckie dla kół naukowych Politechniki Lubelskiej* kołu naukowemu SEMICON przyznane zostało finansowanie w wysokości 8500 złotych dla projektu *Opracowanie urządzenia pomiarowego dla cienkowarstwowych rezystorów wykazujących efekt gigantycznego magnetooporu*. Środki te umożliwiły wykonanie trzech urządzeń pomiarowych wykorzystujących magnesy neodymowe oraz elektromagnesy.

Przynależący do koła studenci, którzy wykonują projekty w ramach grantów lub projekty własne, mają możliwość zaprezentowania swoich pomysłów i wyników badań na organizowanych sympozjach, warsztatach i konferencjach. Studenci mogą podzielić się swoimi ideami, zdobyć cenne wskazówki, spróbować swoich sił w wystąpieniach publicznych oraz nauczyć się opisać stworzone projekty w formie referatów naukowych, które mogą się później znaleźć w monografiach zbiorczych.

## Literatura

- [1] Baraduc C., Chshiev M., Diény B., *Spintronic Phenomena: Giant Magnetoresistance, Tunnel Magnetoresistance and Spin Transfer Torque*, “Smart Sensors, Measurement and Instrumentation” 2013, vol. 6, pp. 1–30.
- [2] Beck R., *Technologia krzemowa*, PWN, Warszawa, 1991.
- [3] Fermon C., *Micro- and Nanofabrication Techniques*, “Spin Electronics”, “Lecture Notes in Physics” 2000, vol. 569, pp. 379–395.
- [4] Inoue J., *GMR, TMR and BMR*, [in:] ed. T. Shinjo, *Nanomagnetism and Spintronics*, Elsevier Science, 2009, pp. 15–92.
- [5] Kan M. A., Sun J., Li B., Przybysz A., Kosel J., *Magnetic sensors-A review and recent technologies*, “Engineering Research Express” 2021, vol. 3, no. 2, pp. 2–22.
- [6] Leitao D. C., Amaral J. P., Cardoso S., Reig C., *Microfabrication Techniques*, “Smart Sensors, Measurement and Instrumentation” 2013, vol. 6, pp. 31–46.
- [7] Leitão D.C., Borme J., Orozco A., Cardoso S., Freitas P. P., *Magneto-resistive Sensors for Surface Scanning, Giant Magnetoresistance (GMR) Sensors: From Basis to State-of-the-Art Applications*, “Smart Sensors, Measurement and Instrumentation” 2013, vol. 6, pp. 275–299.
- [8] Mitra D., Mitra R., *Development of Sensors Based on Giant Magnetoresistance Material*, “Procedia Engineering” 2012, vol. 32, pp. 60–68.
- [9] Pierson, H. O., *Handbook of Chemical Vapor Deposition (CVD) – Principles, Technology and Applications*, William Andrew Publishing Notes, 1999.
- [10] Ren C. et al., *Biomarkers detection with magnetoresistance-based sensors*, “Biosens. Bioelectron” 2020, vol. 165, pp. 1–8. DOI: 10.1016/j.bios.2020.112340.
- [11] Rieger G., Ludwig K., Hauch J., Clemens W., *GMR sensors for contactless position detection*, “Sensors and Actuators A: Physical” 2001, vol. 91, iss. 1–2, pp. 7–11.
- [12] Romero-Arismendi N. O., Pérez-Benítez J. A., Ramírez-Pacheco E., Espina-Hernández J. H., *Design method for a GMR-based eddy current sensor with optimal sensitivity*, “Sensors and Actuators A: Physical” 2020, vol. 314, pp. 1–9.

- [13] Sharko S. A., Serokurova A. I., Zubar T. I., Trukhanov S. V., Tishkevich D. I., Samokhvalov A. A., Kozlovskiy A. L., Zdorovets M. V., Panina L. V., Fedosyuk V. M., Trukhanov A. V., *Multilayer spin-valve CoFeP/Cu nanowires with giant magnetoresistance*, "Journal of Alloy Compounds" 2020, vol. 846.
- [14] Stobiecki T., *Urządzenia elektroniki spinowej*, Wydawnictwo AGH, 2012, pp. 30–59.
- [15] Vacher F., Gilles-Pascaud C., Decitre J.M., Fermon C., Pannetier M., Cattiaux G., *Non destructive testing with GMR magnetic sensor arrays*, [in:] *9th European Conference on NDT*, Berlin, 2006.
- [16] Waite M. M., Chester W., Glocker D., *Sputtering Sources*, [in:] eds. D. M. Mattox, V. Harwood Mattox, *Society of Vacuum Coaters*, 2010, pp. 42–50.
- [17] Hughes Matt, *What is DC Sputtering?*, 26.11.2016, <http://www.semicore.com/news/94-what-is-dc-sputtering>, (dostęp: 05.11.2022).

## Technology and studies of thin-film structures exhibiting the phenomenon of giant magnetoresistance

### Abstract

The work is about production and studies of thin-film structures that exhibit the phenomenon of giant magnetoresistance. As part of the work of the Scientific Circle of Electronics and Mechatronics SEMICON, operating at the Lublin University of Technology, layered structures are produced in which the value of the DC resistance depends on the intensity of the external magnetic field. For this purpose, magnetron sputtering technology is used. In order to be able to observe the phenomenon of giant magnetoresistance, it is important to properly design the structure, since the thicknesses of individual layers reach the order of several nanometers. Structures of this type have potential applications as magnetic field sensors. As part of the circle's activities, devices are also being designed to measure the resistance of the structures under study or to serve as a test station making it possible to obtain a constant magnetic field varying in value by using neodymium magnets or electromagnets.

*Keywords: magnetron sputtering, static magnetic field, thin-film layers, magnetoresistance.*



**Katarzyna Korulczyk<sup>1</sup>, Kamil Kapłon<sup>2</sup>, Magdalena Paśnikowska-Łukaszuk<sup>3</sup>, Aleksandra Prus<sup>4</sup>**

## **Przegląd ekopolimerów dostępnych w polskich sklepach w latach 2018–2022**

### **Streszczenie**

Z roku na rok rośnie światowa produkcja tworzyw polimerowych ropopochodnych. W 2021 roku było to 390,7 mln ton. Co pocieszające, rośnie też produkcja polimerów z tworzyw pochodzących z recyklingu (w 2021 roku – 32,5 mln ton), a także pochodzenia biologicznego (w 2021 roku – 5,9 mln ton). Naturalną konsekwencją tego faktu było wkroczenie na sklepowe półki produktów wykonanych z tworzyw pochodzących z recyklingu (np. ściereczka z butelek wyłowionych z oceanu) czy z polimerów biodegradowalnych (np. słomki jedno-razowe z polilaktydu). Niniejsza praca jest próbą przedstawienia przekroju ofertowanych ekopolimerów dostępnych w sklepach stacjonarnych w Polsce w latach 2018–2022.

*Słowa kluczowe: tworzywa polimerowe, naturalne napelniacze, recykling polimerów, polimery biodegradowalne.*

### **Wstęp**

Wraz ze wzrostem konsumpcji produktów wzrasta również ilość produkowanych opakowań i przedmiotów. Od dłuższego czasu szuka się zamienników, które wpłyną na zmniejszenie produkcji tworzyw polimerowych w ciągu roku. Obecnie organizacje rządowe wprowadzają różnego rodzaju ograniczenia mające na celu zredukowanie wytwarzanych opakowań. Szacuje się, że rocznie wyrabia się około 380 milionów ton tworzyw polimerowych. Niepokojące jest to, że bardzo dużo odpadów ostatecznie trafia do oceanów i wód przybrzeżnych. Tworzywa polimerowe są zagrożeniem dla środowiska i dla człowieka, bo ostatecznie mogą trafiać do żywności i wody w postaci mikroplastików. W celu ochrony zarówno środowiska, jak i ludzkiego zdrowia od dłuższego czasu na rynku można zauważyć

---

<sup>1</sup> Mgr inż. Katarzyna Korulczyk, Wydział Podstaw Techniki, Politechnika Lubelska, katarzyna.korulczyk@pollub.edu.pl.

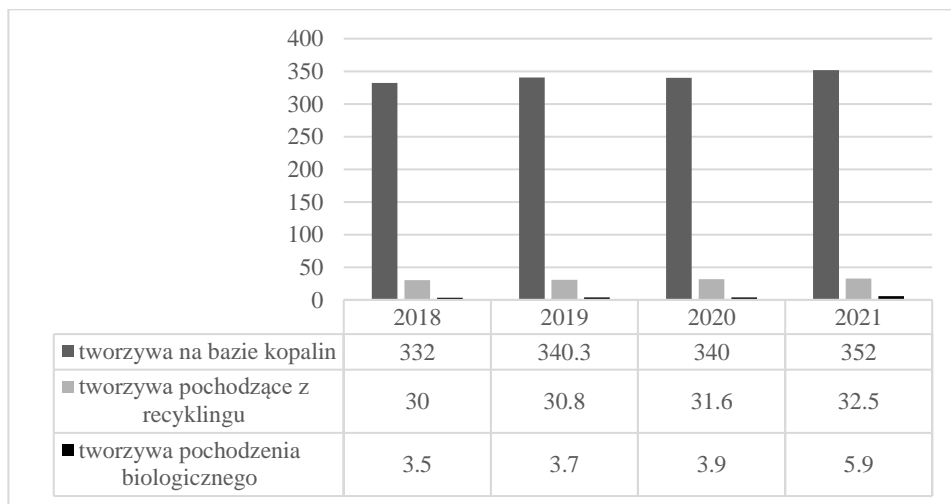
<sup>2</sup> Mgr inż. Kamil Kapłon, Wydział Podstaw Techniki, Politechnika Lubelska, kamil.kaplon@pollub.edu.pl.

<sup>3</sup> Mgr inż. Magdalena Paśnikowska-Łukaszuk, Katedra Podstaw Techniki, Wydział Podstaw Techniki, Politechnika Lubelska, m.pasnikowska-lukaszuk@pollub.pl.

<sup>4</sup> Mgr inż. Aleksandra Prus, Katedra Podstaw Techniki, Wydział Podstaw Techniki, Politechnika Lubelska, a.prus@pollub.pl.



obecność biopolimerów, tworzyw biodegradowalnych oraz zamienianie opakowań z tworzyw polimerowych na papierowe bądź pochodzące całkowicie z recyklingu. Z roku na rok rośnie produkcja tworzyw polimerowych [1]. Wykres poniżej (rys. 1) przedstawia pochodzenie tworzyw polimerowych produkowanych w latach 2018–2021.



**Rysunek 1. Rodzaje pochodzenia tworzyw polimerowych produkowanych w latach 2018–2021 w mln t**

Źródło: opracowanie własne na podstawie *Plastics Europe, Plastics – the Facts 2022*, <https://plasticseurope.org/knowledge-hub/plastics-the-facts-2022/>, dostęp 30.11.2022.

Analizując powyższy wykres można zauważyć ciągły wzrost produkcji tworzyw polimerowych na bazie kopalin. Równocześnie z ich wzrostem rośnie również wyrób tworzyw pochodzących z recyklingu oraz tworzyw pochodzenia biologicznego. Tak duża liczba tworzyw stwarza problem z ich dalszym przetwarzaniem. Aby ograniczyć produkcję zaczęto wprowadzać pewne rozwiązania, jak np. recykling, wytwarzanie tworzyw pochodzenia biologicznego, używanie napelnaczy (zwłaszcza pochodzenia naturalnego) czy używanie dodatków przetwórczych mających wpływ na szybsze uleganie rozpadowi tworzyw [2–4]. Recykling ma na celu zebranie oraz przetworzenie tworzyw już wytworzonych. Zebrane odpady czyści się, rozdrabnia, przygotowuje się z nich regranulat, a następnie używa się go do produkcji czy to w 100% objętości, czy to mieszając z tworzywami wytworzonymi bezpośrednio z kopalin [5]. Tworzywa polimerowe pochodzenia biologicznego uzyskuje się poprzez selektywne przekształcenia roślin i innych niekopalnych surowców biogenych. Dzięki takiemu działaniu powstają tworzywa podobne do tych, które powstają na bazie kopalin [6]. Napelnaczy używa się w celu zmniejszenia użycia tworzywa polimerowego do wypro-

dukowania danego przedmiotu. Stosuje się odpowiedni procent wagowy, dostosowany do właściwości, jakie są pożądane. Stosowane szeroko napełniacze naturalne są używane przede wszystkim ze względu na swoją łatwą dostępność, ale także przez biodegradowalność [7].

## **1. Naturalne i syntetyczne tworzywa polimerowe**

W 2021 roku łącznie wyprodukowano 390,7 mln ton tworzyw polimerowych. Z czego 19,3% stanowił polipropylen (PP), 14,4% – polietylen małej gęstości (PE-LD) i liniowy polietylen małej gęstości (PE-LLD), 12,9% polichlorek winylu (PVC), 12,5% polietylen dużej i średniej gęstości (PE-HD, PE-MD). Wciąż stosunkowo niski odsetek (1,5%) stanowią tworzywa pochodzenia biologicznego [1].

Tworzywem o największym udziale wagowym jest polipropylen. Materiał ten należy do grupy tworzyw termoplastycznych. Otrzymywany jest w wyniku procesu stereospecyficznego polimeryzacji propenu. Polipropylen jest tworzywem twardym oraz odpornym na zarysowania. Ponadto charakteryzuje się bardzo dobrą odpornością chemiczną. Nie jest jednak odporny na działanie ujemnych temperatur. Negatywny wpływ na wytrzymałość tego polimeru ma również promieniowanie ultrafioletowe.

Często wykorzystywanym polimerem jest polietylen. Szeroko wykorzystywane są dwie jego odmiany różniące się gęstością. Pierwszą z nich jest polietylen niskociśnieniowy HD PE, który charakteryzuje się wysoką gęstością, wytrzymałością mechaniczną oraz twardością. Drugi z nich – polietylen wysokociśnieniowy LD PE ma mniejszą gęstość, twardość oraz niższą wytrzymałość mechaniczną. Zaletą polietylenu o małej gęstości jest jego elastyczność.

Polichlorek winylu jest tworzywem otrzymywanym w wyniku polimerizacji chlorku winylu. Cechuje się wysoką twardością, odpornością na chemikalia oraz wytrzymałością. Jest materiałem niepalnym, łatwo obrabialnym mechanicznie.

Polilaktyd (PLA) – jest tworzywem termoplastycznym, charakteryzującym się biodegradowalnością, która może wynosić od kilka miesięcy do kilku lat. Wytwarzany jest z surowców odnawialnych, takich jak kukurydza lub trzcina cukrowa. Właściwości mechaniczne polilaktydu zbliżone są do polipropylenu oraz polietylenu PLA.

## **2. Przegląd wybranych artykułów**

Odwiedzając w ostatnich latach sklepy coraz częściej zauważa się produkty, które mają być w jakiś sposób przyjazne dla środowiska. Obecnie w wielu supermarketach i w drogeriach pojawiają się produkty w opakowaniach z oznaczeniem biodegradowalności. Ponadto, coraz więcej produktów jednorazowych, które najczęściej są wykonane z tworzyw polimerowych, zaczyna być zamieniane na tworzywa z domieszkami np. drewna lub resztek drewnopochodnych. Tabela 1

pokazuje, jak dużo zamienników przyjaznych środowisku wykorzystuje się w produktach bądź ich opakowaniach na co dzień.

Zbiór artykułów trwał od 2018 roku. Pierwszy artykuł (noże jednorazowe z polilaktydu napełniane otrębami) kupiono w 2018 roku. W późniejszym czasie można było zauważyć systematyczne poszerzanie się asortymentu sklepów o produkty, które można byłoby nazwać ekopolimerami. W ciągu czterech lat na potrzeby naukowe zakupiono 33 przedmioty. Poniżej przedstawiono ich zestawienie (tab. 1) wraz z opisem materiału, z jakiego zostały wykonane; ich cechami charakterystycznymi, które producent uznał za ważne dla kupującego i w widoczny sposób zaznaczył je na opakowaniu; a także krajem, w którym dany artykuł został wyprodukowany. Zostały zamieszczone także zdjęcia przedstawiające poszczególne przedmioty (rys. 2–34).

**Tabela 1. Spis zebranych ekopolimerów**

Lp.	Typ	Marka	Materiał	Cecha	Produkcja
1	Ściereczka	Econatural by york	Poliester/ poliamid	Z butelek wylowionych z oceanu	Chiny
2	Ściereczka	Jan Niezbędny	Poliester/ poliamid	Wtórny PET	Chiny
3	Noże jednorazowe	Amigo	C-PLA	Kompostowalne	Polska
4	Szczotka do włosów	KillyS	70% ziarna kawy	Biodegradowalna	Brak danych
5	Długopis	Be happy	40% słoma, 60% plastik		Chiny
6	Nożyczki	eloy	Stal i tworzywo z włókien pszennych		Chiny
7	Widelce jednorazowe	Pepco – party collection	CPLA	Biogedradowalne/ kompostowalne	Polska
8	Ściereczka	Paclan		Z butelek PET	Chiny
9	Słomki jednorazowe	GoDan	PLA	Biodegradowalne	Polska
10	Słomki jednorazowe	Grill&Fun – Lidl	PLA	Biodegradowalne	Brak danych
11	Słomki jednorazowe	EKO GRILL	PLA	Biodegradowalne	UE
12	Sztućce jednorazowe	Anna Zaradna	Mączka kukurydziana – PLA	BIO – produkt naturalny	Brak danych
13	Słomki jednorazowe	EKO GRILL	PLA	Biodegradowalne	UE
14	Kubki jednorazowe	EKO GRILL	PLA	Biodegradowalne	Polska

Lp.	Typ	Marka	Materiał	Cecha	Produkcja
15	Noże jednorazowe	Grill & Fun – Lidl	PLA z dodatkiem otrębów	Biodegradowalne	Brak danych
16	Worki na śmieci	Purio – Lidl	Folia Coex	100% z recyklingu	Brak danych
17	Doniczka	Lamela	PP z napelniaczem z drewna	Do 40% drewna	Polska
18	Kubek	Practic		35% konopi	Polska
19	Spinki do włosów	Elite		Słoma pszeniczna i kukurydza	Chiny
20	Spinka do włosów	Jean Louis David	Fusy kawy i PLA	Naturalne surowce, biodegradowalne	Chiny
21	Cienkopisy	Ziemia obiecana jami		Wyprodukowane z recyklingu, z materiałów postkonsumpcyjnych, korpusy wykonane ze słomy	Chiny
22	Szczoteczka do zębów	Colgate		Uchwyt wykonany z plastiku pochodzącego w 100% z recyklingu, włosie w 100% z surowców roślinnych	Chiny
23	Case na telefon	LILER ECO	ECO PLA	Biodegradowalne	Polska
24	Szczotka do modelowania włosów	Ewa Schmitt	PLA, nylon, aluminium	Biodegradowalne	Chiny
25	Szczotka do masażu	for your Beauty – Rossman	Rączka – 50% materiał z recyklingu; 50% słoma pszeniczna włosie zwykły polimer	Rączka w 50% wyprodukowana z recyklingu	Chiny
26-27	Pojemnik na żywność	Curver	PP z recyklingu	100% polipropylenu z recyklingu	Węgry
28	Maszynki do golenia	Wilkinson Sword	PP	Rączka wykonana z recyklingu	Polska
29	Półmisek	ECO-LINE	PP+mączka drzewna	Min. 30% drewna	Brak danych
30	Organizer	Curver	PP	Wykonany w 100% z recyklingu	Węgry
31	Grzebień do włosów	Ewa Schmitt	PLA	Biodegradowalne	Brak danych

Lp.	Typ	Marka	Materiał	Cecha	Produkcja
32	Miska	ECO-LINE	PP+mączka drzewna	Min. 30% drewna	Polska
33	Szczotka do włosów	Biofriendly	Materiały uzyskane z kukurydzy i manioku	Biodegradowalna w 5 lat	Brak danych

C-PLA, CPLA – polilaktyd skrzystalizowany, PLA – polilaktyd, PP – polipropylen.

Źródło: opracowanie własne.



Rysunek 2. Ściereczka Econatural by york (poz. 1. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 3. Ściereczka Jan Niezbędny (poz. 2. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 4. Noże jednorazowe Amigo (poz. 3. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 5. Szczotka do włosów KillyS (poz. 4. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 6. Długopis Be happy (poz. 5. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 7. Nożyczki eloy (poz. 6. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 8. Widelce jednorazowe Pepco – party collection (poz. 7. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 9. Ściereczka Paclan (poz. 8. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 10. Słomki jednorazowe GoDan (poz. 9. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 11. Słomki jednorazowe Grill & Fun (poz. 10. Tab. 1)

Źródło: fot. K. Korulczyk.



**Rysunek 12. Słomki jednorazowe EKO GRILL (poz. 11. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 13. Sztuczce jednorazowe Anna Zaradna (poz. 12. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 14. Słomki jednorazowe EKO GRILL (poz. 13. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 15. Kubki jednorazowe EKO GRILL (poz. 14. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 16. Noże jednorazowe Grill & Fun (poz. 15. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 17. Worki na śmieci purio (poz. 16. Tab. 1)**

Źródło: fot. K. Korulczyk.



Rysunek 18. Doniczka Lamela (poz. 17. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 19. Kubek practic (poz. 18. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 20. Spinki do włosów elite (poz. 19. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 21. Spinka do włosów Jean Louis David (poz. 20. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 22. Cienkopisy ziemia obiecana jami (poz. 21. Tab. 1)

Źródło: fot. K. Korulczyk.



Rysunek 23. Szczoteczka do zębów Colgate (poz. 22. Tab. 1)

Źródło: fot. K. Korulczyk.





**Rysunek 24. Case na telefon LILER ECO (poz. 23. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 25. Szczotka do modelowania włosów Ewa Schmitt (poz. 24. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 26. Szczotka do masażu for your Beauty (poz. 25. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 27. Pojemnik na żywność Curver (poz. 26. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 28. Pojemnik na żywność Curver (poz. 27. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 29. Maszynki do golenia Wilkinson Sword (poz. 28. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 30. Półmisek ECO-LINE  
(poz. 29. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 31. Organizier Curver  
(poz. 30. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 32. Grzebień do włosów Ewa  
Schmitt(poz. 31. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 33. Miska ECO-LINE  
(poz. 32. Tab. 1)**

Źródło: fot. K. Korulczyk.



**Rysunek 34. Szczotka do włosów Biofriendly  
(poz. 33 . Tab. 1)**

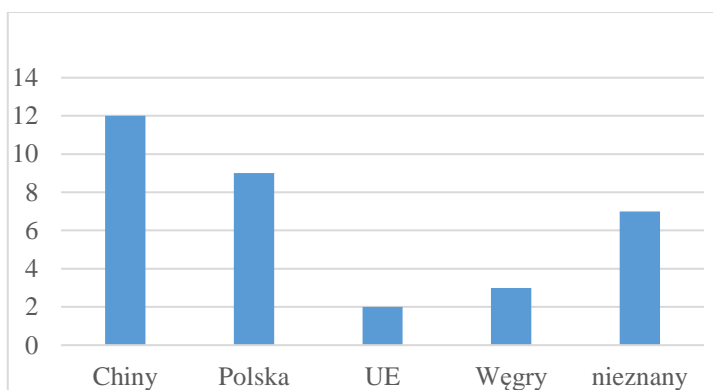
Źródło: fot. K. Korulczyk.

Przedstawione na rysunkach 2–34 artykuły były kupowane przede wszystkim w ogólnodostępnych sieciach sklepów.

### 3. Analiza dostępnych artykułów

Zebrane artykuły można podzielić na trzy główne grupy, tzn. artykuły wytworzone z polimerów biodegradowalnych, wytworzone z naturalnymi napełniaczami oraz wytworzone z materiałów pochodzących z recyklingu. Przeważają produkty z polilaktydu – zarówno bez dodatków, jak i z różnego rodzaju napełniaczami. Zastosowane napełniacze są wyłącznie naturalne, a w większości są odpadami, jak na przykład fusy kawy, słoma czy otręby. Tworzywa pochodzące z recyklingu to w głównej mierze PET (poli(tereftalan etylen)) oraz PP (polipropylen). Producenci troskę o środowisko pragną podkreślić przez wspomnienie o szczególnym pochodzeniu tworzyw z recyklingu, np. butelek wyłowionych z oceanu.

Na wykresie z rysunku 35 przedstawiono liczbę artykułów wyprodukowanych w poszczególnych państwach.



**Rysunek 35. Liczba artykułów wyprodukowanych w poszczególnych państwach**

Źródło: opracowanie własne.

Jak można zauważyć na wykresie, najwięcej przedmiotów zostało wyprodukowanych w Chinach (12), dalej jest Polska z 9 artykułami. 7 producentów nie podaje pochodzenia swoich wyrobów.

Należy zauważyć, że wzrost produkcji biodegradowalnych materiałów wynika przede wszystkim z ograniczeń wprowadzonych m.in. przez Unię Europejską. Od 3 lipca 2021 r. jednorazowe talerze, sztućce, słomki, patyczki do balonów i patyczki higieniczne wykonane z plastiku nie mogą być wprowadzane na rynki państw członkowskich UE. Zasada ta obejmuje m.in. kubki, pojemniki na żywność i napoje wykonane ze styropianu, a także wszystkie produkty z tworzyw sztucz-

nych ulegających oksydegradacji. Nowe przepisy zakazują stosowania wybranych artykułów plastikowych jednorazowego użytku, które można zastąpić rozwiązaniami alternatywnymi np. biodregadowalnymi. Wprowadza się również specjalne środki w celu ograniczenia stosowania niektórych produktów [8].

#### 4. Podsumowanie

Ekopolimery stają się coraz bardziej dostępne dla konsumentów. Wynika to rosnącej troski o środowisko. Przedstawione produkty są najczęściej sprzedawane w dużych sieciach sklepów. Oferowane przedmioty wykonywane są zarówno z polimerów biodegradowalnych, jak i pochodzących z recyklingu, do ich produkcji używa się także naturalnych napelniczy. Dzięki pojawieniu się zamienników tych popularnych tworzyw polimerowych ropopochodnych, kraje Unii Europejskiej mogą dostosowywać się do nowych przepisów i dyrektyw. Biodegradowalne tworzywa możemy również spotkać w technologii druku 3D, gdzie najczęściej używa się polilaktydu PLA, który jest poliestrem alifatycznym.

#### Literatura

- [1] *Plastics Europe, Plastics – the Facts 2022*, <https://plasticseurope.org/knowledge-hub/plastics-the-facts-2022/> (dostęp 30.11.2022).
- [2] Wróbel G., Bortel K., Chmielnicki B., *Wybrane właściwości mieszanin recyklatów PEHD/PLA/oksydegradowalny PE*, „Przetwórstwo Tworzyw” 2013, nr 4, ss. 426–431.
- [3] Marcinkowska A., Rozmysłowicz K., Andrzejewska E., Dach J., Pilarski K., *Wpływ wstępnej degradacji środowiskowej oraz promieniowaniem UV/VIS na właściwości mechaniczne komercyjnych folii oksybiodegradowalnych*, „Rocznik Ochrona Środowiska” 2011, t. 13, ss. 1605–1618.
- [4] Gibas E., Rymarz G., *Termoplastyczne polimery oksybiodegradowalne*, „Przetwórstwo Tworzyw” 2009, nr 2, ss. 24–33.
- [5] Smith R. L., Takkellapati S., Riegerix R. C., *Recycling of Plastics in the United States: Plastic Material Flows and Polyethylene Terephthalate (PET) Recycling Processes*, “ACS Sustainable Chemistry & Engineering” 2022, vol. 10 (6), pp. 2084–2096.
- [6] Cywar R. M., Rorrer N. A., Hoyt C. B. et al., *Bio-based polymers with performance-advantaged properties*, “Nature Reviews Materials” 2022, vol. 7, pp. 83–103.
- [7] Jagadeesh P., Puttegowda M., Gowda Y., Girijappa T., Rangappa S. M., Siengchin S., *Effect of natural filler materials on fiber reinforced hybrid polymer composites: An Overview*, “Journal of Natural Fibers” 2020, vol. 19, pp. 4132–4147.

- [8] Oficjalna strona internetowa Unii Europejskiej, *Ograniczenia UE dotyczące niektórych tworzyw sztucznych jednorazowego użytku*, [https://environment.ec.europa.eu/topics/plastics/single-use-plastics\\_pl](https://environment.ec.europa.eu/topics/plastics/single-use-plastics_pl) (dostęp: 31.12.2022).

## **Overview of ecopolymers available in Polish stores in 2018–2022**

### **Abstract**

The world production of petroleum-derived polymers is growing year by year. In 2021, it was 390.7 million tonnes. What is comforting, the production of polymers from recycled plastics (in 2021 – 32.5 million t) and of biological origin (in 2021 – 5.9 million t) is also growing. A natural consequence of this fact was the entry of products made of recycled plastics (e.g. cloth from bottles fished from the ocean) or biodegradable polymers (e.g. disposable straws made of polylactide) onto store shelves. This work is an attempt to present a cross-section of the offered ecopolymers available in stationary stores in Poland in the years 2018–2022.

*Keywords: polymeric materials, natural fillers, polymer recycling, biodegradable polymers.*

## **Naprawa czy na śmietnik?**

### **Streszczenie**

W obecnych czasach temat naprawy komputerów jest rozpatrywany pod wieloma aspektami, ze względu na szybki rozwój technologii i brak części zamiennych. Z dnia na dzień opracowywane są nowe technologie, stosowane w coraz to nowszych komponentach komputerowych, co sprawia, że naprawa starych jest nieopłacalna, ze względu na cenę. Biorąc pod uwagę wydarzenia z ostatnich lat, możemy stwierdzić, że inflacja dotknęła przedsiębiorców, którzy – oczywiście w zależności od usterki – nie kupują przeważnie nowego osprzętowania, tylko dokonują renowacji starych. Poprzez renowację nie tylko oszczędzamy na komponentach informatycznych, ale i zmniejszamy ilość elektrośmieci.

*Słowa kluczowe: elektrośmieci, technologie informatyczne, komponenty informatyczne.*

### **Wprowadzenie**

W XXI wieku żyjemy w przeświadczeniu, że wszystko co nas otacza, zwłaszcza sprzęt elektroniczny, informatyczny i mechaniczny, z każdym dniem się starzeje. Firmy zajmujące się produkcją takowych sprzętów codziennie opracowują kolejne urządzenia, które zastąpią nasze obecne (rys. 1). W momencie uszkodzenia, czy typowo technicznych problemów przy korzystaniu z danego osprzętowania, większość ludzi decyduje się raczej na kupno urządzeń zastępczych aniżeli na samą naprawę.

### **1. Wykorzystanie kart graficznych**

Kwestia naprawy jest oczywiście uzależniona od wielu czynników, w tym od rodzaju wyposażenia użytkowego, rodzaju samej usterki oraz kosztów, jakie ponosi tak konsument, jak i serwis. W obecnych czasach problemem są przede wszystkim braki części. Na rynku komputerowym jest ogromny deficyt kart graficznych oraz procesorów. Przykładowo, gdy Nvidia wypuściła kolejną generację kart, RTX serii 40, to rynek zaczął się bardzo gwałtownie zmieniać, przez co dla wielu użytkowników towaru zabrakło. Karty graficzne RTX serii 20, które miały premierę 20 września 2020 roku, obecnie są uważane za przestarzałe i do niczego się już nienadające, pomimo że do grania w gry komputerowe,

---

<sup>1</sup> Marek Nowosad, Studenckie Koło Naukowe Zastosowań Informatyki w Nauce i Biznesie „InfoNaBi”, Wydział Podstaw Techniki, Politechnika Lubelska.

wystarczają w zupełności. Zwłaszcza zważywszy na to, że żadna gra komputerowa dostępna na rynku nie zużywa tyle pamięci graficznej, żeby chociaż w połowie wykorzystać moc obliczeniową najtańszej karty z owej serii 20 [7].



**Rysunek 1. Anteny satelitarne pochodzące z demontażu**

Źródło: fot. M. Nowosad.

Z tego powodu bardzo duża liczba użytkowników wystawia stare karty graficzne na aukcje, na których skupują je liczni przedsiębiorcy oraz firmy zajmujące się pracą biurową. Wyposażają w nie komputery swoich pracowników, aby te miały większą moc obliczeniową. Zapewniają w ten sposób większy komfort pracownikom podczas pracy. Z czasem, kiedy pracodawcy decydują się na udoskonalenie komputerów, przekazują stary sprzęt informatyczny do odpowiednich firm, które zajmują się odzyskiem. Skutkiem takiego działania jest ponowne wykorzystanie rdzeni krzemowych, będących podstawą działania procesorów graficznych, do produkcji coraz to nowszych kart graficznych [7].

## **2. Koszt naprawy a koszt wymiany**

Z reguły człowiek skłania się do wydania jak najmniejszej sumy pieniędzy, aby posiadać sprzęt komputerowy spełniający jego oczekiwania. Dzieje się tak

nawet wtedy, gdy sama kwota zakupu nie jest wygórowana. Sprowadza się to do odmowy naprawy urządzeń informatycznych, gdy cenowo wychodzi to dużo drożej aniżeli kupno nowego, nieużywanego sprzętu. To zjawisko występuje nie tylko w Polsce, ale i na całym świecie. Jest sprzeczne z filozofią *zero waste* [9].

Wielu przedsiębiorców zajmujących się sprzedażą komputerów, laptopów oraz innych urządzeń techniki komputerowej sprowadza towar najczęściej z krajów skandynawskich, gdzie owe nastawienie jest jeszcze bardziej rozpowszechnione niż w Polsce. Gdyby zadać pytanie “Skąd państwo macie tak duże ilości komputerów/laptopów?”, to co drugi przedsiębiorca śmiało odpowie, iż z importu właśnie ze Skandynawii, gdzie większość komputerów leżących wśród śmieci to najczęściej urzędzenia wciąż zdadne do użytku, stosunkowo nowe [2].

Procesory znajdujące się w takich komputerach bardzo często okazują się być z ósmej lub nawet dziewiątej generacji, gdzie najnowszą jest generacja trzynasta, kosztująca obecnie na rynku od dwóch do pięciu tysięcy złotych, w zależności od modelu. Sam proces wytwarzania procesora jest długotrwały i bardzo kosztowny. Do produkcji procesorów wykorzystuje się wiele metali szlachetnych, w tym złota. W trakcie produkcji, przy użyciu bardzo dokładnych lutownic, lutowane są bardzo cienkie wiązki z owego metalu szlachetnego. Mają one za zadanie przepuszczać bardzo wysokie wartości napięć elektrycznych tak, aby wiązka elektronowa nie uszkodziła przy tym pozostałych mikroelementów procesora. Ta technologia stosowana jest przede wszystkim w nowszych procesorach, zapewniając przy tym bardzo dużą moc obliczeniową. Jednoznacznie ma to wpływ na wysokość ceny takich procesorów. Oprócz procesorów i wcześniej wspomnianych kart graficznych, do prawidłowej pracy komputera niezbędne są również kości pamięci RAM. Jest to pamięć krótkotrwała, w której przechowywane są najważniejsze dane przeanalizowane przez procesor [1].

Zważywszy na szeroki dostęp do surowców, z których wytwarzane są kości pamięci, stanowią one jeden z najtańszych elementów budowy komputera. Należy również wziąć pod uwagę tzw. szkielet komputera, czyli płytę główną (rys. 2), do której wpinane są wszystkie inne elementy. Tutaj, podobnie jak przy procesorach, do produkcji wykorzystywane są metale szlachetne, służące do przeprowadzania wysokich wartości napięć elektrycznych między poszczególnymi obwodami elektrycznymi. Płyty główne służą przede wszystkim do komunikacji między poszczególnymi komponentami oraz układami elektrycznymi. W przypadku uszkodzenia, konsument nie ma alternatywy i zmuszony jest do wymiany tego modułu. W przypadkach wielu uszkodzeń, konieczność nie kończy się na wymianie samej płyty. Bardzo często przez wysokie napięcia, uszkodzeniu podlega wiele innych segmentów układu elektrycznego, przez co koszty naprawy wzrastają [6].





**Rysunek 2. Płyty główne pochodzące z demontażu**

Źródło: fot. M. Nowosad.

### 3. Wpływ sytuacji geopolitycznej

Agresja ze strony narodu rosyjskiego wobec Ukrainy to tragedia, która dotknęła cały świat. Skala zniszczeń, do jakich dochodzi w dalszym ciągu, przeobraziła się w problemy zarówno ekonomiczne, jak i gospodarcze. Mimo licznych niedogodności, ukraińska branża IT dalej funkcjonuje i nie poddaje się [5].

Wielu przedsiębiorców doskwiera problem, jakim jest zamknięcie importu ze wschodnich krajów, zaopatrujących pozostałą część świata w liczne surowce potrzebne do produkcji modułów informatycznych, mechanicznych czy elektronicznych. Zauważono również, jak wiele sprzętu leży pod gruzami zawalonych budynków czy na śmietnikach. Firmy zajmujące się odzyskiem oraz demontażem, w porozumieniu z przedsiębiorcami z Ukrainy, rozpoczęły niedawno próby odzysku sprzętu informatycznego. Ma on zostać przekazany odpowiednim korporacjom w celu prób odzysku jak największej ilości surowca tak, aby został on wykorzystany ponownie do wytworzenia kolejnych komponentów. W przypadku wciąż działających komputerów, zostają one przywrócone do jak najlepszego stanu, a następnie odesłane do zakładów pracy [6].

## 4. Dostępne rozwiązania

Rodzi się pytanie, czy bardziej opłacalna jest naprawa czy wyrzucenie problematycznego osprzętowania informatycznego? Otóż jest to uzależnione od rodzaju problemu – jeśli usterka jest poważna, a jej naprawa przekracza koszt kupna nowego, nieużywanego komputera, lepszym wyjściem jest inwestycja w nieużywane urządzenia. Sprowadza się to do problemu elektrośmieci, które w tym konkretnym przypadku stanowią masowy problem, obejmujący skalą prawie cały świat [4].

W wielu miastach na całym świecie znajdują się punkty zbiórki. Są one tworzone specjalnie z myślą o utylizacji elektroniki oraz przekazaniu jej na recykling. Proces ten polega na demontażu i odzyskaniu wielu metali, takich jak miedź, srebro czy złoto, które zostaną później wykorzystane ponownie. Dzięki tego rodzaju działaniom ograniczana jest konieczność wydobywania surowców naturalnych. Ostatecznie z elektrośmieci pozostają tylko opakowania ze szkła, stali czy plastiku (rys. 3 i 4), które podlegają standardowemu recyklingowi [4].



Rysunek 3. Zniszczone telefony stacjonarne

Źródło: fot. M. Nowosad.



Rysunek 4. Obudowy komputerów pozostałe po demontażu

Źródło: fot. M. Nowosad.

## 5. Inny aspekt

Mając za sobą jedną z największych katastrof, jaka dotknęła współczesny świat, czyli pandemię Covid-19, wiele placówek oświaty nakłada duży nacisk na edukację informatyczną. Skutkiem takiego działania jest przede wszystkim zmiana nastawienia młodzieży do technologii informatycznej. Bardzo częstym zjawiskiem jest zakup przez rodziców droższego komputera/laptopa, pod pretekstem większego komfortu pracy oraz nauki. W związku z czym stare urządzenie trafia przeważnie do kąta na strychu czy do piwnicy. Podobnie jest z telefonami komórkowymi oraz urządzeniami typu smartwatch. Młodzież w obecnych czasach bardzo często, działając pod presją rówieśników, wydaje duże kwoty pieniężne, aby kupować coraz to nowsze urządzenia i w ten sposób zaskarbić sobie popularność wśród rówieśników. Takie przypadki występują coraz częściej, zwłaszcza w odniesieniu do urządzeń mobilnych. Jest to ogromny problem, ze względu na ograniczoną świadomość społeczeństwa na temat postępowania z nieużywanymi urządzeniami informatycznymi czy telekomunikacyjnymi [8]. Dotyczy to również przedsiębiorstw, które wykorzystują sprzęt komputerowy do prze-

chowywania i przetwarzania danych. Tymczasem istnieje konieczność ochrony informacji zapisanych na dyskach twardych i pamięciach, które nie są już używane [10].

## 6. Wnioski

Reasumując, obecnie stajemy wobec wielkiego wyzwania, jakim jest potrzeba postępowania w zgodzie z zasadami ekologii [9], przy jednoczesnej konieczności oszczędzania, szczególnie w przedsiębiorstwach dotkniętych skutkami pandemii. Przedstawione w rozdziale nastawienie społeczeństwa oraz przedsiębiorstw do użytkowania sprzętu komputerowego sugeruje, że istnieje szansa w naprawianiu komputerów, pozyskiwaniu ich podzespołów i ich ponownym wykorzystywaniu. Możemy zauważyć, iż jest to pomysł na prowadzenie działalności gospodarczej. Młodzi przedsiębiorcy, którzy podejmą się tego wyzwania, mogą liczyć na pomoc w rozwinięciu biznesu, oferowaną przez Inkubator Przedsiębiorczości Politechniki Lubelskiej. Zapoczątkowanie takiego działania nie jest trudne, zważywszy na obecne niskie ceny zużytych, lecz wciąż funkcjonalnych komponentów informatycznych oraz możliwości składania z nich kolejnych komputerów, które następnie będą sprzedawane firmom. Młody przedsiębiorca ma zatem możliwość wejść na rynek informatyczny, jeden z najbardziej prosperujących na świecie. Będzie mógł tam przeprowadzać następne działania finansowe w celu rozbudowy sieci swojej firmy, a tym samym mieć szansę uzyskiwania większych zysków jako wyników swojej pracy.

## Literatura

- [1] Gramwzielone.pl, *Chińczycy podnieśli ceny płytek krzemowych i ogniw PV*, <https://www.gramwzielone.pl/energia-sloneczna/106193/chinczycy-podniesli-ceny-plytek-krzemowych-i-ogniw-pv> (dostęp: 9.03.2023).
- [2] Dudko Z., *Kiedy warto kupić nowy komputer?*, <https://www.zuzkapisze.pl/2022/03/kiedy-warto-kupic-nowy-komputer.html> (dostęp: 9.03.2023).
- [3] Garbarek N., *Z tego powodu ludzie zmieniają smartfony. Robisz tak samo?*, <https://www.benchmark.pl/aktualnosci/najczestsze-usterki-smartfonow-dlaczego-kupujemy-nowe.html> (dostęp: 9.03.2023).
- [4] Półchłopek J., *Elektrośmieci – co z nimi robić?*, <https://swiatoze.pl/elektrosmieci-co-z-nimi-zrobic/> (dostęp: 9.03.2023).
- [5] Pietrzak A., *Konflikt na Wschodzie nie taki straszny polskim eksporterom i importerom?*, <https://www.akcenta.pl/clanky-konflikt-na-wschodzie-nie-taki-straszny-polskim-eksporterom-i-importerom.html> (dostęp: 12.03.2023).
- [6] HitIT, *Zadbaj o nasze wspólne dobro – recykling sprzętu komputerowego*, <https://hitit.pl/zadbaj-o-nasze-wspolne-dobro-recykling-sprzetu-komputerowego/> (dostęp: 12.03.2023).

- [7] Wąsowski M., *Ceny na jednym rynku oszalały. Klienci wściekli, a firmy w kropce*, <https://businessinsider.com.pl/finanse/handel/ceny-kartograficznych-nvidia-rtx-3000-amd-radeon-6000-sytuacja-na-rynku/4v8f42w> (dostęp: 12.03.2023).
- [8] Chomiuk M., Bodecka M., *Zagrożenia i uzależnienia komputerowe*, <https://www.uzaleznieniabehawioralne.pl/siecioholizm/zagrozenia-i-uzaleznienia-komputerowe/> (dostęp: 12.03.2023)
- [9] Paśnikowska-Łukaszuk M., *Filozofia stylu życia „zero waste” w aspekcie ochrony środowiska*, [w:] red. D. Wójcicka-Migasiuk, *Wybrane technologie informatyczne w aspektach zrównoważonego rozwoju*, Wydawnictwo Politechniki Lubelskiej, Lublin 2020, ss. 138–147.
- [10] Buraczyńska B., *Wybrane aspekty zarządzania bezpieczeństwem sklepów internetowych*, [w:] red. M. Śniadkowski, *Wyniki badań interdyscyplinarnych w aspekcie edukacji techniczno-informatycznej i bezpieczeństwa*, Wydawnictwo Politechniki Lubelskiej, Lublin, 2020, ss. 69–79.

## **Repair or trash**

### **Abstract**

Nowadays, the topic of computer repair is considered from many aspects, given the rapid development of technology and the shortage of spare parts. New technologies are being developed day by day, applied to newer and newer computer components making it unprofitable to repair old ones, due to the price. Taking into account the events of recent years, we can clearly say how inflation has affected entrepreneurs, who, depending on the defect, do not buy new hardware, but renovate the old ones. By refurbishing, we not only save on IT components, but also reduce the amount of electro-waste.

*Keywords: electro-waste, information technology, IT components.*

**Emilia Popławska<sup>1</sup>**

## **Hipergrafy mieszane i ich kolorowanie<sup>2</sup>**

### **Streszczenie**

Niniejsza praca dokonuje przeglądu najważniejszych pojęć teorii grafów oraz teorii hipergrafów. Opisano wybrane własności hipergrafów mieszanych oraz ich kolorowania. Przedstawiono przykłady hipergrafów mieszanych kolorowalnych i niekolorowalnych. Następnie zaprezentowano wybrane zastosowania hipergrafów mieszanych.

*Słowa kluczowe: hipergraf mieszany, kolorowanie hipergrafów mieszanych, dolna liczba chromatyczna.*

### **Wstęp**

Temat tej pracy dotyczy zagadnienia hipergrafów mieszanych, które powstały niedawno, bo w latach 90. XX wieku. Chociaż hipergrafy mieszane są nowym odkryciem, to mają już wiele interesujących zastosowań.

Kolorowanie grafów jest jednym z podstawowych pojęć teorii grafów. Pierwsze wzmianki sięgają 1852 roku, kiedy Francis Guthrie zapytał, czy cztery kolory wystarczą, aby pokolorować dowolną mapę geograficzną w taki sposób, aby żadne dwa państwa mające wspólną granicę nie miały tego samego koloru (hipoteza o czterech kolorach – obecnie twierdzenie o czterech kolorach).

Z biegiem czasu zauważono, że wiele problemów dałoby się modelować hipergrafami, które stanowią uogólnienie grafów. Claude Berge zdefiniował najważniejsze pojęcia teorii grafów w monografii zatytułowanej *Grafy i hipergrafy* [1] w 1973 roku.

W 1992 roku wprowadzono pojęcie *C*-krawędzi (wtedy nazywanej antykrawędzią). Hipergrafy mieszane zawierają dwie rodziny hiperkrawędzi: krawędzie klasyczne i ich przeciwieństwa, które można traktować jako antykrawędzie. W każdym kolorowaniu klasyczne krawędzie mają co najmniej dwa wierzchołki różnych kolorów, podczas gdy antykrawędzie mają co najmniej dwa wierzchołki tego samego koloru. Zależności między krawędziami i antykrawędziami stanowią główną cechę kolorowania hipergrafów mieszanych oraz wnoszą do teorii kolo-

---

<sup>1</sup> Mgr inż. Emilia Popławska, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>2</sup> Niniejszy tekst stanowi fragment pracy dyplomowej pod tytułem *Kolorowanie i wybrane zastosowania hipergrafów mieszanych*, obronionej w dn. 7 lipca 2022 na Wydziale Podstaw Techniki Politechniki Lubelskiej.

rowania wiele nowych własności, takich jak kolorowalność, jednoznaczą kolorowalność, dolną i górną liczbę chromatyczną.

## 1. Wstęp do teorii grafów

Od 1852 roku przeprowadzono wiele analiz dotyczących kolorowania grafów, co doprowadziło do powstania jednej z najszerzej badanych i stosowanych dziedzin, jaką jest teoria grafów.

Przedstawimy teraz podstawowe pojęcia związane z grafami i ich kolorowaniem.

Niech  $V = \{v_1, v_2, \dots, v_n\}$  będzie niepustym, skończonym zbiorem, a zbiór  $E$  będzie zbiorem dwuelementowych podzbiorów zbioru  $V$ .

Grafem prostym  $G$  nazywamy uporządkowaną parę  $G = (V, E)$ , gdzie zbiór  $V$  nazywamy zbiorem wierzchołków, a  $E$  zbiorem krawędzi.

**Definicja 1.1.** [4, 5] Grafem prostym  $G$  nazywamy uporządkowaną parę  $G = (V, E)$ , gdzie zbiór  $V$  nazywamy zbiorem wierzchołków, a  $E$  zbiorem krawędzi.

**Definicja 1.2.** [5] Wierzchołki  $v_i$  i  $v_j$  nazywamy incydentnymi (sąsiednimi), jeżeli istnieje krawędź, która je łączy.

Mówi się również, że wierzchołki te są incydentne z tą krawędzią.

**Definicja 1.3.** [5] Grafem pełnym nazywamy graf, w którym każde dwa wierzchołki są sąsiednie.

W przypadku grafu krawędź jest incydentna z dwoma wierzchołkami. Rozszerzeniem pojęcia grafu jest hipergraf, w którym krawędź może być incydentna z dowolną liczbą wierzchołków.

Przedstawimy teraz podstawowe definicje związane z pojęciem hipergrafu.

**Definicja 1.4.** [2, 4] Hipergrafem  $H$  nazywamy uporządkowaną parę  $H = (V, \mathcal{E})$ , gdzie  $V$  jest zbiorem wierzchołków, a  $\mathcal{E}$  jest rodziną podzbiorów zbioru  $V$ .

Elementy  $E_1, E_2, \dots, E_m$  rodziny  $\mathcal{E}$  nazywamy hiperkrawędziami.

Zatem w grafie pełnym każde dwa wierzchołki są połączone krawędzią.

**Definicja 1.5.** [4]  $k$ -krawędzią nazywamy krawędź zawierająca  $k$  wierzchołków.

Będziemy używali oznaczeń:

- $V(G), V(H)$  – zbiór wierzchołków odpowiednio grafu  $G$ , hipergrafu  $H$ ,
- $E(G), \mathcal{E}(H)$  – zbiór odpowiednio krawędzi grafu  $G$ , hiperkrawędzi hipergrafu  $H$ .

**Definicja 1.6.** [6] Rzędem hipergrafu  $H$  nazywamy liczbę jego wierzchołków  $n(H) = |V(H)|$ .

**Definicja 1.7.** [6] Rozmiarem hipergrafu  $H$  nazywamy liczbę jego krawędzi  $m(H) = |\mathcal{E}(H)|$ .

**Definicja 1.8.** [4] Podhipergrafem hipergrafu  $H_1$  nazywamy hipergraf  $H_2$ , jeśli spełnione są warunki  $V(H_1) \subseteq V(H_2)$  oraz  $E(H_1) \subseteq E(H_2)$ .

**Definicja 1.9.** [4] Podhipergrafem indukowanym krawędziowo hipergrafu  $H_1$  nazywamy hipergraf  $H_2$ , jeżeli krawędzie ze zbioru  $E(H_2)$  należą również do zbioru  $E(H_1)$  oraz zbiór wierzchołków  $V(H_2)$  zawiera tylko wierzchołki incydentne ze zbiorem krawędzi  $E(H_2)$ .

## 2. Kolorowanie hipergrafów mieszanych

### 2.1. Hipergrafy mieszane – podstawowe definicje

Niech  $V = \{v_1, v_2, \dots, v_n\}$  dla  $n \geq 1$  będzie dowolnym zbiorem skończonym. Niech  $\mathcal{C} = \{C_1, C_2, \dots, C_l\}$  oraz  $\mathcal{D} = \{D_1, D_2, \dots, D_m\}$ , gdzie  $I = \{1, 2, \dots, l\}$  oraz  $J = \{1, 2, \dots, m\}$  będą rodzinami podzbiorów zbioru  $V$  takimi, że dla każdego z elementów rodziny  $\mathcal{C}$  oraz  $\mathcal{D}$  zachodzi warunek  $|C \cup D| \geq 2$ .

**Definicja 2.1.** [3] Hipergrafem mieszanym nazywamy trójkę  $(V, \mathcal{C}, \mathcal{D})$ , gdzie zbiór  $V$  jest zbiorem wierzchołków,  $\mathcal{C}$  oraz  $\mathcal{D}$  są rodzinami podzbiorów zbioru  $V$  nazywanych odpowiednio  $C$ -krawędziami oraz  $D$ -krawędziami.

Analogicznie jak w przypadku grafów i hipergrafów stosujemy oznaczenia:

- $V(H)$  – zbiór wierzchołków hipergrafu mieszanego  $H$ ,
- $\mathcal{C}(H)$  –  $C$ -krawędzie hipergrafu mieszanego  $H$ ,
- $\mathcal{D}(H)$  –  $D$ -krawędzie hipergrafu mieszanego  $H$ .

Niech zbiór  $\{1, 2, \dots, \lambda\}$  będzie zbiorem kolorów, gdzie  $\lambda \in \mathbb{N}$ .

**Definicja 2.2.** [3] Właściwym  $\lambda$ -kolorowaniem hipergrafu mieszanego  $H = (V, \mathcal{C}, \mathcal{D})$  nazywamy odwzorowanie  $c: V \rightarrow \{1, 2, \dots, \lambda\}$  spełniające warunki:

- każda  $C$ -krawędź ma co najmniej dwa wierzchołki tego samego koloru,
- każda  $D$ -krawędź ma co najmniej dwa wierzchołki różnych kolorów.

Jeżeli hipergraf mieszany posiada co najmniej jedno kolorowanie, to mówimy, że jest kolorowalny, w przeciwnym wypadku, że jest niekolorowalny.

Będziemy mówić, że dwa kolorowania  $c_1, c_2$  hipergrafu mieszanego  $H$  są różne, jeśli istnieje wierzchołek  $v$  taki, że  $c_1(v) \neq c_2(v)$ . Zatem każda permutacja kolorów daje nowe właściwe kolorowanie.

Rysunek 1 przedstawia hipergraf mieszany  $H = (V, \mathcal{C}, \mathcal{D})$  rzędu  $n(H) = 6$ , gdzie

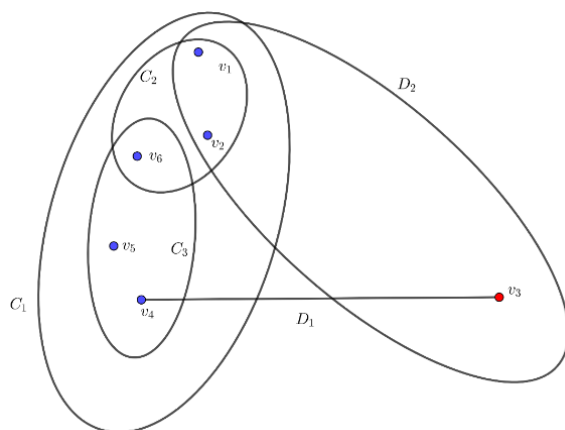
$$V(H) = \{v_1, v_2, v_3, v_4, v_5, v_6\}, \quad (1)$$

$$\mathcal{C}(H) = \{C_1, C_2, C_3\} = \{\{v_1, v_2, v_4, v_5, v_6\}, \{v_1, v_2, v_6\}, \{v_4, v_5, v_6\}\}, \quad (2)$$

$$\mathcal{D}(H) = \{D_1, D_2\} = \{\{v_1, v_2, v_3\}, \{v_3, v_4\}\}. \quad (3)$$

Widzimy, że jest on kolorowalny, gdyż istnieje dla niego kolorowanie właściwe  $c_1$  zobrazowane na rysunku 1. Wszystkie  $D$ -krawędzie mają co najmniej dwa wierzchołki o różnych kolorach, zaś  $C$ -krawędzie mają przynajmniej dwa wierzchołki o tych samych barwach.





Rysunek 1. Hipergraf mieszany  $H$  oraz jego kolorowanie właściwe  $c_1$

Źródło: opracowanie własne.

Zauważmy, że w przypadku właściwego  $\lambda$ -kolorowania nie wszystkie  $\lambda$  kolory muszą zostać użyte. Gdy użyjemy dokładnie  $\lambda$  kolorów, to będziemy mieli czynienia za ścisłym  $\lambda$ -kolorowaniem.

Z liczbą użytych kolorów związane jest pojęcie dolnej liczby chromatycznej (ang. *lower chromatic number*) oraz górnej liczby chromatycznej (ang. *upper chromatic number*).

**Definicja 2.3.** [3] Dolną liczbą chromatyczną  $\chi(H)$  hipergrafu mieszanego  $H$  nazywamy minimalną wartość  $\lambda$ , dla której istnieje właściwe  $\lambda$ -kolorowanie hipergrafu.

**Definicja 2.4.** [3] Ścisłym  $i$ -kolorowaniem hipergrafu mieszanego  $H$  nazywamy takie właściwe  $i$ -kolorowanie, w którym użyto wszystkich  $i$  kolorów, gdzie  $i \in N$ .

Ścisłe  $i$ -kolorowania mogą istnieć dla  $1 \leq i \leq \lambda$ . Zauważmy, że kolorowanie przy użyciu  $\chi(H)$  kolorów jest ścisłym kolorowaniem, ponieważ  $\chi(H)$  jest dolną granicą liczby możliwych kolorów dla właściwego  $\lambda$ -kolorowania. Górną granicę liczby możliwych kolorów dla właściwego  $\lambda$ -kolorowania możemy zdefiniować następująco:

**Definicja 2.5.** [3] Górną liczbą chromatyczną  $\bar{\chi}(H)$  hipergrafu mieszanego  $H$  nazywamy największą liczbę  $i$ , dla której istnieje ścisłe  $i$ -kolorowanie.

Niech  $Y \subset V$ . Zbiór  $Y$  nazywamy monochromatycznym, jeśli wierzchołki do niego należące są takiego samego koloru. Jeżeli dla  $y \in Y$  kolory  $c(y)$  są parami różne, to zbiór  $Y$  będziemy nazywali polichromatycznym.

W przypadku właściwego kolorowania hipergrafów mieszanych,  $D$ -krawędzie nie są zbiorami monochromatycznymi, zaś  $C$ -krawędzie nie są zbiorami polichromatycznymi.

Każde ściśle  $i$ -kolorowanie hipergrafu mieszanego  $H$  prowadzi do podziału zbioru wierzchołków  $V$  na  $i$  niepustych monochromatycznych podzbiorów  $V_1, V_2, \dots, V_i$ , zwanych klasami kolorów, w taki sposób, że każda  $C$ -krawędź ma przynajmniej dwa wierzchołki w tym samym podzbiorku oraz każda  $D$ -krawędź ma przynajmniej dwa wierzchołki znajdujące się w różnych podzbiorkach. Taki podział nazywamy podziałem wykonalnym (ang. *feasible partition*) hipergrafu mieszanego  $H$ .

**Definicja 2.6.** [3] Spektrum chromatycznym hipergrafu mieszanego  $H$  nazywamy wektor liczb całkowitych  $R(H) = (r_1, r_2, \dots, r_n)$ , gdzie dla  $k \in \{1, \dots, n\}$   $r_k$  jest liczbą możliwych  $k$ -kolorowań. Kolorowanie, które dzieli zbiór wierzchołków na te same podzbiory, traktujemy jako jedno kolorowanie.

Każdy kolorowalny hipergraf mieszany posiada dolną i górną liczbę chromatyczną, więc spektrum chromatyczne ma postać

$$R(H) = (0, \dots, 0, r_{\bar{\chi}}, \dots, r_{\bar{\chi}}, 0, \dots, 0). \quad (4)$$

Każdy podział wykonalny na  $i$  klas kolorów określa  $i$  ściśle  $i$ -kolorowań różniących się między sobą permutacją kolorów. Stąd liczba ściśle  $i$ -kolorowań jest równa  $r_i \cdot i!$ . Jeśli mamy  $\lambda > i$  kolorów, to otrzymujemy  $\binom{\lambda}{i}$  sposobów wyboru podzbioru  $i$  kolorów. Zatem liczba właściwych  $\lambda$ -kolorowań generowanych przez wszystkie możliwe podziały na  $i$  podzbiorów wynosi

$$\binom{\lambda}{i} r_i \cdot i! = \frac{\lambda!}{i! \cdot (\lambda-i)!} r_i \cdot i! = \frac{\lambda \cdot (\lambda-1) \cdot (\lambda-2) \cdot \dots \cdot (\lambda-i+1) \cdot (\lambda-i)!}{(\lambda-i)!} r_i = \lambda^{(i)} \cdot r_i. \quad (5)$$

**Definicja 2.7.** [3, 8] Wielomianem chromatycznym  $P(H, \lambda)$  hipergrafu mieszanego  $H$  nazywamy wielomian zmiennej  $\lambda$  określający liczbę właściwych  $\lambda$ -kolorowań hipergrafu mieszanego  $H$ . Ma on następującą postać

$$P(H, \lambda) = \sum_{i=\bar{\chi}(H)}^{\bar{\chi}(H)} r_i(H) \cdot \lambda^{(i)}. \quad (6)$$

Dla kolorowalnych hipergrafów mieszanych maksymalny stopień wielomianu chromatycznego jest równy górnej liczbie chromatycznej  $\bar{\chi}$ , zaś współczynnik przy najwyższej potędze wielomianu chromatycznego wynosi  $r_{\bar{\chi}}$ . Dla niekolorowalnego hipergrafu mieszanego  $H$  mamy

$$\chi(H) = 0, \quad \bar{\chi}(H) = 0, \quad R(H) = (0, 0, \dots, 0), \quad P(H, \lambda) = 0. \quad (7)$$

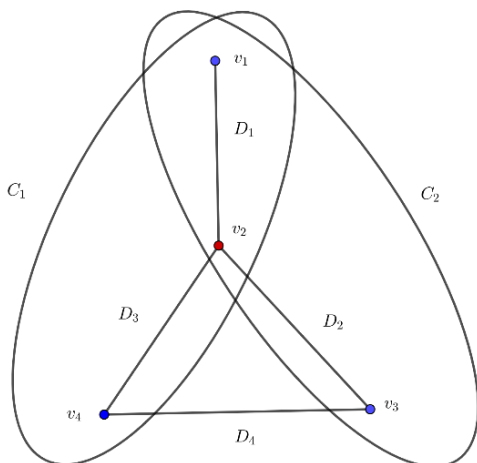
Niech dany będzie hipergraf mieszany  $H_1 = (V, \mathcal{C}, \mathcal{D})$ , gdzie

$$V(H_1) = \{v_1, v_2, v_3, v_4\}, \quad (8)$$

$$\mathcal{C}(H_1) = \{C_1, C_2\} = \{\{v_1, v_2, v_4\}, \{v_1, v_2, v_3\}\}, \quad (9)$$

$$\mathcal{D}(H_1) = \{D_1, D_2, D_3, D_4\} = \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_2, v_4\}, \{v_3, v_4\}\}. \quad (10)$$

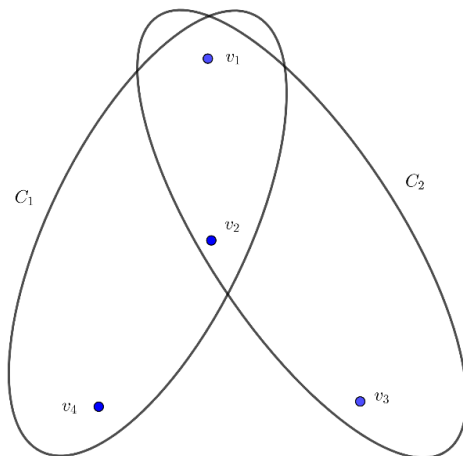
Został on przedstawiony na rysunku 2. Dla tego hipergrafu nie istnieje właściwe kolorowanie, ponieważ nadając wierzchołkom kolory napotykamy sytuację, w której krawędź  $D_4$  ma wierzchołki o tych samych kolorach.



**Rysunek 2. Niekolorowalny hipergraf mieszany  $H_1$**

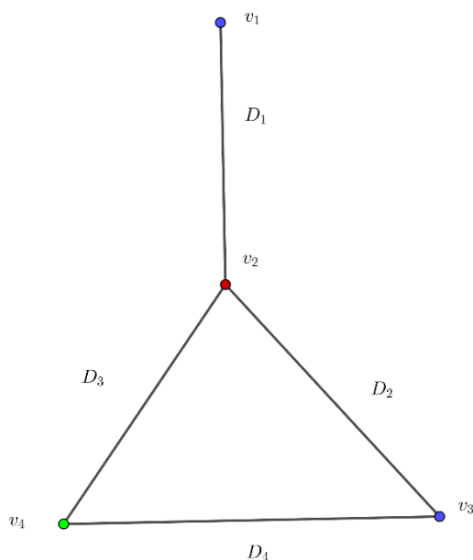
Źródło: opracowanie własne.

Zauważmy, że jeśli  $\mathcal{C} = \emptyset$ , to kolorowanie hipergrafu mieszanego  $H = (V, \emptyset, \mathcal{D})$  będzie kolorowaniem hipergrafu. Jeśli dla hipergrafu mieszanego  $H$  zachodzi, że  $\mathcal{C} = \mathcal{D}$ , to taki hipergraf będziemy nazywać *bi*-hipergrafem.



**Rysunek 3.  $\mathcal{C}$ -hipergraf hipergrafu mieszanego  $H$**

Źródło: opracowanie własne.



**Rysunek 4**  $D$ -hipergraf hipergrafu mieszanego  $H$

Źródło: opracowanie własne.

Dla dowolnego hipergrafu mieszanego  $H = (V, \mathcal{C}, \mathcal{D})$  możemy mówić o częściowym podhipergrafie mieszanym (ang. *partial mixed subhypergraph*)  $(V, \emptyset, \mathcal{D})$  oznaczanym przez  $H_D = (V, \mathcal{D})$  i nazywanym  $D$ -hipergrafem oraz o częściowym podhipergrafie mieszanym (ang. *partial mixed subhypergraph*)  $(V, \mathcal{C}, \emptyset)$  oznaczanym przez  $H_C = (V, \mathcal{C})$ , i nazywanym  $C$ -hipergrafem.

Warto zwrócić uwagę,  $D$ -hipergraf z krawędziami o rozmiarze 2 jest klasycznym grafem, zaś kolorowanie  $D$ -hipergrafów to kolorowanie zwykłych hipergrafów.

Gdybyśmy w hipergrafie na rysunku 2 usunęli  $C$ -krawędzie, to otrzymalibyśmy  $D$ -hipergraf, który jest kolorowalny. Postępując analogicznie z  $D$ -krawędziami otrzymalibyśmy  $C$ -hipergraf, który również jest kolorowalny. Na rysunkach 3 i 4 przedstawiliśmy podhipergrafy częściowe hipergrafu mieszanego  $H$ . W  $C$ -hipergrafie zamieszczonym na rysunku 3 moglibyśmy pokolorować wszystkie wierzchołki tym samym kolorem otrzymując właściwe  $C$ -kolorowanie.  $D$ -hipergraf z rysunku 4 jest zwykłym grafem. Aby go właściwie pokolorować, należy wierzchołkowi  $v_4$  z rysunku 2 nadać nowy kolor.

Dla dowolnego hipergrafu mieszanego  $H$  (kolorowalnego lub niekolorowalnego) częściowe hipergrafy mieszane są kolorowalne. Trywialnym kolorowaniem w przypadku  $C$ -hipergrafu jest użycie jednego koloru dla wszystkich wierzchołków, zaś w przypadku  $D$ -hipergrafu użycie  $n(H_D)$  kolorów. Mamy oczywiście, że:

$$\chi(H_C) = 1, r_1 = 1, \bar{\chi}(H_D) = n, r_n(H_D) = 1. \quad (11)$$

Dla kolorowalnego hipergrafu mieszanego  $H$  prawdziwa jest nierówność

$$1 \leq \chi(H_D) \leq \chi(H) \leq \bar{\chi}(H) \leq \bar{\chi}(H_C) \leq n. \quad (12)$$

Nierówność  $\chi(H_D) \leq \bar{\chi}(H_C)$  jest pierwszym warunkiem kolorowalności hipergrafów mieszanych. Jeśli nie jest spełniona dla hipergrafu mieszanego  $H$ , to hipergraf mieszany nie jest kolorowalny.

## 2.2. Hipergrafy mieszane niekolorowalne

Problem rozstrzygnięcia, czy istnieje co najmniej jedno właściwe kolorowanie hipergrafu mieszanego nazywamy problemem kolorowalności.

**Definicja 2.8.** [8] Minimalnym niekolorowalnym hipergrafem mieszanym nazywamy spójny hipergraf mieszany, który staje się kolorowalny po usunięciu  $C$ -krawędzi lub  $D$ -krawędzi.

Minimalny niekolorowalny hipergraf mieszany staje się również kolorowalny po usunięciu wierzchołka.

Mówiąc o usuwaniu wierzchołków będziemy mieli na myśli usuwanie silne, tzn. tworzymy podhipergraf mieszany indukowany przez pozostałe nieusunięte wierzchołki. Natomiast usuwaniem krawędzi będzie usuwanie słabym, tzn. tworzymy częściowy podhipergraf mieszany z pozostałych krawędzi.

**Definicja 2.9.** [8] Maksymalnym kolorowalnym częściowym podhipergrafem nazywamy częściowy podhipergraf  $H'$  hipergrafu mieszanego  $H$ , jeżeli dodanie dowolnej  $C$ -krawędzi lub  $D$ -krawędzi (lub dowolnego wierzchołka z incydentnymi krawędziami obu typów) hipergrafu mieszanego  $H$  do hipergrafu  $H'$  sprawia, że  $H'$  staje się niekolorowalny.

Zatem minimalne niekolorowalne podhipergrafy i maksymalne kolorowalne podhipergrafy niekolorowalnych hipergrafów mieszanych wyznaczają krytyczną granicę kolorowalności.

Jeśli  $H$  jest minimalnym hipergrafem niekolorowalnym, to zawiera  $|C| + |D|$  maksymalnych częściowych podhipergrafów kolorowalnych oraz  $|V|$  maksymalnych indukowanych podhipergrafów kolorowalnych.

Najbardziej niekolorowalnym przykładem (jeśli nie jest wymagana minimalność kolorowania) jest hipergraf mieszany  $\mathcal{U}_n = \left( V, \binom{n}{2}, \binom{n}{2} \right)$ , który jest niekolorowalny wraz z każdym ze swoich indukowanych podhipergrafów rzędu  $\geq 2$  dla każdego  $n = |V| > 2$ .

**Definicja 2.10.** [8] Dla hipergrafu mieszanego  $H = (V, C, D)$ , liczba niekolorowalności wierzchołków  $\Omega_V(H)$  (liczba niekolorowalności  $C$ -krawędzi  $\Omega_C(H)$ , liczba niekolorowalności  $D$ -krawędzi  $\Omega_D(H)$ ) jest minimalną liczbą wierzchołków ( $C$ -krawędzi,  $D$ -krawędzi), które należy usunąć, aby uzyskać hipergraf mieszany kolorowalny.

Dla hipergrafu mieszanego  $H$  spełniona jest nierówność

$$0 \leq \Omega_V(H) \leq n - 1. \quad (13)$$

Zauważmy, że skoro hipergraf mieszany  $\mathcal{U}_n$  jest niekolorowalny oraz jego indukowane podhipergrafy są również niekolorowalne, to liczba niekolorowalności wierzchołków wynosi

$$\Omega_V(\mathcal{U}_n) = n - 1. \quad (14)$$

Jeśli hipergraf mieszany  $H$  jest kolorowalny, to

$$\Omega_V(H) = \Omega_C(H) = \Omega_D(H) = 0. \quad (15)$$

Dla minimalnego kolorowalnego hipergrafu mieszanego  $H$  prawdziwa jest zależność

$$\Omega_V(H) = \Omega_C(H) = \Omega_D(H) = 1. \quad (16)$$

### 2.3. Wyznaczanie kolorowania hipergrafu mieszanego

Do wyznaczania kolorowania hipergrafu mieszanego, dolnej liczby chromatycznej, górnej liczby chromatycznej, spektrum chromatycznego oraz wielomianu chromatycznego stosujemy algorytm rozdzielania-ściągnięcia (ang. *splitting-contraction*). Podczas wyznaczania wielomianu chromatycznego  $P(H, \lambda)$  oraz spektrum chromatycznego  $R(H)$  hipergrafu mieszanego  $H = (V, \mathcal{C}, \mathcal{D})$ , dla którego  $\mathcal{C} = \{C_1, C_2, \dots, C_l\}$  oraz  $\mathcal{D} = \{D_1, D_2, \dots, D_m\}$ , gdzie  $I = \{1, 2, \dots, l\}$  oraz  $J = \{1, 2, \dots, m\}$ , stosujemy zasady: zasadę eliminacji, zasadę  $C$ -usuwania, zasadę  $D$ -usuwania, zasadę rozdzielania, zasadę ściągnięcia. Dokładny opis algorytmu można znaleźć w [8]. Za pomocą wymienionych zasad algorytm sprowadza wielomian chromatyczny hipergrafu mieszanego do sumy wielomianów chromatycznych grafów pełnych. Otrzymane liczby grafów pełnych odpowiadają wartościom  $r_i$  (liczby właściwych kolorowań hipergrafu mieszanego  $H$  przy użyciu  $i$  kolorów z pominięciem permutacji kolorów) [8].

Problem kolorowania hipergrafu mieszanego można modelować za pomocą programowania liniowego całkowitoliczbowego. Więcej na ten temat w [7, 8].

## 3. Wybrane zastosowanie kolorowania hipergrafów mieszanych

Za pomocą kolorowania hipergrafów mieszanych można modelować wiele problemów. Dziedzinami, w których wykorzystuje się hipergrafy mieszane są m.in. biologia molekularna, obliczenia równoległe, zarządzanie bazami danych, programowanie całkowitoliczbowe oraz genetyka [3, 7].

Każde pojęcie, które może być modelowane przez hipergraf mieszany  $H = (V, \mathcal{C}, \mathcal{D})$  jest wynikiem interakcji między dwiema rodzinami przeciwnych ograniczeń wyrażonych przez rodzinę  $\mathcal{C}$  oraz rodzinę  $\mathcal{D}$ . Jeśli hipergraf mieszany  $H$  jest niekolorowalny, to nie mamy rozwiązania problemu, który jest modelowany kolorowaniem hipergrafu mieszanego, a odpowiadający problem jest trak-

towany jako samozaprzeczący [8]. Im bardziej niekolorowalny jest hipergraf mieszany modelujący dane pojęcie, tym bardziej pojęcie to jest samozaprzeczące.

Mówiąc, że hipergraf mieszany  $H_1$  jest bardziej niekolorowalny niż hipergraf mieszany  $H_2$ , mamy na myśli sytuację, kiedy liczba usuniętych  $D$ -krawędzi bądź  $C$ -krawędzi w celu otrzymania hipergrafu mieszanego kolorowalnego jest większa dla hipergrafu mieszanego  $H_1$  niż dla hipergrafu mieszanego  $H_2$ .

Opiszemy teraz kilka wybranych zastosowań hipergrafów mieszanych.

### 3.1. Interpretacja liczb naturalnych

Kolorowanie hipergrafu mieszanego pozwala na interpretację pojęcia liczby naturalnej. Liczenie może być postrzegane jako sekwencyjne wybieranie obiektów, które w rzeczywistości są parami różne (różne kolory), ale z definicji są uważane za parami identyczne (wspólnego koloru). Jeśli zatem  $V$  jest zbiorem takich obiektów, to liczenie obiektów ze zbioru  $V$  odpowiada stopniowej konstrukcji hipergrafu mieszanego  $\mathcal{U}_n = (V, \binom{n}{2}, \binom{n}{2})$ . Pojęcie liczby naturalnej  $n = |V|$  powstaje w wyniku  $\binom{n}{2}$  elementarnych sprzeczności (minimalnych niekolorowalności) generowanych przez porównania parami. Jak widzieliśmy w podrozdziale 2.2, spośród wszystkich niekolorowalnych hipergrafów mieszanych o  $n$  wierzchołkach,  $\mathcal{U}_n$  jest najbardziej niekolorowalny, ponieważ każdy z jego nietrywialnych indukowanych podhipergrafów jest również niekolorowalny. Co więcej,  $\mathcal{U}_n$  jest jedynym takim hipergrafem mieszanym dla każdego  $n$ . W rzeczywistości prawdziwe liczenie zaczyna się w momencie rozpoczęcia porównań, czyli od liczby 2. Zatem dla każdego  $n$  mamy  $n - 1$  porównań, co pokrywa się z liczbą niekolorowalności wierzchołków  $\Omega_V(\mathcal{U}_n)$  [8].

### 3.2. Genetyka populacji

W populacji, która rozmnaża się płciowo występują dwa rodzaje cech dziedzicznych: dominujące i recesywne. Podczas krzyżowania dziecko dziedziczy cechę recesywną wtedy i tylko wtedy, gdy jest ona obecna u obojga rodziców. Niech  $V$  będzie zbiorem rodziców, a  $\mathcal{C}$  rodziną podzbiorów  $V$  takich, że każdy podzbiór ma jedną, wspólną cechę recesywną. Czy można podzielić zbiór  $V$  w taki sposób, aby cechy recesywne przeszły do następnego pokolenia? Niech  $V = M \cup F$ , gdzie  $M$  jest zbiorem samców, a  $F$  jest zbiorem samic. W rozmnażaniu płciowym parę tworzą osobniki o różnych płciach. Konstruujemy hipergraf mieszany w następujący sposób  $H = (V, \mathcal{C}, \binom{M}{2} \cup \binom{F}{2})$ , gdzie  $K_{|F|}$  oraz  $K_{|M|}$  tworzą dwa oddzielne pełne  $D$ -hipergrafy indukowane przez  $M$ ,  $F$ . Jeśli  $H$  jest kolorowalny, to każde właściwe kolorowanie daje pewien podział wykonalny zbioru rodziców  $V$ . Dla każdej cechy recesywnej  $C \in \mathcal{C}$  istnieje para  $m(C) \in M$ ,  $f(C) \in F$  o różnych płciach. Maksymalną liczbą klas w takim podziale jest  $\bar{\chi}(H)$ . Zauważmy, że

w każdym podziale na więcej niż  $\bar{\chi}(H)$  klas przynajmniej jedna cecha recesywna nie zostanie odziedziczona w następnym pokoleniu [8].

Rozważmy grupę osób składającą się z 3 kobiet i 3 mężczyzn. Wprowadzamy następujące oznaczenia:

- $v_1$  – kobieta o jasnych włosach,
- $v_2$  – kobieta z brakiem umiejętności zwijania języka w rurkę,
- $v_3$  – kobieta bez piegów na twarzy, o rudych włosach i niebieskich oczach,
- $v_4$  – mężczyzna o jasnych włosach,
- $v_5$  – mężczyzna o jasnych włosach i niebieskich oczach,
- $v_6$  – mężczyzna z brakiem umiejętności zwijania języka w rurkę.

W opisach osób w grupie wymieniliśmy tylko ich cechy recesywne, ponieważ dziedziczenie tych cech jest przedmiotem naszych badań.

W celu modelowania dziedziczenia cech recesywnych w danej grupie możemy zbudować hipergraf mieszany  $H = (V, \mathcal{C}, \mathcal{D})$ .  $D$ -krawędzie tego hipergrafu są krawędziami  $D$ -hipergrafów pełnych indukowanych przez zbiór trzech kobiet  $F$  oraz zbiór trzech mężczyzn  $M$ . Mamy

$$\begin{aligned} D_1 = \{v_1, v_2\}, D_2 = \{v_2, v_3\}, D_3 = \{v_1, v_3\}, D_4 = \{v_4, v_5\}, \\ D_5 = \{v_4, v_6\}, D_6 = \{v_5, v_6\}. \end{aligned} \quad (17)$$

$C$ -krawędzie powstają jako zbiory osób posiadających daną cechę recesywną. Mamy

- $C_1 = \{v_1, v_4, v_5\}$  dla cechy recesywnej jasne włosy,
- $C_2 = \{v_2, v_6\}$  dla cechy recesywnej brak umiejętności zwijania języka w rurkę,
- $C_3 = \{v_3, v_5\}$  dla cechy recesywnej niebieskie oczy.

Dla pozostałych cech nie tworzymy  $C$ -krawędzi, ponieważ występują tylko u jednej z 6 rozważanych osób, zatem nie jest możliwe ich odziedziczenie w kolejnym pokoleniu.

Powstały w ten sposób hipergraf mieszany oraz jego właściwe kolorowanie zostały przedstawione na rysunku 5.

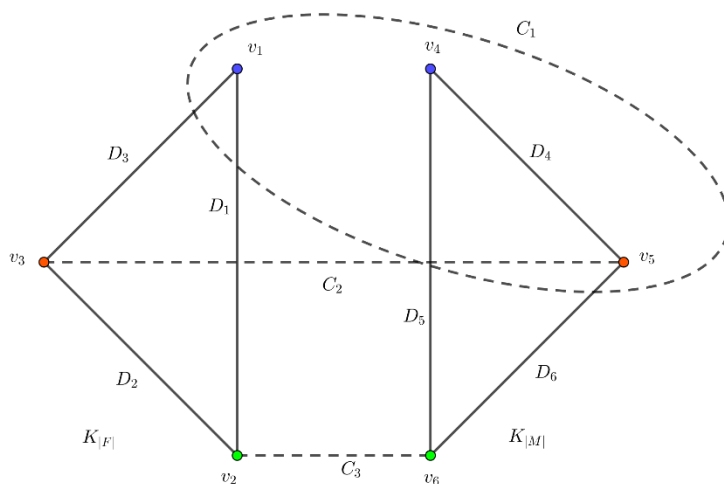
Właściwe kolorowanie dzieli zbiór wierzchołków na podzbiory

$$\{v_1, v_4\}, \quad \{v_2, v_6\}, \quad \{v_3, v_5\}. \quad (18)$$

Optymalną liczbą kolorów użytych do kolorowania jest 3. Jest to jedyne ściśle kolorowanie hipergrafu mieszanego  $H$ . Zatem hipergraf mieszany  $H$  jest jednoznacznie kolorowalny.

Warto zwrócić uwagę, że podział rodziców na podzbiory w wyniku kolorowania zapewnia dziedziczenie cech recesywnych jasne włosy, brak umiejętności zwijania języka w rurkę oraz niebieskie oczy w następnym pokoleniu.





**Rysunek 5. Hipergraf mieszany  $H$  modelujący dziedziczenie cech recesywnych oraz jego kolorowanie**

Źródło: opracowanie własne.

### 3.3. Obliczenia równoległe

Obliczenia równoległe są stosowane w celu przyspieszenia algorytmu. Większość komputerów posiada procesory wielordzeniowe, gdzie każdy rdzeń takiego procesora może realizować niezależne obliczenia. Obliczenia równoległe są modelem obliczeniowym dzielącym program na mniejsze operacje. Operacje te są wykonywane w tym samym czasie (równoległe) na kilku procesorach [9].

Niech będzie dany zbiór  $n$  problemów  $V = \{v_1, v_2, v_3, \dots, v_n\}$ . Załóżmy, że mamy nieograniczoną liczbę procesorów. Każdy problem może być rozwiązany na dowolnym procesorze. Istnieją jednak pewne zależności między problemami, ponieważ dane wyjściowe jednych reprezentują dane wejściowe innych. W związku z tym nie możemy używać procesorów po jednym procesorze dla każdego problemu, choć byłby to najlepszy wariant obliczeń równoległych. Ze względu na relacje między danymi wejściowymi i wyjściowymi wprowadzamy rodzinę zbiorów  $\mathcal{C} = \{C_1, C_2, \dots, C_l\}$ , gdzie  $C_i \subseteq V$  oraz  $i \in \{1, 2, \dots, l\}$ , będzie rodziną podzbiorów  $V$  takich, że co najmniej dwa problemy z każdego zbioru  $C_i$  muszą być rozwiązane na tym samym procesorze. Ze względu na związki między danymi wejściowymi i wyjściowymi, niech  $\mathcal{D} = \{D_1, D_2, \dots, D_m\}$ , gdzie  $D_j \subseteq V$ ,  $j \in \{1, 2, \dots, m\}$ , będzie rodziną podzbiorów  $V$  takich, że co najmniej dwa zadania z każdego  $D_j$  muszą być rozwiązane na różnych procesorach. Jaka jest maksymalna liczba procesorów, które można wykorzystać do rozwiązania wszystkich problemów ze zbioru  $V$ ? Odpowiedzią na to pytanie jest górna liczba chromatuyczna  $\bar{\chi}(H)$  hipergrafu mieszanego  $H = (V, \mathcal{C}, \mathcal{D})$  [8].

## 4. Podsumowanie

W tej pracy zdefiniowano pojęcia związane z teorią grafów oraz teorią hipergrafów niezbędne do analizy tematyki kolorowania hipergrafów mieszanych. Następnie przedstawiono najważniejsze własności hipergrafów mieszanych oraz ich kolorowania. Zaprezentowano wybrane zastosowania hipergrafów mieszanych.

Hipergrafy mieszane mają zastosowanie w genetyce. Mogą modelować dziedziczenie cech recesywnych. Wtedy w hipergrafie mieszanym wierzchołki reprezentują rodziców,  $D$ -krawędzie są krawędziami  $D$ -hipergrafu pełnego indukowanego przez zbiór kobiet oraz zbiór mężczyzn, zaś  $C$ -krawędzie „zbierają” osoby posiadające wspólną cechę recesywną. Właściwe kolorowanie hipergrafu mieszanego zamodelowanego w ten sposób dzieli zbiór rodziców na podzbiory tak, aby cechy recesywne zostały odziedziczone w następnym pokoleniu.

Liczby naturalne mogą być interpretowane za pomocą pojęć związanych z kolorowaniem hipergrafów mieszanych. Innym zastosowaniem hipergrafów mieszanych zaprezentowanym w pracy są obliczenia równoległe.

Reasumując, hipergrafy mieszane są nowym odkryciem, które bardzo szybko znalazło zastosowanie w wielu dziedzinach. Są dobrym narzędziem modelowania sytuacji w przypadku wprowadzenia dwóch kontrastowych względem siebie ograniczeń.

## Literatura

- [1] Berge C., *Graphs and Hypergraphs*, North-Holland Publishing Company, Amsterdam 1976.
- [2] Berge C., *Hypergraphs: Combinatorics of Finite Sets*, Elsevier Science Publishing Company, Inc., Amsterdam 1989.
- [3] Voloshin V., *The mixed hypergraphs*, “Computer Science Journal of Moldova” 1993, vol. 1, no. 1, pp. 45–52.
- [4] Łazuka E., *Wielomian chromatyczny jako narzędzie klasyfikacji hipergrafów*, [w:] red. I. Gorgol, *Matematyczne miscellanea*, t. 3, Wydawnictwo Politechniki Lubelskiej, Lublin 2018.
- [5] Wilson R., *Wprowadzenie do teorii grafów*, PWN, Warszawa, 2007.
- [6] Tomescu I., *Chromatic coefficients of linear uniform hypergraphs*, “Journal of Combinatorial Theory” 1998, vol. 72, pp. 229–235.
- [7] Lozovanu D., Voloshin V., *Integer programming models for colorings of mixed hypergraphs*, “Computer Science Journal of Moldova” 2000, vol. 8, no. 1 (22), pp. 64–74.
- [8] Voloshin V. I., *Coloring Mixed Hypergraphs: theory, algorithms and applications*, American Mathematical Society, Providence 2002.

- [9] *Lekcja 4: Obliczenia równoległe i rozproszone*, [w:] Kurs on-line: *Podstawy informatyki – program rozszerzony*, <https://pl.khanacademy.org/computing/ap-computer-science-principles/algorithms-101/x2d2f703b37b450a3:parallel-and-distributed-computing/a/parallel-computing> (dostęp: 29.05.2022).

## Mixed hypergraphs and their coloring

### Abstract

In the paper, selected terms from graph theory and hypergraph theory are defined. The most important properties of mixed hypergraphs and their coloring are described. Examples of colorable and uncolorable mixed hypergraphs are introduced. Then selected applications of mixed hypergraphs are presented.

*Keywords: mixed hypergraph, coloring mixed hypergraphs, lower chromatic number.*

**Albert Rachwał<sup>1</sup>**

## **Wybrane metody klasyfikacji w zadaniach uczenia maszynowego**

### **Streszczenie**

Niniejszy rozdział przedstawia przegląd popularnych metod uczenia maszynowego wykorzystywanych w procesie klasyfikacji danych. Wśród zaprezentowanych metod znajdują się: drzewo decyzyjne, las losowy, regresja logistyczna, analiza dyskryminacyjna, maszyna wektorów nośnych oraz metoda  $k$  najbliższych sąsiadów. Poza omówieniem powyższych metod, wykonano także ich porównanie w praktycznym zadaniu klasyfikacji. Zbiór danych użyty do analizy nosi nazwę Breast Cancer Wisconsin. Celem analizy jest określenie tego, czy nowotwór jest złośliwy na podstawie rozmiaru guza i innych parametrów związane ze zmianami nowotworowymi. Modele uczenia maszynowego mogą również pozwolić na wyodrębnienie najbardziej istotnych cech wpływających na określenie typu nowotworu.

*Słowa kluczowe: klasyfikacja, drzewo decyzyjne, las losowy, SVC, KNN, regresja logistyczna, analiza dyskryminacyjna.*

### **Wstęp**

Choroby nowotworowe są nieustannie jednym z największych utrapień współczesnych społeczeństw. Przykładowo w 2018 roku na całym świecie zdiagnozowano 18,1 miliona nowych przypadków raka, a 9,6 miliona osób zmarło z powodu tej choroby. Dane te podkreślają globalne obciążenie rakiem, z wyraźnymi różnicami regionalnymi w zakresie typów i śmiertelności nowotworów [11]. Nowotwory najprościej można sklasyfikować jako łagodne, złośliwe oraz miejscowo złośliwe.

Wcześniejsza identyfikacja pozwala na poprawę przeżywalności i zmniejszenie śmiertelności. Dlatego niezbędne staje się zaprojektowanie modelu komputerowego wspomaganego do wykrywania i klasyfikacji nowotworów [1].

Uważa się, że wczesna identyfikacja raka piersi jest kluczowa do obniżania śmiertelności kobiet, dlatego wykonuje się kilka różnych metod badań przesiewowych piersi, aby poprawić skuteczność diagnostyki tej choroby. Obecnie do wykrywania oraz lokalizowania nowotworu wykorzystywana jest histopatologia przy wsparciu sztucznej inteligencji i konwolucyjnych sieci

---

<sup>1</sup> Mgr inż. Albert Rachwał, Wydział Elektrotechniki i Informatyki, Politechnika Lubelska.

neuronowych [13], co w badaniu przesiewowym może umożliwić zidentyfikowanie obecności guza.

## 1. Opis zbioru badawczego

Do przeprowadzenia niniejszych badań został wykorzystany zbiór Breast Cancer Wisconsin [4]. Zbiór ten zawiera dane dotyczące 569 obserwacji guza u pacjentek, u których stwierdzono chorobę nowotworową, spośród nich 212 ma raka złośliwego, a 357 ma raka sklasyfikowanego jako łagodny. Na dane dotyczące pacjentów składają się 30 atrybutów opisujących przypadek nowotworu, który posiada badany, oraz jeden atrybut oznaczony cyfrą 1 lub 0, w zależności od tego, czy nowotwór został przez lekarzy uznany za łagodny lub złośliwy. Cechy są obliczane na podstawie cyfrowego obrazu guza piersi i opisują jądra komórkowe obecne na obrazie. Zbiór ten nie posiada żadnych braków danych. Poza identyfikatorem i polem oznaczającym rodzaj raka, dla każdego jądra komórkowego wyodrębniono dziesięć cech o wartościach rzeczywistych:

- a) promień (średnia odległości od środka do punktów na obwodzie),
- b) tekstura (odchylenie standardowe wartości skali szarości),
- c) obwód,
- d) obszar,
- e) gładkość (lokalna zmienność długości promienia),
- f) zwartość (obwód podniesiony do kwadratu / powierzchnia – 1,0),
- g) wklęsłość (dotkliwość wklęsłych fragmentów konturu),
- h) punkty wklęsłe (liczba wklęsłych części konturu),
- i) symetria,
- j) wymiar fraktalny („przybliżenie linii brzegowej” – 1).

Dla każdej z powyższych cech obliczono średnią, błąd standardowy oraz wartość dla przypadku „najgorszego”, co daje łącznie 30 atrybutów opisujących guza [4].

## 2. Dobór metod

Wśród metod wykorzystanych do klasyfikacji zbioru znajdują się: algorytm  $k$  najbliższych sąsiadów, regresja logistyczna, maszyna wektorów nośnych, drzewo decyzyjne, las losowy, liniowa analiza dyskryminacyjna oraz kwadratowa analiza dyskryminacyjna. Na początku pracy zbiór danych został podzielony na zbiory uczący i testowy, stanowiący odpowiednio 70% oraz 30% całego zbioru. Dla każdego z algorytmów przeprowadzono uczenie oraz dopasowanie zbiorów uczącego i testowego, na podstawie których otrzymano dokładność dopasowania, będącą kluczowym parametrem przy ocenie, która spośród użytych metod może mieć dobre wyniki w klasyfikacji nowych danych. Dla każdego zastosowanego algorytmu przeprowadzono selekcję cech, identyfikując pięć atrybutów o największym wpływie, ocenianym na podstawie ich ważności w modelu. W każdym rozdziale

została przedstawiona macierz pomyłek, jaką uzyskał dany model, dzięki czemu można dokładnie przyjrzeć się jakiego rodzaju błędy popełnia częściej.

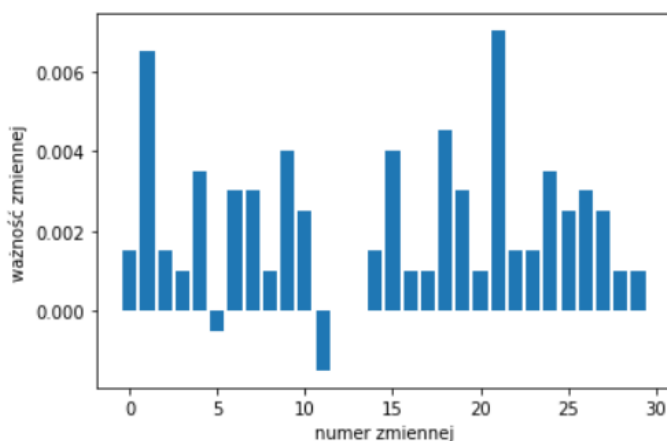
Pole „Stan Faktyczny” w macierzy pomyłek oznacza, czy przy przypadku w rzeczywistości był złośliwy, natomiast pole „Predykcja” mówi, w jaki sposób sklasyfikował obserwację model. W zbiorze testowym znajdowały się 63 przypadki nowotworu złośliwego, oznaczanego jako „Tak” i 108 niezłośliwego oznaczanego jako „Nie”.

### 3. $K$ najbliższych sąsiadów

Jedną z najbardziej podstawowych metod używanych do klasyfikacji jest algorytm  $k$  najbliższych sąsiadów (ang. *k nearest neighbors*, KNN). Należy on do domeny uczenia nadzorowanego i znajduje szerokie zastosowanie w rozpoznawaniu wzorców, eksploracji danych i wykrywaniu anomalii. Zyskał on popularność, ponieważ jest nieparametryczny, co oznacza, że nie przyjmuje żadnych podstawowych założeń dotyczących przynależności danych do specyficznego rozkładu.

Podstawowym założeniem metody  $k$  najbliższych sąsiadów jest to, że podobne obserwacje znajdują się blisko siebie, a odstające są zwykle obserwacjami samotnymi, znajdującymi się dalej od skupienia podobnych sobie obserwacji. Właściwy wybór parametru  $k$  zależy od rodzaju danych. Z jednej strony im większą wartość przypiszemy  $k$ , tym mniejszy wpływ na klasyfikację będą miały obecne w danych szumy. Z drugiej strony jednak podział na większą liczbę skupień sprawi, że granice między poszczególnymi klasami staną się mniej wyraźne.

Wykres na rysunku 1 przedstawia ważność zmiennych dla modelu KNN. Na jego podstawie 5 najważniejszych cech oraz ich wartości zostały ukazane w tabeli 1.



Rysunek 1. Ważność poszczególnych zmiennych dla modelu  $k$  najbliższych sąsiadów

Źródło: opracowanie własne.

**Tabela 1. Najważniejsze zmienne badanego zbioru uzyskane na podstawie modelu  $k$  najbliższych sąsiadów**

Nazwa zmiennej	Ważność zmiennej
Najgorsza tekstura	0,00704
Średni promień	0,00653
Błąd symetrii	0,00452
Błąd zwartości	0,00402
Średni wymiar fraktalny	0,00402

Źródło: opracowanie własne.

W tabeli 2 znajduje się macierz pomyłek dla  $k$  najbliższych sąsiadów. Algorytm wykonał jedną klasyfikację fałszywie pozytywną oraz 6 fałszywie negatywnych. Całkowita dokładność modelu wyniosła 95,9%.

**Tabela 2. Macierz pomyłek uzyskana na podstawie modelu  $k$  najbliższych sąsiadów**

		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	57	6
	Nie	1	107

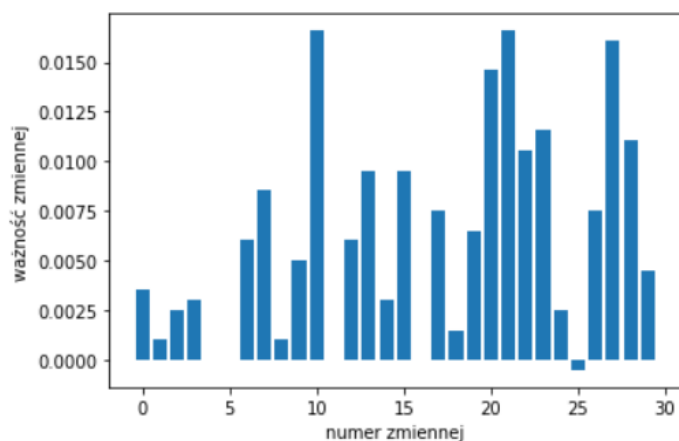
Źródło: opracowanie własne.

## 4. Regresja logistyczna

Regresja logistyczna (ang. *logistic regression*), zwana także regresją logitową, jest często używana do klasyfikacji i analiz predykcyjnych. Szacuje prawdopodobieństwo wystąpienia zdarzenia na podstawie zestawu danych zmiennych niezależnych, ponieważ wynik jest prawdopodobieństwem, zmienna zależna jest ograniczona od 0 do 1. Transformacja logitowa, używana w regresji logistycznej, przekształca prawdopodobieństwo na logarytm szansy (log-odds), co matematycznie wyraża się jako stosunek prawdopodobieństwa sukcesu do prawdopodobieństwa niepowodzenia [7].

Wykres na rysunku 2 ilustruje wagę poszczególnych zmiennych w modelu regresji logistycznej, a najbardziej istotne z nich są szczegółowo wymienione w tabeli 3, odnoszącej się do tego wykresu.

W tabeli 4 znajduje się macierz pomyłek dla regresji logistycznej. Algorytm wykonał trzy klasyfikacje fałszywie pozytywne oraz dwie fałszywie negatywne. Całkowita dokładność modelu wyniosła 97,07%.



Rysunek 2. Ważność poszczególnych zmiennych dla modelu regresji logistycznej

Źródło: opracowanie własne.

Tabela 3. Najważniejsze cechy badanego zbioru danych wyznaczone z modelu regresji logistycznej

Nazwa zmiennej	Ważność zmiennej
Błąd promienia	0,01658
Najgorsza tekstura	0,01658
Najgorsze punkty wklęsłe	0,01608
Najgorszy promień	0,01457
Najgorsze pole	0,01156

Źródło: opracowanie własne.

Tabela 4. Macierz pomyłek uzyskana na podstawie modelu regresji logistycznej.

		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	61	2
	Nie	3	105

Źródło: opracowanie własne.

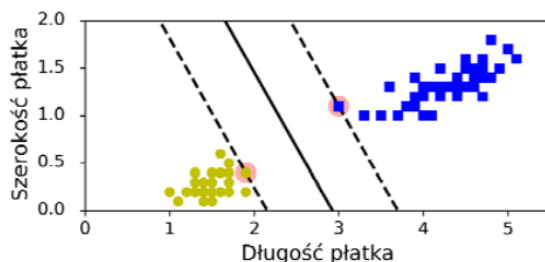
## 5. Maszyna wektorów nośnych

Maszyny wektorów nośnych (ang. *support vector machines*, SVM) są nadzorowanymi modelami uczenia się, które analizują dane do celów analizy klasyfikacji i regresji. Jest to jedno z najpopularniejszych rozwiązań w świecie uczenia maszynowego, ponieważ dzięki swojej wszechstronności może być zastosowane do przeprowadzania klasyfikacji liniowej, nieliniowej, regresji, a nawet do wykry-



wania elementów odstających [8]. SVM odwzorowuje przykłady treningowe na punktach w przestrzeni, aby zmaksymalizować szerokość luki między dwiema kategoriami. Nowe przykłady są następnie mapowane w tej samej przestrzeni i przewiduje się, że będą należeć do kategorii na podstawie tego, po której stronie luki się znajdują [12].

Na rysunku 3 został zaprezentowany model SVM w zadaniu klasyfikacji zbioru danych Iris [5], a jego centralną część zajmuje granica decyzyjna klasyfikatora. Obserwacje znajdujące się na liniach przerywanych są nazywane wektorami nośnymi. Dodawanie nowych obserwacji na zewnątrz granicy decyzyjnej nie spowoduje jej przesunięcia. Wewnętrzna linia ciągła stanowi granicę podziału, przy czym utrzymuję możliwie duży dystans od wektorów nośnych. Ujmując najprościej, zadaniem SVM jest wyznaczenie jak najszerszego przedziału pomiędzy klasami utworzonymi przez pogrupowane obserwacje. Proces ten nazywamy klasyfikowaniem maksymalnego marginesu (ang. *large margin classification*).

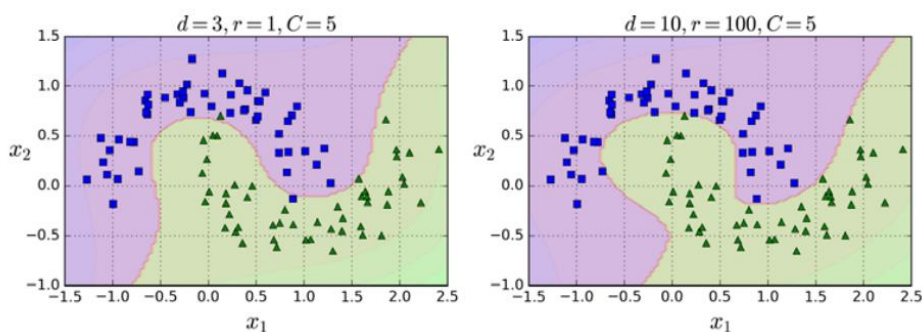


**Rysunek 3. Klasyfikowanie maksymalnego marginesu dla zbioru Iris**

Źródło: A. Geron, *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Helion SA, Gliwice 2020.

Model nie zawsze jest w stanie podzielić zbiór danych liniowo, dlatego może dochodzić do naruszenia marginesu tam, gdzie poszczególne obserwacje nachodzą na granicę decyzyjną lub nawet znajdują się po jej drugiej stronie. Wyróżniamy dlatego klasyfikację marginesu twardego, którego obserwacje z danej klasy nie mogą przekroczyć, oraz miękkiego, która dopuszcza, aby obserwacje przekraczały granicę swojej klasy.

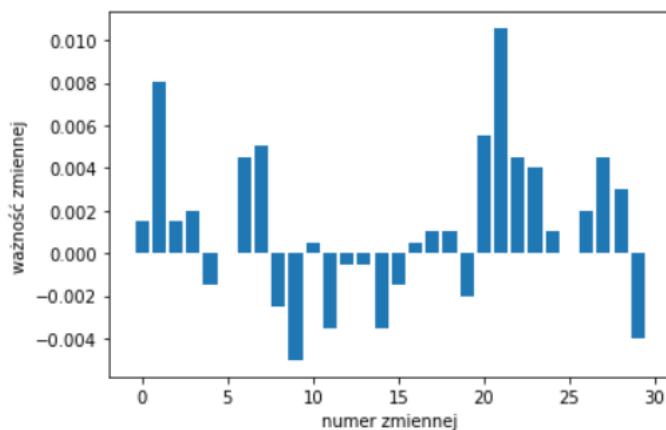
Rozwinięciem liniowego modelu SVM jest model wielomianowy, który zamiast prostej dobiera wielomian  $n$ -tego stopnia i na jego podstawie klasyfikuje dane na klasy analogicznie do liniowego odpowiednika. Rysunek 4 obrazuje przykłady wielomianów, którymi zostały rozdzielone obserwacje. Wykres po lewej stronie jest wielomianem 3 stopnia, a po prawej 10 stopnia.



**Rysunek 4.** Przykłady klasyfikatorów SVM wykorzystujących jądro wielomianowe

Źródło: A. Geron, *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Helion SA, Gliwice 2020.

Na rysunku 5 przedstawiono ważności poszczególnych zmiennych dla klasyfikatora opartego na maszynie wektorów nośnych. W tabeli 5 widzimy, które zmienne zostały przez ten klasyfikator uznane za najbardziej istotne.



**Rysunek 5.** Ważność poszczególnych zmiennych dla modelu maszyny wektorów nośnych

Źródło: opracowanie własne.

W tabeli 6 znajduje się macierz pomyłek dla SVM. Zauważono, że wszystkie przypadki raka niezłośliwego zostały przez model prawidłowo zaklasyfikowane jako niezłośliwe. Natomiast w grupie obserwacji z diagnozą raka złośliwego trzy przypadki zostały błędnie zidentyfikowane jako niezłośliwe. Dokładność uzyskana na podstawie modelu maszyny wektorów nośnych wynosi 98,24%, co jest bardzo dobrym wynikiem na tle innych metod.

**Tabela 5. Najważniejsze cechy badanego zbioru danych wyznaczone z modelu maszyny wektorów nośnych**

Nazwa zmiennej	Ważność zmiennej
Najgorsza tekstura	0,0155
Średnia tekstura	0,00804
Średnia punktów wklęsłych	0,00503
Najgorszy promień	0,00553
Średnia symetria	0,00503

Źródło: opracowanie własne.

**Tabela 6. Macierz pomylek uzyskana na podstawie modelu maszyny wektorów nośnych**

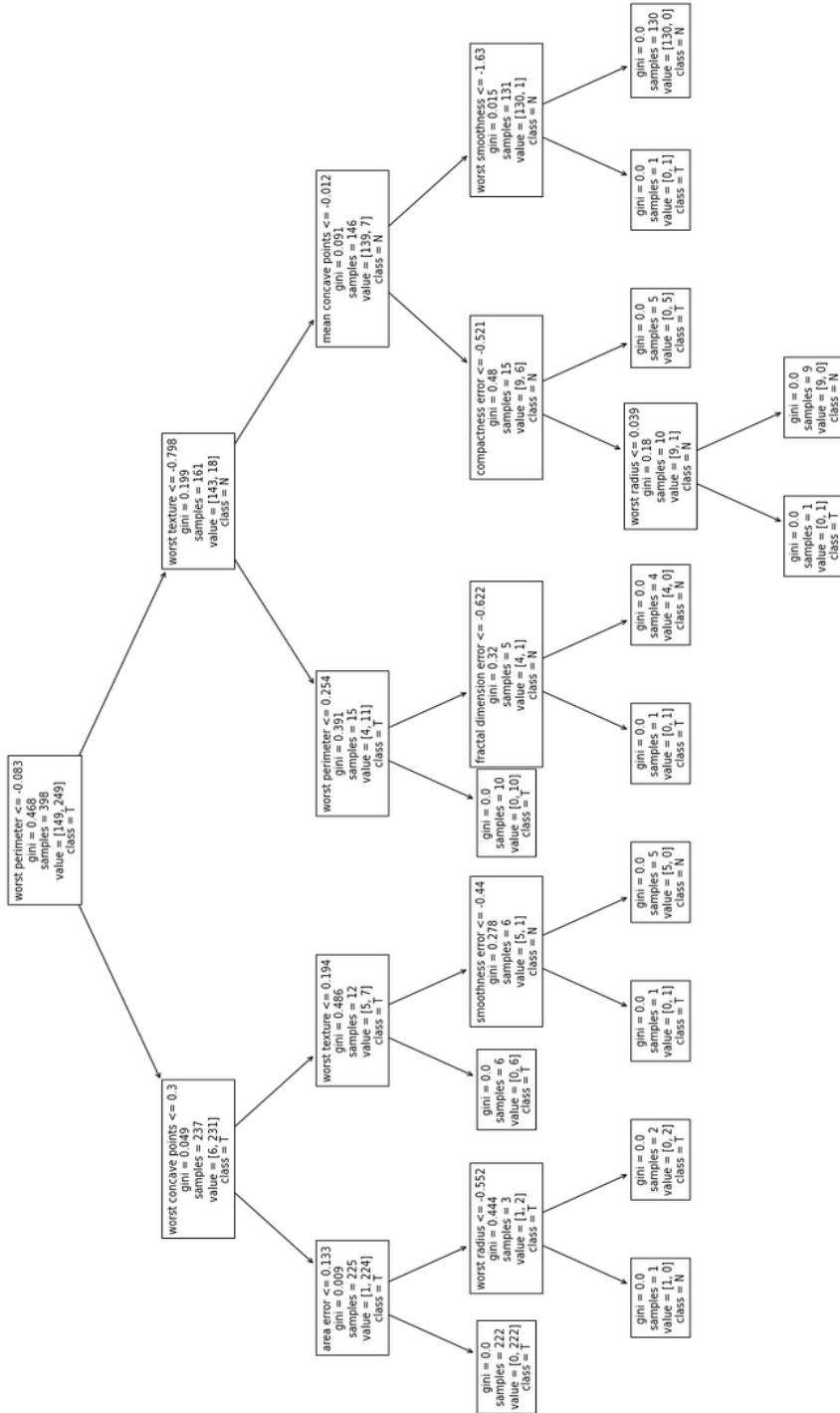
		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	60	3
	Nie	0	108

Źródło: opracowanie własne.

## 6. Drzewo decyzyjne

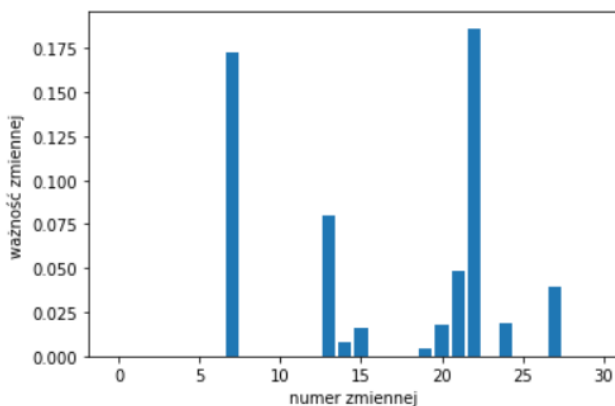
Drzewo decyzyjne (ang. *decision tree*) jest to prosty, lecz użyteczny w uczeniu maszynowym, model zbudowany w oparciu o algorytm drzewa. Najczęściej stosowany do klasyfikacji, czyli przypisania obserwacji zbioru danych do jednej z klas. Każdy węzeł w drzewie decyzyjnym reprezentuje atrybut, a liść reprezentuje wynik. Na każdym węźle drzewa dokonujemy podziału zbioru na mniejsze zbiory, mając na celu jak najlepsze odseparowanie obserwacji należących do różnych klas. Gałęzie, które łączą węzły z liśćmi, to decyzje lub reguły przewidywania. Węzeł główny to atrybut, który najlepiej opisuje zestaw danych szkoleniowych. W ten sposób cały proces jest zobrazowany w strukturze drzewiastej (rys. 6). Wadą drzewa jest to, że niewielka zmiana danych może niekiedy prowadzić do uzyskania całkowicie odmiennego drzewa.

Na rysunku 7 przedstawiono ważność poszczególnych zmiennych zbioru Breast Cancer dla modelu drzewa decyzyjnego, w tabeli 7 zostały wypisane najważniejsze spośród tych zmiennych.



Rysunek 6. Graficzne przedstawienie modelu drzewa decyzyjnego dla badanego zbioru

Źródło: opracowanie własne.



**Rysunek 7. Ważność poszczególnych zmiennych dla modelu drzewa decyzyjnego**

Źródło: opracowanie własne.

**Tabela 7. Najważniejsze cechy badanego zbioru danych wyznaczone z modelu drzewa decyzyjnego**

Nazwa zmiennej	Ważność zmiennej
Najgorszy promień	0,18643
Średnia punktów wklęsłych	0,17286
Błąd pola	0,07990
Najgorsza tekstura	0,04874
Najgorsze punkty wklęsłe	0,03920

Źródło: opracowanie własne.

W tabeli 8 znajduje się macierz pomyłek dla drzewa decyzyjnego. Model poprawnie sklasyfikował wszystkie przypadki raka łagodnego, jednak widać tendencje do predykcji fałszywie negatywnych, czyli klasyfikowania przypadków złośliwych jako łagodnych, co miało miejsce siedem razy. Dokładność uzyskana na podstawie modelu drzewa decyzyjnego wynosi 95,9%.

**Tabela 8. Macierz pomyłek uzyskana na podstawie modelu drzewa decyzyjnego**

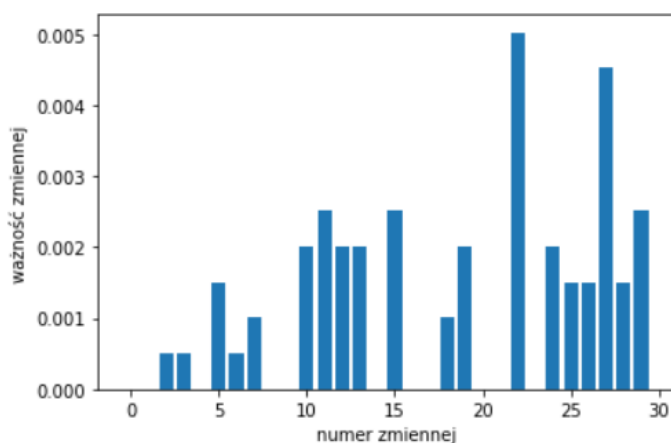
		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	56	7
	Nie	0	108

Źródło: opracowanie własne.

## 7. Las losowy

Las losowy (ang. *random forest*) można postrzegać jako zbiór wielu drzew decyzyjnych z losowym próbkowaniem. Ten algorytm ma na celu wyeliminowanie wad algorytmu drzewa decyzyjnego [10]. Losowe lasy mogą ograniczać niestabilności drzew decyzyjnych oraz ich tendencje do nadmiernego dopasowywania się do zestawu uczącego poprzez uśrednianie prognoz uzyskiwanych z wielu takich drzew [5].

Na rysunku 8 przedstawiono ważność poszczególnych zmiennych dla modelu lasu losowego, a tabela 9, odnosząca się do tego samego rysunku, zawiera zestawienie najważniejszych cech.



Rysunek 8. Ważność poszczególnych zmiennych dla modelu lasu losowego

Źródło: opracowanie własne.

Tabela 9. Najważniejsze cechy badanego zbioru danych wyznaczone z modelu lasu losowego

Nazwa zmiennej	Ważność zmiennej
Najgorszy promień	0,00503
Najgorsze punkty wklęsłe	0,00452
Najgorsza powierzchnia fraktalna	0,00251
Błąd zwartości	0,00251
Błąd promienia	0,00201

Źródło: opracowanie własne.

W tabeli 10 znajduje się macierz pomyłek dla lasu losowego, dokładność uzyskana na podstawie tego modelu wynosi 95,32%. Niestety otrzymany rezultat nie poprawia wyniku drzewa decyzyjnego, jednak widoczne są różnice w macierzy pomyłek.

Tabela 10. Macierz pomylek uzyskana na podstawie modelu lasu losowego

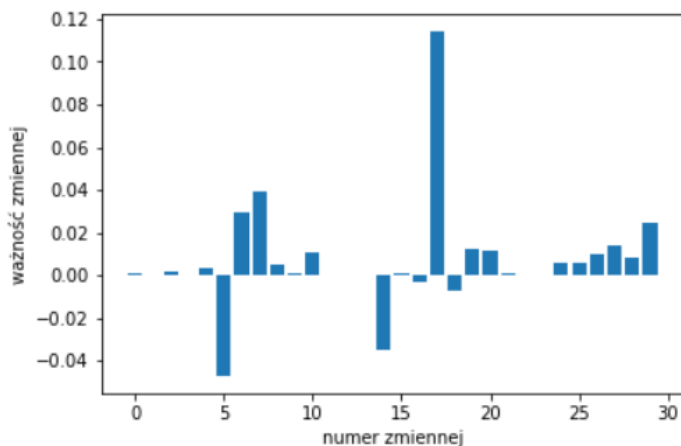
		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	58	5
	Nie	3	105

Źródło: opracowanie własne.

## 8. Liniowa analiza dyskryminacyjna

Liniowa analiza dyskryminacyjna (ang. *linear discriminant analysis*, LDA) jest używana w uczeniu maszynowym do znalezienia liniowej kombinacji cech, które najlepiej rozróżniają dwie lub więcej klas obiektów [9]. Samo zagadnienie dyskryminacji można zdefiniować jako problem znalezienia charakterystyki klas, poprzez identyfikację sposobów ich odseparowania [8]. LDA jest to algorytmem klasyfikacyjnym, ale w trakcie uczenia wykrywa on najbardziej charakterystyczne osie pomiędzy klasami, dzięki którym jesteśmy w stanie zdefiniować hiperpłaszczyzny rzutowania danych. Zaletą tego rozwiązania jest maksymalne możliwe rozdzielenie klas, zatem analiza LDA stanowi dobry sposób redukcji wymiarowości przed uruchomieniem innego algorytmu, np. SVM.

Na rysunku 9 przedstawiono ważność poszczególnych zmiennych dla modelu liniowej analizy dyskryminacyjnej. W tabeli 11 zostały wymienione nazwy najważniejszych cech wraz z wyznaczonymi wartościami ich ważności.



Rysunek 9. Waga poszczególnych zmiennych dla modelu liniowej analizy dyskryminacyjnej

Źródło: opracowanie własne.

**Tabela 11. Najważniejsze cechy badanego zbioru danych wyznaczone z modelu liniowej analizy dyskryminacyjnej**

Nazwa zmiennej	Ważność zmiennej
Błąd punktów wklęsłych	0,11407
Średnia punktów wklęsłych	0,03920
Średnia wklęsłość	0,02915
Najgorszy wymiar fraktalny	0,02412
Najgorsze punkty wklęsłe	0,01357

Źródło: opracowanie własne.

W tabeli 12 znajduje się macierz pomyłek dla liniowej analizy dyskryminacyjnej. Jak widać spośród obserwacji pozytywnych blisko połowa została sklasyfikowana fałszywie negatywnie. Obserwacji fałszywie negatywnych jest stosunkowo mniej, choć ich liczba 30 jest wciąż bardzo duża na tle innych metod. Całkowita dokładność uzyskana na podstawie tego modelu wynosi 65,4%.

**Tabela 12. Macierz pomyłek uzyskana na podstawie modelu liniowej analizy dyskryminacyjnej**

		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	34	29
	Nie	30	78

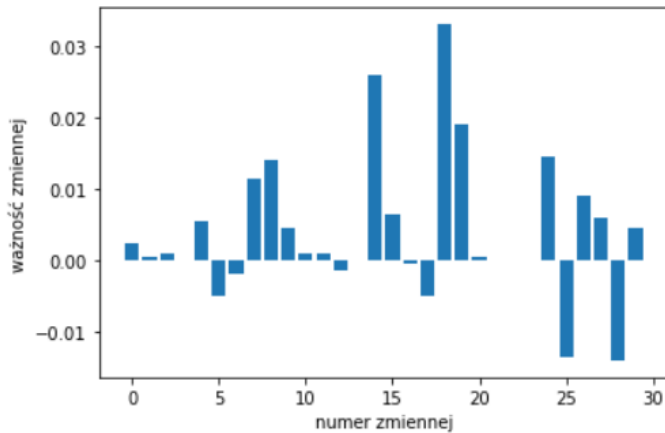
Źródło: opracowanie własne.

## 9. Kwadratowa analiza dyskryminacyjna

Kwadratowa analiza dyskryminacyjna (ang. *quadratic discriminant analysis*, QDA) jest to metoda podobna do LDA, ponieważ obie są technikami klasyfikacji z założeniami Gaussa. LDA wymaga jednak założenia, że macierze kowariancji cech obu klas są takie same, podczas gdy QDA nie wymaga takiego założenia i pozwala na różne macierze kowariancji cech dla różnych klas [2].

Na rysunku 10 przedstawiono ważność poszczególnych zmiennych dla modelu kwadratowej analizy dyskryminacyjnej. W tabeli 13 zostały umieszczone cechy uznane przez model jako najważniejsze.





Rysunek 10. Ważność poszczególnych zmiennych dla modelu kwadratowej analizy dyskryminacyjnej

Źródło: opracowanie własne.

Tabela 13. Najważniejsze cechy badanego zbioru danych wyznaczone z modelu kwadratowej analizy dyskryminacyjnej

Nazwa zmiennej	Ważność zmiennej
Błąd symetrii	0,03317
Błąd gładkości	0,02613
Błąd wymiaru fraktalnego	0,01910
Najgorsza gładkość	0,01457
Średnia symetrii	0,01407

Źródło: opracowanie własne.

W tabeli 14 znajduje się macierz pomyłek dla kwadratowej analizy dyskryminacyjnej. Widać w tym przypadku silną tendencję do klasyfikowania nowej obserwacji jako raka łagodnego. Widzimy 12 klasyfikacji fałszywie dodatnich, jednak bardziej niepokoi fakt, że spośród obserwacji faktycznie będących rakiem złośliwym, niemalże 2/3 zostało sklasyfikowane jako rak łagodny. Całkowita dokładność tego modelu wynosi 69%.

Tabela 14. Macierz pomyłek uzyskana na podstawie modelu kwadratowej analizy dyskryminacyjnej

		Predykcja	
		Tak	Nie
Stan Faktyczny	Tak	22	41
	Nie	12	96

Źródło: opracowanie własne.

## 10. Porównanie modeli i wnioski

W tabeli 15 znajduje się porównanie wyników poszczególnych metod pod kątem dokładności, jaką osiągnięto, budując na ich podstawie model klasyfikacji.

**Tabela 15. Porównanie dokładności otrzymanych dla poszczególnych metod**

Metoda	Dokładność
KNN	95,9%
Regresja logistyczna	97%
Maszyna wektorów nośnych	98,2%
Drzewo decyzyjne	95,9%
Las losowy	95,3%
Liniowa analiza dyskryminacyjna	65,4%
Kwadratowa analiza dyskryminacyjna	69%

Źródło: opracowanie własne.

Jak widać w tabeli 15 większość metod uzyskała dokładność powyżej 95%. Najlepiej ze wszystkich metod w tym przypadku sprawdziła się maszyna wektorów nośnych, dając prawdopodobieństwo poprawnej klasyfikacji 98,2%, na drugim miejscu znalazła się regresja logistyczna, która uzyskała prawdopodobieństwo 97%. Drzewo decyzyjne i KNN uzyskały identyczny wynik 95,9%. Co ciekawe, las losowy wypadł gorzej niż drzewo decyzyjne z wynikiem 95,3%, może się to wydawać dziwne, ponieważ las losowy, korzystając z wielu drzew decyzyjnych, powinien w założeniu uzyskiwać lepsze rezultaty, co jednak nie sprawdza się w tym przypadku.

Najniższe wartości zostały otrzymane dla liniowej analizy dyskryminacyjnej oraz kwadratowej analizy dyskryminacyjnej. Kwadratowa analiza uzyskuje o niecałe 4% wyższy wynik niż jej liniowy odpowiednik, jednak prawdopodobieństwo poprawnej klasyfikacji dla żadnej z tych dwóch metod nie przekroczyło wartości 70%.

W kwestii ważności zmiennych nie otrzymano jednoznacznego wyniku, a każdy model wypracowywał odmienne podejście. Można jednak wypisać kilka cech, które wielokrotnie były w procesie uczenia modelu uznawane za wysoko istotne, są to: najgorszy wymiar fraktalny, średnia punktów wklęsłych, najgorsze punkty wklęsłe, najgorsza tekstura oraz najgorszy promień. Widać zależność, że przeważnie najczęściej cechami ważnymi były te ze słowem „najgorszy”, które opisują najbardziej niepożądane dla zdrowia pacjenta parametry.

## Literatura

- [1] Al Duhayyim M., Malibari A. A., Dhahbi S., Nour, M. K., Al-Turaiki I., Obayya M., Mohamed A., *Sailfish optimization with deep learning based oral cancer classification model*, “Computer Systems Science and Engineering” 2023, vol. 45 (1), pp. 753–767.
- [2] Ghodsi A., *Lec 2: Machine learning. classification, Linear and quadratic discriminant analysis*, [https://www.youtube.com/watch?v=\\_m7TMkzZzus&t=1856s](https://www.youtube.com/watch?v=_m7TMkzZzus&t=1856s) (dostęp: 05.11.2022).
- [3] Education-WIKI.com, *Algorytm lasu losowego*, <https://pl.education-wiki.com/2890284-random-forest-algorithm> (dostęp: 05.11.2022).
- [4] UCI Machine Learning, *Breast Cancer Wisconsin, (Original) Data Set*, <https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data> (dostęp: 05.11.2022).
- [5] Fisher R.A., *Iris data set*, <https://archive.ics.uci.edu/ml/datasets/Iris/> (dostęp: 05.11.2022).
- [6] Gatnar E., *Analiza dyskryminacyjna – stan aktualny i kierunki rozwoju*, “Studia Ekonomiczne” 2013, nr 152, ss. 42–58.
- [7] IBM, *Interpreting logistic regression, IBM.Logistic Regression*, <https://www.ibm.com/topics/logistic-regression> (dostęp: 05.11.2022).
- [8] Geron A., *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Wydanie II, Helion, Gliwice, 2020.
- [9] Tharwat A., Gaber T., Ibrahim A., Hassanien A. E., *Linear discriminant analysis: A detailed tutorial*, “AI Communications” 2017, vol. 30, iss. 2, pp. 169–190.
- [10] Cutler A., Cutler D. R., Stevens J. R., *Random Forests*, [in:] eds. C. Zhang, Y. Ma, *Ensemble Machine Learning*, Springer, New York 2012, pp. 157–175.
- [11] Bray F., Ferlay J., Soerjomataram I., Siegel R. L., Torre L. A., Jemal A., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*, “CA: A Cancer Journal for Clinicians” 2018, vol. 68, iss. 6, pp. 394–424.
- [12] Evgeniou T., Pontil M., *Support Vector Machines: Theory and Applications*, [in:] eds. G. Paliouras, V. Karkaletsis, C. D. Spyropoulos, *Machine Learning and Its Applications*, Springer Berlin Heidelberg, 2001, pp. 249–257.
- [13] Tayel M. B., Mokhtar M. A., Kishk A. F., *Breast cancer diagnosis using histopathology and convolution neural network CNN method*, 2023.

---

## Selected classification methods in machine learning tasks

### Abstract

This chapter provides an overview of popular machine learning methods used in the data classification process. The presented methods include decision tree, random forest, logistic regression, discriminant analysis, support vector machine and k-nearest neighbors method. In addition to the discussion of the above methods, a comparison was also made in a practical classification task. The dataset used for the analysis is called "Breast Cancer Wisconsin". The purpose of the analysis is to determine whether the cancer is malignant based on the size of the tumor and other parameters related to the neoplastic changes. Machine learning models can also allow for the extraction of the most important features affecting the determination of the type of cancer.

*Keywords: classification, decision tree, random forest, SVC, KNN, logistic regression, discriminant analysis.*



**Alicja Rachwał<sup>1</sup>**

## **Autoenkodery w nienadzorowanym uczeniu maszynowym<sup>2</sup>**

### **Streszczenie**

Niniejszy rozdział dokonuje przeglądu zastosowań autoenkoderów stosowanych w nienadzorowanym uczeniu maszynowym. Przedstawione zostało pojęcie autoenkodera i jego rodzaje, a wybrane z rodzajów zostały opisane bardziej szczegółowo. Pokazanych zostało sześć przykładów użycia autoenkoderów w uczeniu nienadzorowanym. Zastosowania zaprezentowano na przykładzie zbioru MNIST składającego się z czarno-białych obrazów ręcznie pisanych cyfr. Pokazano działanie autoenkodera prostego i splotowego, a następnie ich użycie do odsumiania obrazu, wyszukiwania obrazu, wykrywania anomalii i generowania nowych obrazów.

*Słowa kluczowe: autoenkodery, sztuczne sieci neuronowe, przetwarzanie obrazu, uczenie nienadzorowane.*

### **Wstęp**

Jeszcze niedawno komputery nie były w stanie wykonywać wielu zadań, które dla człowieka wydają się trywialne, jak na przykład rozpoznanie obiektu znajdującego się na zdjęciu. W ostatnich latach następuje jednak coraz większy rozwój w dziedzinie sztucznych sieci neuronowych, inspirowanych do pewnego stopnia biologicznym mózgiem człowieka. Zainteresowanie sztucznymi sieciami neuronowymi jest potęgowane przez dostępność ogromnej ilości danych uczących oraz znaczny wzrost mocy obliczeniowej urządzeń dostępnych dla użytkowników. Algorytmy uczące sieci neuronowe są stale udoskonalane, a pewne teoretyczne ograniczenia sztucznych sieci okazały się w praktyce nieistotne.

Kluczowym rodzajem sieci neuronowych przedstawionym w tym artykule są autoenkodery. Stanowią one formę uczenia nienadzorowanego. Są szkolone po to, aby kopiować swoje dane wejściowe na wyjściu, przy zachowaniu pewnych ograniczeń, tak aby w procesie uczenia poznać istotne informacje o rozkładzie danych. Autoenkodery uczą się tzw. reprezentacji ukrytych, które zazwyczaj mają znacznie mniejszy rozmiar niż dane wejściowe. Dzięki temu mogą służyć do re-

---

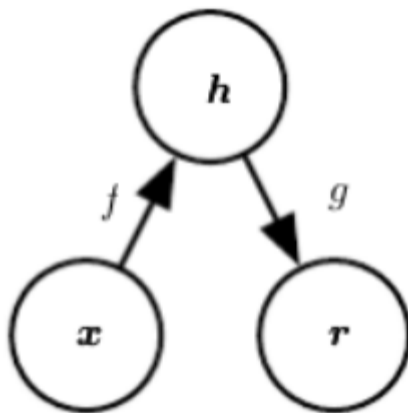
<sup>1</sup> Mgr inż. Alicja Rachwał, Wydział Elektrotechniki i Informatyki, Politechnika Lubelska.

<sup>2</sup> Niniejszy tekst stanowi fragment pracy dyplomowej pod tytułem *Przegląd autoenkoderów stosowanych w nienadzorowanym uczeniu maszynowym*, obronionej w dn. 14 lipca 2022 na Wydziale Podstaw Techniki Politechniki Lubelskiej.

dukowania wymiarowości, zwłaszcza w zadaniach wizualizacji. Potrafią również generować nowe dane przypominające te ze zbioru uczącego. Niniejszy rozdział przedstawi zagadnienia związane z autoenkoderami, w szczególności skupiając się na ich zastosowaniach w nienadzorowanym uczeniu maszynowym.

## 1. Autoenkoder i jego wybrane rodzaje

Autoenkoder (inaczej autokoder, z ang. *autoencoder*, *auto-encoder*) jest rodzajem sieci neuronowej przeznaczonym głównie do kodowania danych wejściowych do skompresowanej i znaczącej reprezentacji, a następnie dekodowania ich z powrotem w taki sposób, aby zrekonstruowane dane były jak najbardziej podobne do oryginalnych [2]. Autoenkodery uczą się gęstych reprezentacji danych, tzw. **reprezentacji ukrytych** (ang. *latent representations*) lub kodowań (ang. *codings*) w formie uczenia nienadzorowanego. Kodowania często mają mniejszą wymiarowość od danych wejściowych, dzięki czemu autoenkodery mogą służyć do redukcji wymiarowości. Mają też zastosowanie w modelach generatywnych (ang. *generative models*), które potrafią losowo generować nowe dane przypominające zbiór uczący [3]. Autoenkoder jest zatem siecią neuronową szkoloną po to, aby kopiować dane wejściowe do wyjścia. Zawiera ukrytą warstwę  $h$ , która opisuje kodowanie używane do reprezentowania wejścia. Sieć można postrzegać jako składającą się z dwóch części: kodującej funkcji  $h(x)$  i dekodera, który tworzy rekonstrukcję  $r = g(h)$  [4]. Ogólna struktura autoenkodera jest przedstawiona na rysunku 1.



**Rysunek 1.** Struktura autoenkodera odwzorowującego wejście  $x$  na wyjście  $r$  (nazywane rekonstrukcją) poprzez reprezentację ukrytą (kodowanie)  $h$ . Autoenkoder składa się z dwóch składników: kodera  $f$  (odwzorowującego  $x$  na  $h$ ) i dekodera  $g$  (odwzorowującego  $h$  na  $r$ )

Źródło: I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning. Systemy uczące się*, PWN, Warszawa 2018.

## 1.1. Rodzaje autoenkoderów

Gdyby autoenkoder nauczył się po prostu, aby na wyjściu ustawiać zawsze  $g(g(x)) = x$ , to nie byłby zbyt przydatny. Z tego powodu autoenkodery są projektowane tak, aby nie potrafiły kopiować w sposób doskonały. Zazwyczaj nakładane są ograniczenia, aby autoenkoder mógł kopiować jedynie w przybliżeniu i tylko takie dane, które są podobne do danych ze zbioru uczącego. Dzięki temu model musi wybierać jedynie pewne aspekty danych wejściowych, które powinny być kopiowane. W ten sposób może on wydobyć użyteczne informacje o strukturze danych. Ze względu na nakładane ograniczenia, wyróżniamy wiele rodzajów autoenkoderów – w kolejnych podrozdziałach zostaną opisane wybrane z nich.

## 1.2. Autoenkodery niedopelnione

Jednym ze sposobów, aby uzyskać przydatne cechy z autoenkodera, jest ograniczenie  $h$  do mniejszego wymiaru niż  $x$ . Autoenkoder, w którym wymiar kodu jest mniejszy niż wymiar wejściowy, jest nazywany niekompletnym (niedopełnionym). Poznawanie niekompletnych reprezentacji zmusza autoenkoder do przechwycenia najistotniejszych cech danych szkoleniowych. Proces poznawania jest opisywany jako minimalizowanie funkcji straty  $L(x, g(f(x)))$ , gdzie  $L$  jest funkcją straty z karą dla  $g(f(x))$  za niepodobieństwo do  $x$ , jak np. błąd średniokwadratowy.

## 1.3. Autoenkodery z regularyzacją

Regularyzacja jest stosowana głównie w autoenkoderach nadkompletnych, tzn. takich, w których reprezentacja ukryta ma większy wymiar niż wejście. Zamiast ograniczać pojemność modelu przez zachowywanie płytkości kodera i dekodera oraz małego rozmiaru kodu, autoenkodery z regularyzacją używają funkcji straty, dzięki której model może posiadać inne właściwości oprócz możliwości kopiowania swojego wejścia do wyjścia. Do tych właściwości należą m.in. rzadkość reprezentacji oraz odporność na szum lub brakujące dane wejściowe [4].

### Rzadkie autoenkodery

Rzadkie autoenkodery są zwykle używane do tego, aby uczyć się cech do innego zadania, takiego jak klasyfikacja. Autoenkoder, który dzięki regularyzacji jest rzadki, musi reagować na unikatowe statystyczne cechy zbioru danych, na jakim został wyszkolony, a nie tylko działać jak funkcja tożsamościowa.

Kryterium szkolenia autoenkodera rzadkiego obejmuje karę rzadkości  $\Omega(h)$  na warstwie kodu  $h$  oprócz błędu rekonstrukcji  $L(x, g(h)) + \Omega(h)$ , gdzie  $g(h)$  to wyjście dekodera, a zwykle mamy  $h = f(x)$ , czyli wyjście kodera [4]. Dodanie tego składnika do funkcji kosztu zmusza autoenkoder do zmniejszenia liczby aktywnych neuronów w warstwie kodowania. W ten sposób każde wejście musi być reprezentowane jako kombinacja niewielkiej liczby pobudzeń. Dzięki temu



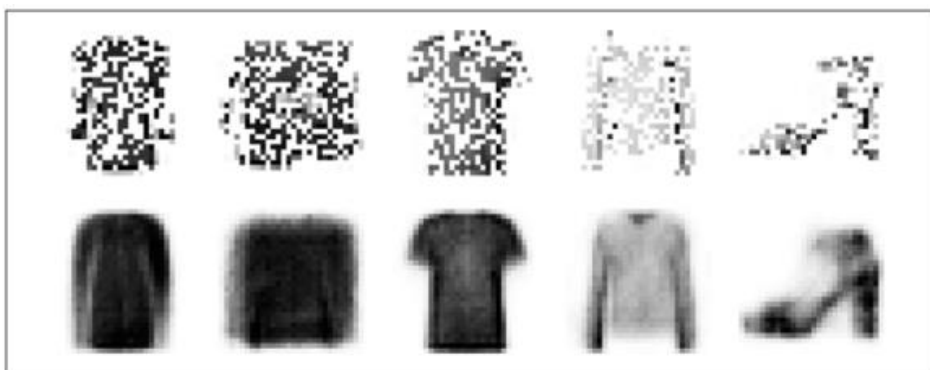
każdy neuron warstwy kodowania zazwyczaj uczy się wykrywać jakąś przydatną cechę [3].

### Autoenkodery z odszumianiem

Zamiast dodawać karę  $\Omega$  do funkcji kosztów, możemy uzyskać autoenkoder, który uczy się czegoś przydatnego, zmieniając składnik błędu rekonstrukcji w funkcji kosztów. Tradycyjne autoenkodery minimalizują jakąś funkcję  $L(x, g(f(x)))$ , gdzie  $L$  to funkcja straty z karą dla  $g(f(x))$  za niepodobieństwo do  $x$ , jak np. norma  $L^2$  ich różnicy. Sprzyja to temu, aby  $g \circ f$  uczyła się być jedynie funkcją tożsamościową, jeśli ma do tego odpowiednią pojemność.

Autoenkoder z odszumianiem zamiast tego minimalizuje  $L(x, g(f(\tilde{x})))$ , gdzie  $\tilde{x}$  to kopia  $x$ , która została zniekształcona przez jakiegoś rodzaju postać szumu. Autoenkodery mają tym samym za zadanie odwrócić to zniekształcenie, a nie po prostu przekopiarować swoje wejście [4].

Zniekształcenie może być szumem gaussowskim dodawanym do danych wejściowych lub może przybrać postać losowo wyłączanych wejść za pomocą metody porzucania. Rysunek 2 przedstawia przykłady zaszumionych obrazów (połowa pikseli została „wyłączona”), a także ich rekonstrukcje uzyskane za pomocą autokodera odszumiającego. Autokoder „odgaduje” szczegóły niewystępujące w obrazach wejściowych, na przykład górną część białej sukienki (czwarty obraz w dolnym rzędzie). Wejściem jest zatem zanieczyszczony obraz, a docelową wartością jest oryginał bez zniekształcenia. Sieć uczy się rozpoznawać i usuwać zniekształcenia, aby wygenerować rekonstrukcję. Autoenkoder nie widzi oryginalnego, czystego obrazu, jedynie ten ze zniekształceniem.



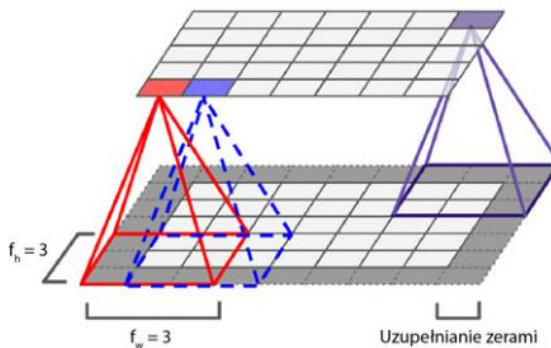
Rysunek 2. Zaszumione obrazy (na górze) i ich rekonstrukcje (na dole)

Źródło: A. Géron, *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Helion SA, Gliwice 2020.

## 1.4. Autoenkodery plotowe

Autoenkodery plotowe są typem autoenkoderów, który najlepiej nadaje się do przetwarzania obrazów. W takim rodzaju autoenkodera, kodem jest sieć konwolucyjna składająca się z warstw plotowych i redukujących. Zwykle zmniejsza ona wymiarowość obrazu (wysokość i szerokość), a zwiększa głębokość (liczbę map cech). Dekoder przeprowadza operację odwrotną: musi zwiększyć rozdzielczość i zredukować głębokość do pierwotnego wymiaru. W tym celu można wykorzystać transponowane warstwy plotowe lub łączyć warstwy ekspansji (*UpSampling*) ze zwykłymi warstwami plotowymi [3].

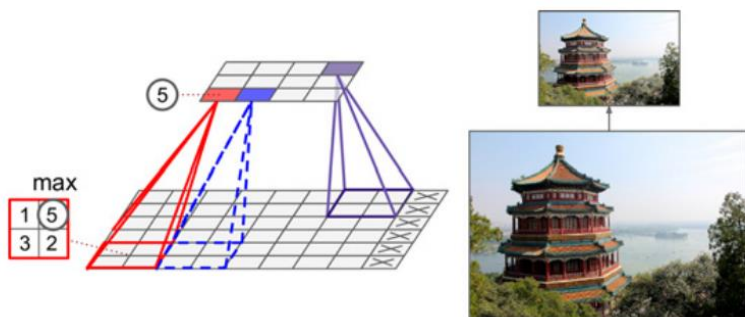
**Warstwa plotowa** to taka warstwa, w której dany neuron jest połączony z wyjściami neuronów poprzedniej warstwy zlokalizowanymi w polu recepcyjnym o wysokości  $f_h$  i szerokości  $f_w$  (rys. 3). W celu uzyskania takich samych wymiarów każdej warstwy najczęściej są dodawane zera wokół wejść. Proces ten nazywamy uzupełnianiem zerami (ang. *zero padding*). Odległość pomiędzy dwoma kolejnymi polami recepcyjnymi nosi nazwę kroku (ang. *stride*).



Rysunek 3. Warstwy plotowe z uzupełnianiem zerami

Źródło: A. Géron, *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Helion SA, Gliwice 2020.

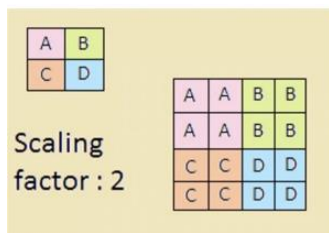
**Warstwa redukująca** zmniejsza przestrzenny wymiar mapy aktywacji, pozostawiając jej głębokość bez zmian [5]. Warstwa łącząca nie zawiera żadnych wag; jej jedynym zadaniem jest gromadzenie danych wejściowych za pomocą jakiejś funkcji agregacyjnej, np. maksymalizującej lub uśredniającej. Na rysunku 4 przedstawiony jest najpopularniejszy rodzaj warstwy łączącej – **maksymalizująca warstwa łącząca** (ang. *max pooling layer*).



**Rysunek 4. Maksymalizująca warstwa łącząca (jądro łączące:  $2 \times 2$ , krok: 2, brak uzupełniania zerami)**

Źródło: A. Géron, *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Helion SA, Gliwice 2020.

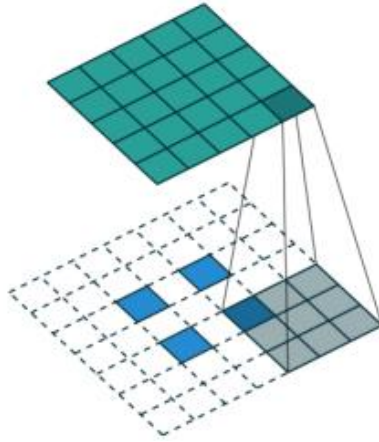
**Warstwa ekspansji** wykonuje nadpróbkowanie poprzez powtórzenie każdej wartości  $m \times n$  razy, gdzie  $m$  to współczynnik nadpróbkowania dla wierszy, a  $n$  to współczynnik nadpróbkowania dla kolumn. Przykład działania warstwy ekspansji widać na rysunku 5.



**Rysunek 5. Działanie warstwy nadpróbkowania (ekspansji)**

Źródło: Artificial Intelligence in Plain English, *Convolutional Autoencoders (CAE) with Tensorflow*, <https://ai.plainenglish.io/convolutional-autoencoders-cae-with-tensorflow-97e8d8859cbe> (dostęp: 15.05.2022).

O **transponowanych warstwach splotowych** można powiedzieć, że zachowują się jak zwykle warstwy splotowe z ułamkową długością kroku, tzn. stosują filtr do obszaru, który jest mniejszy niż rozmiar filtru. W ten sposób transponowane warstwy konwolucyjne wykonują operację w przeciwnym kierunku niż zwykle warstwy konwolucyjne [6]. Działanie transponowanej warstwy splotowej widać na rysunku 6.



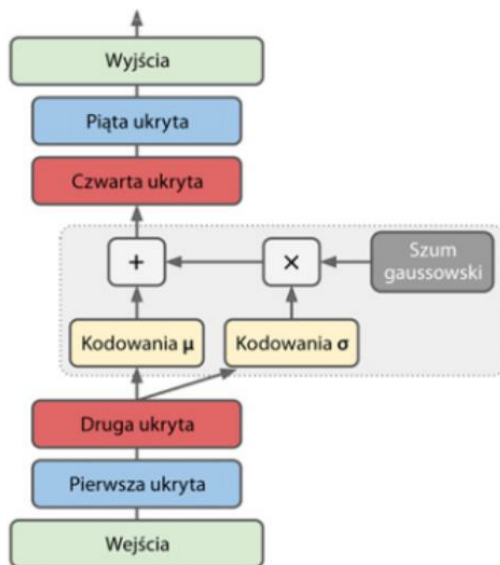
**Rysunek 6. Działanie transponowanej warstwy splotowej**

Źródło: P. L. Pröve, *An Introduction to different Types of Convolutions in Deep Learning*, <https://towardsdatascience.com/types-of-convolutions-in-deep-learning-717013397f4d> (dostęp:15.05.2022).

## 1.5. Autoenkodery wariancyjne

Autoenkodery wariancyjne istotnie różnią się od opisywanych wcześniej rodzajów autoenkoderów. Są one autoenkoderami probabilistycznymi, czyli generują częściowo losowe wyniki. Stanowią klasę modeli generatywnych, co oznacza, że są w stanie tworzyć nowe dane przypominające te ze zbioru uczącego.

Na rysunku 7 przedstawiona jest przykładowa struktura autoenkodera wariancyjnego. Można zauważyć tutaj elementy podstawowej architektury autoenkoderów: koder i dekodek składają się z dwóch warstw ukrytych. Mamy też do czynienia z pewną modyfikacją: koder nie generuje bezpośredniego kodowania próbki wejściowej, lecz uśrednione kodowanie  $\mu$  oraz odchylenie standardowe  $\sigma$ . Rzeczywiste kodowanie jest następnie losowane z rozkładu normalnego o parametrach właśnie  $\mu$  i  $\sigma$ . Następnie dekodek w standardowy sposób dekoduje wylosowane kodowanie. W czasie uczenia funkcja kosztu zmusza kodowania do stopniowego poruszania się po przestrzeni kodowania (nazywanej również przestrzenią ukrytą, z angielskiego *latent space*) w poszukiwaniu miejsca wewnątrz obszaru przypominającego chmurę punktów gaussowskich [3].



**Rysunek 7. Przykładowa struktura autoenkodera wariacyjnego**

Źródło: A. Géron, *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Helion SA, Gliwice 2020.

Funkcja kosztu autoenkodera wariacyjnego składa się z dwóch członów. Pierwszy jest tradycyjną funkcją straty rekonstrukcji, która zmusza autoenkoder do rekonstruowania danych wejściowych. Drugi element nazywany jest funkcją straty ukrytej (ang. *latent loss*). Sprawia on, że autoenkoder uzyskuje reprezentacje przypominające te uzyskiwane z rozkładu normalnego. Stosujemy w tym celu dywergencję Kullbacka–Leiblera [3] pomiędzy docelowym rozkładem (normalnym) a rzeczywistym rozkładem kodowań. Funkcja straty ukrytej wygląda wówczas następująco:

$$-\frac{1}{2} \sum_{i=1}^n [1 + \log(\sigma_i^2) - \sigma_i^2 - \mu_i^2], \quad (1)$$

gdzie  $n$  – wymiarowość kodowań,  $\mu_i$  i  $\sigma_i$  to średnia i odchylenie standardowe  $i$  tej składowej kodowania. Koder generuje na wyjściu wektory  $\mu$  i  $\sigma$ , przechowujące wszystkie wartości  $\mu_i$  i  $\sigma_i$ .

Warto zauważyć, że po wytrenowaniu autoenkodera wariacyjnego generowanie nowych próbek jest bardzo łatwe – wystarczy wylosować kodowanie z rozkładu normalnego, a następnie rozkodować je przy użyciu dekodera [3].

## 2. Przykłady zastosowań autoenkoderów

Zbiorem, na przykładzie którego będziemy pokazywać zastosowania autoenkoderów, jest zbiór danych MNIST [7]. Zbiór ten znajduje szerokie zastosowanie w trenowaniu i testowaniu modeli uczenia maszynowego. Często stanowi punkt odniesienia do porównania różnego rodzaju modeli [8]. Przykładowe dane z części treningowej tego zbioru są przedstawione na rysunku 8. W zbiorze uczącym znajdują się 60 tysięcy obserwacji, a w testowym 10 tysięcy.

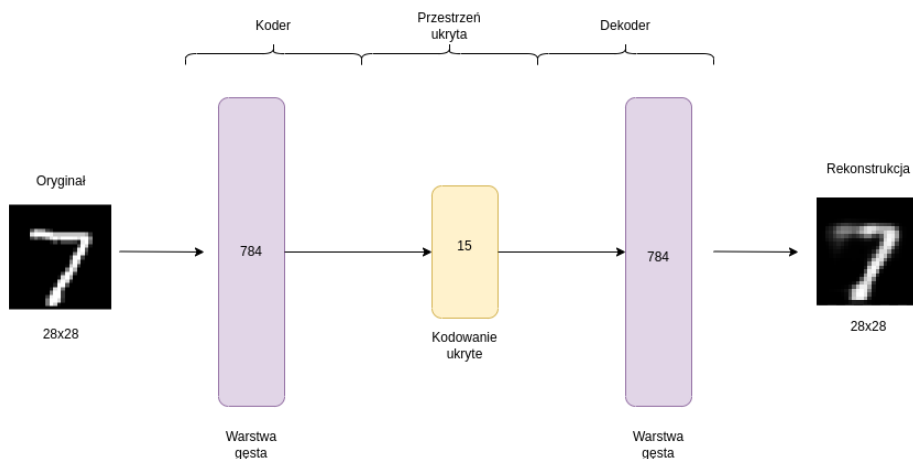


Rysunek 8. Pięć pierwszych obserwacji ze zbioru cyfr MNIST

Źródło: opracowanie własne.

### 2.1. Prosty autoenkoder

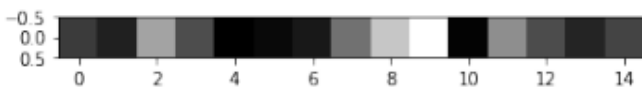
Autoenkoder prosty składa się on z kodera i dekodera, z których każdy jest pojedynczą warstwą gęstą. Warstwa wyjściowa będzie przyjmować obrazy „spłaszczone” do wektora o długości 784 (obrazy 2D były wymiaru  $28 \times 28$ ). Reprezentacja ukryta ma wymiar 15, jest to więc autoenkoder niedopełniony (kodowania mają mniejszy wymiar niż dane wejściowe). Następnie dekodery przywraca wejściowy wymiar 784. Struktura użytego autoenkodera jest przedstawiona na rysunku 9.



Rysunek 9. Struktura autoenkodera prostego

Źródło: opracowanie własne.

Rysunek 10 przedstawia kodowanie w ukrytej przestrzeni piętnastowymiarowej dla pierwszej obserwacji ze zbioru testowego. Rysunek 11 przedstawia pięć pierwszych obrazów ze zbioru testowego (górnny rząd) oraz ich rekonstrukcje po odkodowaniu przez dekodery (dolny rząd). Można zauważyć, że po przejściu przez autoenkoder niedopełniony, obrazy znacznie tracą na jakości.



**Rysunek 10. Reprezentacja 15-wymiarowa dla pierwszej obserwacji ze zbioru testowego MNIST**

Źródło: opracowanie własne.



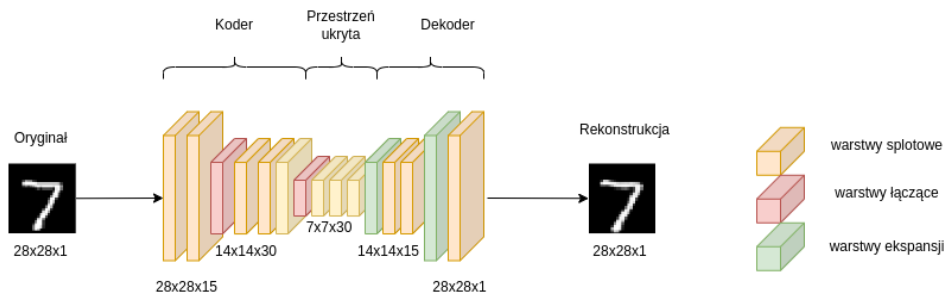
**Rysunek 11. Pięć pierwszych obserwacji ze zbioru testowego MNIST: w górnym rzędzie oryginalne obrazy, w dolnym rekonstrukcje z autoenkodera niedopełnionego**

Źródło: opracowanie własne.

## 2.2. Autoenkoder splotowy

W przypadku danych ze zbioru MNIST obrazy są dość małe, więc nawet prosty autoenkoder dał dość dobre wyniki. Jednak w zastosowaniu do danych wejściowych będących obrazami, zamiast autoenkodera prostego można użyć autoenkodera splotowego, który jest przeznaczony przede wszystkim do takiego typu danych.

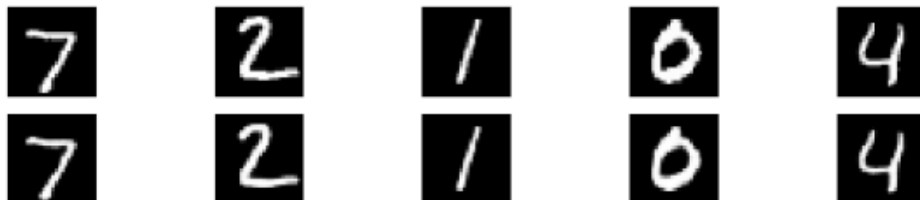
Rysunek 12 przedstawia strukturę użytego autoenkodera splotowego. Koder składa się z dwóch warstw splotowych i dwóch warstw łączących. Warstwy łączące zmniejszają wymiary obrazu z  $28 \times 28$  kolejno do  $14 \times 14$  i  $7 \times 7$ . Dekoder składa się z dwóch warstw splotowych i dwóch warstw ekspansji, które zwiększają wymiary skompresowanego obrazu kolejno do  $14 \times 14$  i  $28 \times 28$ , osiągając oryginalny wymiar. W przypadku warstw splotowych, jądro jest wymiaru  $3 \times 3$ . Warstwy 1 i 4 mają po 15 filtrów, a warstwy 2 i 3 po 30 filtrów. Warstwa wyjściowa również jest warstwą splotową. Ma ona jeden filtr. W warstwach łączących stosowana jest metoda redukcowania *MaxPooling*, gdzie jądro łączące ma rozmiar  $2 \times 2$ . W warstwach ekspansji współczynnik nadpróbkowania jest równy 2 dla wierszy i tyle samo dla kolumn.



**Rysunek 12. Struktura autoenkodera splotowego**

Źródło: opracowanie własne.

Na rysunku 13 widać efekty działania autoenkodera splotowego. W górnym rzędzie przedstawione są oryginalne obrazy ze zbioru testowego MNIST, a w dolnym rzędzie ich rekonstrukcje z autoenkodera. Widać, że obrazy w obu rzędach są do siebie podobne – strata jakości jest mniejsza niż w przypadku autoenkodera niedopełnionego.



**Rysunek 13. Pięć pierwszych obserwacji ze zbioru testowego MNIST: w górnym rzędzie oryginalne obrazy, w dolnym rekonstrukcje z autoenkodera splotowego**

Źródło: opracowanie własne.

### 2.3. Autoenkoder odsumiający

Pokażemy teraz zastosowanie autoenkodera splotowego do odsumiania zanieczyszczonych obrazów. W poprzednich przykładach zbiór uczący stanowił zarówno dane wejściowe, jak i oczekiwane dane wyjściowe w modelu. W przypadku autoenkodera odsumiającego zbiór uczący również jest traktowany jako docelowy wynik, ale danymi wejściowymi są zaszumione obserwacje z tego zbioru.



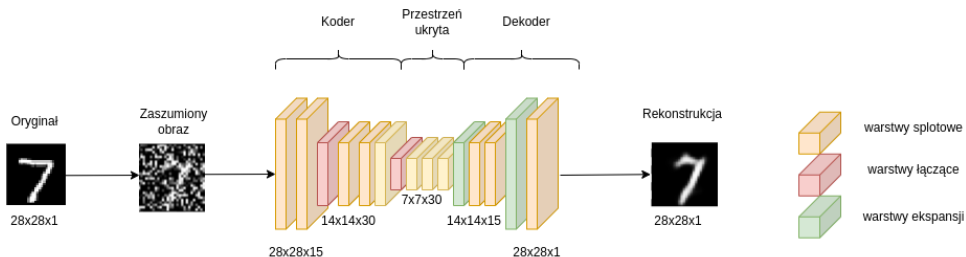
Rysunek 14 przedstawia obrazy ze zbioru testowego MNIST po zaszumieniu. Zaszumienie polegało na dodaniu do wartości każdego piksela losowej liczby z rozkładu  $N(0,1)$ , pomnożonej przez 0,7. Parametr 0,7 został przyjęty arbitralnie i można go zmieniać w celu dalszych eksperymentów, jednakże przy większej jego wartości rekonstrukcje były już znacznie mniej czytelne przy użytym autoenkoderze.



**Rysunek 14. Zaszumione obrazy ze zbioru testowego MNIST**

Źródło: opracowanie własne.

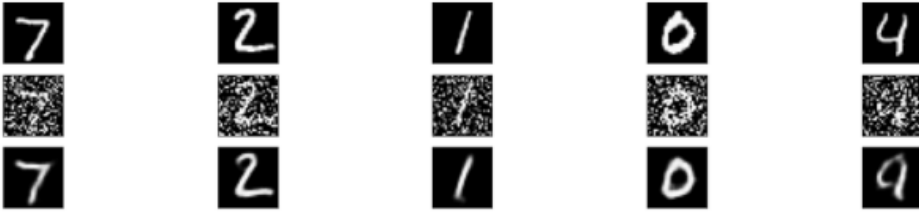
Struktura autoenkodera odszumiającego jest taka sama jak autoenkodera spłotowego (rys. 12). Różnica polega jedynie na tym, że w tym przykładzie danymi wejściowymi są zaszumione obrazy, co zostało pokazane na rysunku 15.



**Rysunek 15. Struktura autoenkodera odszumiającego**

Źródło: opracowanie własne.

Rysunek 16 przedstawia efekt działania autoenkodera odszumiającego. W górnym rzędzie widać oryginalne, niezaszumione obrazy ze zbioru testowego. Środkowy rząd przedstawia zanieczyszczone obrazy, które stanowią dane wejściowe dla autoenkodera. W dolnym rzędzie widać obrazy bez szumu, zrekonstruowane przez autoenkoder. Widać, że w większości autoenkoder dobrze radzi sobie z usuwaniem zanieczyszczeń – jedynie na piątym obrazie widać w jego rekonstrukcji wyraźną pozostałość po szumie.



Rysunek 16. Pięć pierwszych obserwacji ze zbioru testowego MNIST: w górnym rzędzie oryginalne obrazy, w środkowym zaszumione, w dolnym rekonstrukcje z autoenkodera odsumiającego

Źródło: opracowanie własne.

## 2.4. Wyszukiwanie obrazu

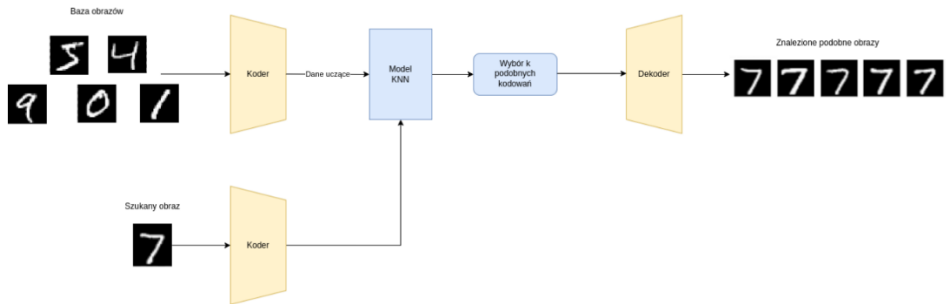
Do wyszukiwania obrazu zostanie użyty autoenkoder prosty opisany w podrozdziale 2.1. Elementy ze zbioru treningowego zostały przy użyciu kodera skompresowane do wymiaru 15. Dla skompresowanych danych budowany jest model  $k$  najbliższych sąsiadów, gdzie parametr  $k$  oznacza również liczbę podobnych obrazów do zadanego, jakie zostaną wyszukane. Zbiorem testowym są w tym przypadku obrazy, do których chcemy wyszukać obrazy podobne. Użyty zbiór testowy jest przedstawiony na rysunku 17. Dla każdego z tych obrazów chcemy wyszukać  $k = 5$  obrazów podobnych.



Rysunek 17. Zbiór testowy dla zadania wyszukiwania obrazu

Źródło: opracowanie własne.

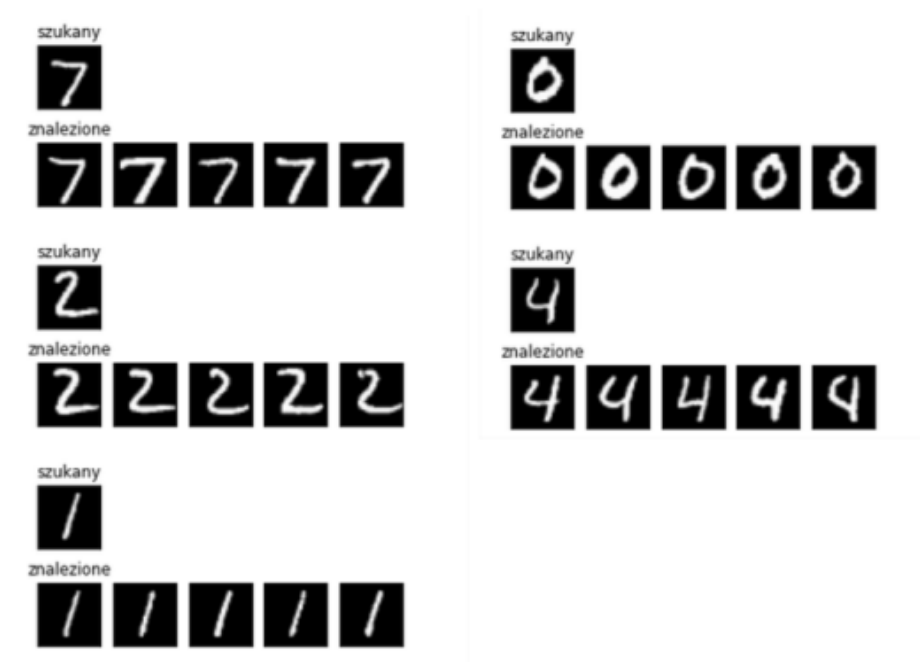
Obrazy ze zbioru testowego również są kompresowane przy użyciu kodera. Następnie dla skompresowanej formy obrazów znajdujemy  $k = 5$  najbliższych sąsiadów według modelu KNN. Jako wynik wyświetlane są oryginalne wersje znalezionych obrazów. Schemat tego zadania jest przedstawiony na rysunku 18.



Rysunek 18. Schemat wyszukiwania podobnych obrazów za pomocą autoenkodera

Źródło: opracowanie własne.

Efekt zastosowania autoenkodera oraz modelu  $k$  najbliższych sąsiadów dla opisanego zbioru testowego znajduje się na rysunku 19. Widać, że dla każdego elementu ze zbioru testowego, znalezione obrazy zawierają odpowiednią cyfrę, zgodną z wyszukiwaniem (np. dla pierwszego przykładu z liczbą 7, wszystkie znalezione obrazy również przedstawiają tą cyfrę).

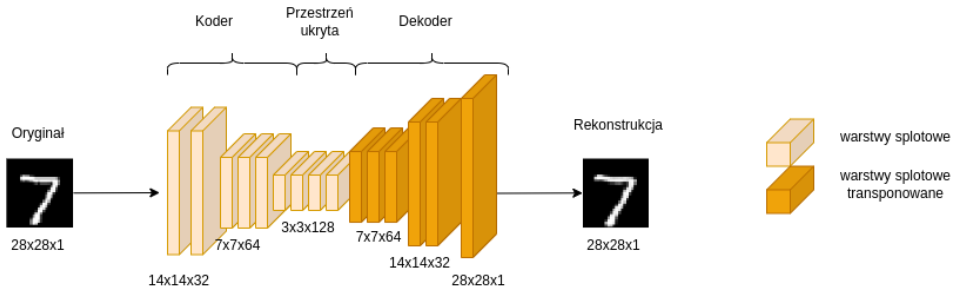


Rysunek 19. Wyniki zastosowania prostego autoenkodera przy wyszukiwaniu obrazu

Źródło: opracowanie własne.

### 2.5. Wykrywanie anomalii przy użyciu autoenkodera

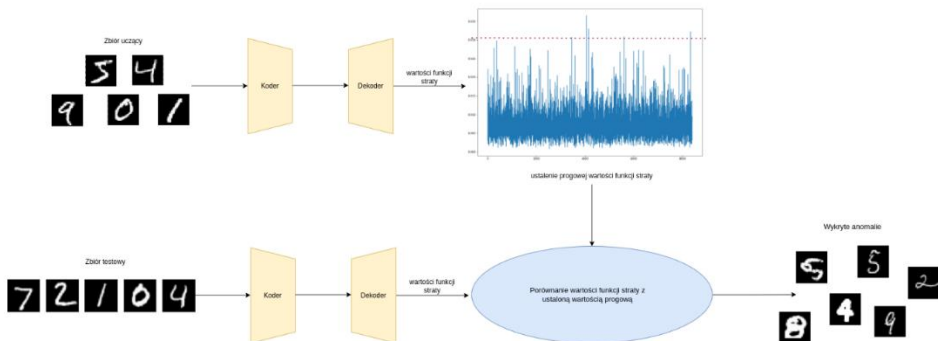
Wykrywanie anomalii przy użyciu autoenkodera zwykle polega na tym, że sieć autoenkodera jest trenowana na zbiorze „prawidłowych” danych – bez anomalii. Następnie, jeśli w zbiorze testowym model trafi na anomalię, funkcja straty dla niej powinna mieć znacznie większe wartości niż dla obserwacji podobnych do tych „prawidłowych”. W przypadku nienadzorowanym, kiedy nie wiemy, które obserwacje są anormalne, w zbiorze uczącym mogą się znaleźć także anomalie. Na podstawie wartości funkcji straty dla obserwacji zbioru uczącego, ustalamy wartość graniczną, powyżej której przykłady testowe będą uznawane za anomalię.



Rysunek 20. Struktura autoenkodera splotowego użytego do wykrywania anomalii

Źródło: opracowanie własne.

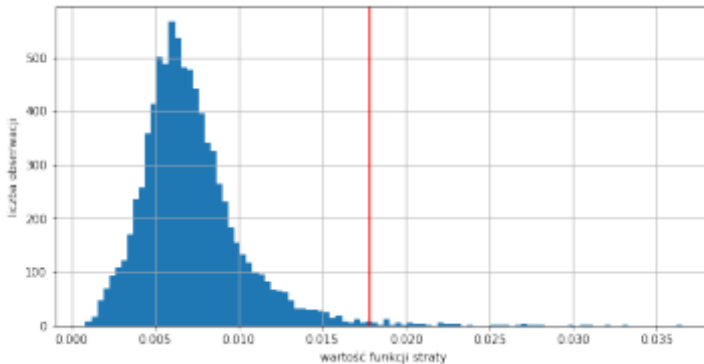
Struktura autoenkodera splotowego wykorzystanego w zadaniu wykrywania anomalii w zbiorze MNIST jest przedstawiona na rysunku 20. W każdej warstwie jądro ma rozmiar 3 x 3, a krok wynosi 2. Funkcją aktywacji w każdej warstwie jest funkcja ReLU. Zastosowano uzupełnianie zerami. Schemat zadania wykrywania anomalii za pomocą autoenkodera jest przedstawiony na rysunku 21.



Rysunek 21. Schemat wykrywania anomalii za pomocą autoenkodera

Źródło: opracowanie własne.

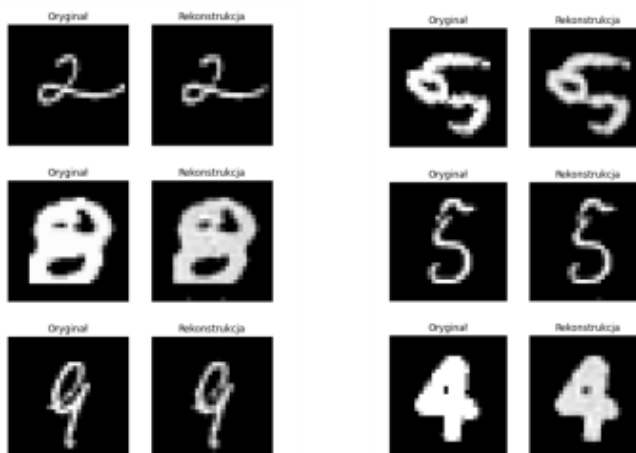
Funkcja straty bazuje na wskaźniku podobieństwa strukturalnego SSIM [10], a dokładniej jest to jeden minus wartość tego wskaźnika. Autoenkoder jest trenowany przez 5 epok. Na rysunku 22 widoczny jest histogram wartości funkcji straty dla obserwacji ze zbioru uczącego. Graniczną wartością straty, przy której uznamy obserwację za anormalną, jest dziewięćdziesiąty dziewiąty percentyl wartości funkcji straty ze zbioru treningowego (wynosi on około 0,017821). Wartość ta na wykresie jest zaznaczona za pomocą pionowej czerwonej linii.



**Rysunek 22.** Wartości funkcji straty z autoenkodera do wykrywania anomalii, z zaznaczoną wartością progową, powyżej której obserwacja jest uznawana za anomalię

Źródło: opracowanie własne.

Przykłady obserwacji uznanych przez opisany model za anomalie zaprezentowano na rysunku 23. Dla każdej obserwacji przedstawiono jej oryginał oraz rekonstrukcję z autoenkodera.

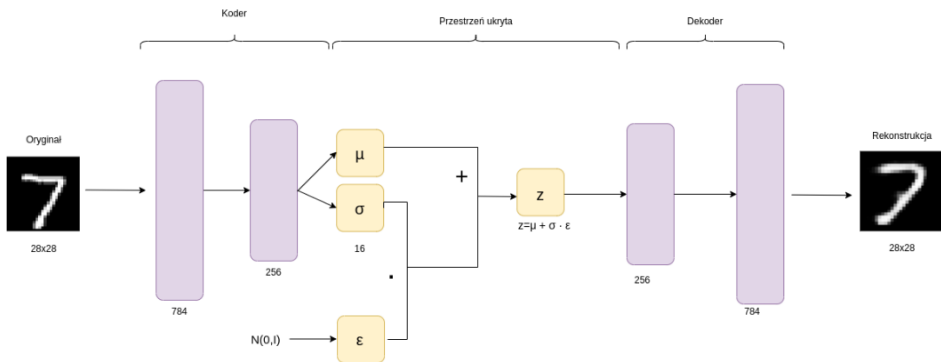


**Rysunek 23.** Przykłady anomalii w zbiorze MNIST wykrytych za pomocą autoenkodera

Źródło: opracowanie własne.

## 2.6. Generowanie obrazów przy użyciu autoenkodera wariacyjnego

Do generowania obrazów podobnych do obserwacji ze zbioru uczącego MNIST zostanie użyty autoenkoder o warstwach gęstych. Na wejściu model dostaje zatem obrazy „spłaszczone” do wektorów o długości 784. Koder zawiera jedną warstwę ukrytą, której wymiar wyjściowy to 256, a funkcją aktywacji jest ReLU. W kolejnej warstwie zachodzi kodowanie ukryte do wymiaru 16. Na podstawie wartości  $\mu$  i  $\sigma$  dokonywane jest próbkowanie z rozkładu normalnego w przestrzeni ukrytej. Wynik tego próbkowania jest przekazywany do dekodera. Dekoder składa się z dwóch warstw ukrytych (gęstych). Pierwsza zwiększa wymiar do 256, jej funkcją aktywacji jest ReLU. Druga zwiększa wymiar do oryginalnego 784, a jej funkcją aktywacji to funkcja sigmoidalna. Strukturą tego autoenkodera jest przedstawiona na rysunku 24. Funkcją straty jest binarna entropia krzyżowa. Funkcją straty ukrytej jest dywergencja Kullbacka–Leiblera. Liczba epok uczenia wynosi 50.



Rysunek 24. Struktura autoenkodera wariacyjnego

Źródło: opracowanie własne.

Na rysunku 25 widoczne są obrazy liczb wygenerowane przez autoenkoder wariacyjny. Generowanie każdego obrazu polega na wylosowaniu z rozkładu normalnego wektora o rozmiarze kodowania ukrytego (w naszym przykładzie jest to 16), a następnie na dokonaniu dla tego wektora predykcji z dekodera. W ten sposób otrzymujemy wektor długości 784, który następnie przekształcany jest do wymiaru 28 x 28 i prezentowany w formie czarno-białego obrazu.



Rysunek 25. Przykłady obrazów cyfr wygenerowanych przez autoenkoder wariacyjny

Źródło: opracowanie własne.

### 3. Podsumowanie

W tej pracy przedstawiono większość ważniejszych zastosowań autoenkoderów związanych z przetwarzaniem obrazu. Warto zaznaczyć, że opisane zastosowania da się również przenieść na dane innego rodzaju – na przykład autoenkodera odsumiającego można użyć do uzupełniania braków danych, a autoenkodera niedopełnionego do przeprowadzenia uogólnienia analizy składowych głównych na danych liczbowych. Autoenkodery mają też wiele zastosowań, których przykłady nie zostały podane w tej pracy – przykładowo autoenkoder rekurencyjny może być używany chociażby do danych w formie szeregu czasowego albo filmu czy mowy. Podsumowując, autoenkodery da się zastosować w szerokiej gamie problemów, można je również łączyć z innymi modelami uczenia maszynowego.

### Literatura

- [1] Artificial Intelligence in Plain English, *Convolutional Autoencoders (CAE) with Tensorflow*, <https://ai.plainenglish.io/convolutional-autoencoders-cae-with-tensorflow-97e8d8859cbe> (dostęp: 15.05.2022).
- [2] Bank D., Koenigstein N., Giryes R., *Autoencoders*, 2021.
- [3] Géron A., *Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow*, Wydanie II, Helion, Gliwice, 2020.
- [4] Goodfellow I., Bengio Y., Courville A., *Deep Learning. Systemy uczące się*, Wydawnictwo Naukowe PWN, Warszawa 2018.
- [5] Krohn J., Beyleveld G., Bassens A., *Uczenie głębokie i sztuczna inteligencja. Interaktywny przewodnik ilustrowany*, Helion, Gliwice, 2022.

- [6] Mishra D., *Transposed Convolution Demystified*, <https://towardsdatascience.com/transposed-convolution-demystified-84ca81b4baba> (dostęp: 15.05.2022).
- [7] LeCun Y., Cortes C., Burges C. J. C., *THE MNIST DATABASE of handwritten digits*, <http://yann.lecun.com/exdb/mnist/> (dostęp: 29.05.2022).
- [8] Papers With Code, *MNIST Benchmark (Image Classification on MNIST)*, <https://paperswithcode.com/sota/image-classification-on-mnist> (dostęp: 31.05.2022).
- [9] Pröve P. L., *An Introduction to different Types of Convolutions in Deep Learning*, <https://towardsdatascience.com/types-of-convolutions-in-deep-learning-717013397f4d> (dostęp: 15.05.2022).
- [10] Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P., *Image Quality Assessment: From Error Visibility to Structural Similarity*, “IEEE Transactions on Image Processing” 2004, vol. 13, no. 4, pp. 600–612.

## Autoencoders in unsupervised machine learning

### Abstract

This chapter reviews the applications of autoencoders used in unsupervised machine learning. The concept of autoencoder and its types are introduced, and selected ones are described in more detail. Six examples of the use of autoencoders in unsupervised learning are shown. Applications are demonstrated using the example of the MNIST dataset consisting of black and white images of handwritten numbers. The functionalities of simple and convolutional autoencoder are shown, followed by their use for image de-noising, image retrieval, anomaly detection and generation of new images.

*Keywords:* autoencoders, artificial neural networks, computer vision, unsupervised learning.





**Agnieszka Szczepaniak<sup>1</sup>, Milena Welna<sup>2</sup>, Ewelina Stąpór<sup>3</sup>,  
Jolanta Wójtowicz<sup>4</sup>**

## **Inspektor ochrony danych osobowych jako droga rozwoju studentów**

### **Streszczenie**

W rozdziale został przedstawiony temat ochrony danych osobowych, historii ich powstania oraz zadania inspektora ochrony danych osobowych. Opisane zostały także regulacje prawne i zasady dotyczące przetwarzania danych osobowych. Przybliżona została działalność koła naukowego „InfoNaBi” oraz organizowany przez nie projekt dotyczący ochrony danych osobowych.

*Słowa kluczowe: ochrona danych osobowych, historia IODO, prywatność, inspektor ochrony danych osobowych, regulacje prawne.*

### **Wstęp**

Zadaniem kół naukowych jest rozwój zainteresowań i przedsiębiorczości studentów oraz poszerzanie ich doświadczeń. Studenckie Koło Naukowe Zastosowań Informatyki w Nauce i Biznesie „InfoNaBi” zostało założone dnia 3 października 2011 roku. Koło ma swoją siedzibę w Katedrze Metod i Technik Nauczania Politechniki Lubelskiej. Przeważają w nim studenci Wydziału Podstaw Techniki, a opiekunem Koła jest Robert Lis. Tematyką zainteresowań naukowych Koła jest wykorzystanie nowoczesnych systemów informatycznych w biznesie, wpływ nowych technologii na efektywność kształcenia, rozwijanie form komunikacji poza personalnej oraz trendy w kształceniu na odległość [1].

Tematem, jaki członkowie Koła chcą teraz zgłębić, jest zagadnienie dotyczące ochrony danych osobowych. Aktualnie Koło ma w planach przystąpienie do szkolenia, w celu zwiększenia świadomości na temat ochrony danych osobowych oraz

---

<sup>1</sup> Agnieszka Szczepaniak, Studenckie Koło Naukowe „InfoNaBi”, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>2</sup> Milena Welna, Studenckie Koło Naukowe „InfoNaBi”, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>3</sup> Ewelina Stąpór, Studenckie Koło Naukowe „InfoNaBi”, Wydział Podstaw Techniki, Politechnika Lubelska.

<sup>4</sup> Jolanta Wójtowicz, Studenckie Koło Naukowe „InfoNaBi”, Wydział Podstaw Techniki, Politechnika Lubelska.

uzyskania tytułu IODO zwiększającego naszą konkurencyjność na rynku pracy po skończonych studiach.

## 1. Rozwój pojęcia ochrony danych osobowych

Ochrona danych osobowych, inaczej RODO, to różnego rodzaju informacje, które pozwalają zidentyfikować daną osobę fizyczną na podstawie danych takich jak: numer identyfikacyjny, cechy fizyczne, fizjologiczne, ekonomiczne, społeczne, kulturowe czy umysłowe. Wokół danych osobowych skupia się prawo do ochrony tych danych. Rozumienie tych danych wyznaczone jest w Dyrektywie 95/46/WE. Społeczeństwo coraz bardziej interesuje się prawem ochrony danych osobowych, co sprawia, że możemy natknąć się na wiele poświęconych temu zagadnieniu artykułów w internecie czy prasie, co pomaga zwiększyć świadomość w tym zakresie. Odpowiednie organy wnikliwie przyglądają się różnym instytucjom, zwracając uwagę, czy ich rozwiązania są wystarczająco dobre, aby nasze dane były bezpieczne.

Idea ochrony danych osobowych ściśle łączy się z koncepcją prywatności. Obie mają na celu umożliwienie jednostkom podejmowanie dobrowolnych decyzji, niezależnych od wszelkich instytucji, bez ingerencji i konsekwencji ze strony np. państwa, przedsiębiorstw, pracodawcy. Prawo do prywatności łączy się silnie z wolnością człowieka. Stawia na niezależność człowieka we wszystkich dziedzinach życia. Każda jednostka ma możliwość dobrowolnego i indywidualnego udostępniania informacji dotyczących swojej osoby wybranym podmiotom. Prawo do prywatności jest podstawą ochrony danych osobowych, a także jest pojęciem szerszym niż sama ochrona danych osobowych. Pomiędzy prawem do prywatności a prawem do ochrony danych osobowych istnieje jeszcze jeden znaczący związek: ich źródłem jest godność ludzka (co wynika wprost z treści Konstytucji) [2], a ich ochrona ma zapewnić nienaruszalność tej godności. Bezpośrednią zaś funkcją, jaką niewątpliwie spełnia ochrona danych osobowych, jest zabezpieczenie samego prawa do prywatności [3].

Rozwój technologii spowodował znacznie łatwiejszy dostęp do informacji. W poprzednich wiekach, kiedy informacje przekazywane i gromadzone były na papierze, dostęp do nich miały jedynie osoby, którym ten dokument został przekazany. Udostępnianie i pozyskiwanie danych było więc manualne, wymagało przepisywania, przesyłania dokumentów, co zabierało zdecydowanie więcej czasu niż obecnie. Rozwój cywilizacyjny, wraz z rozwojem technologii informatycznych, spowodował tworzenie baz danych, które znacznie szybciej mogą być przetwarzane oraz udostępniane. Ze względu na wygodę i łatwość wyparły one tradycyjne manualne metody. Dostępność informacji jest potrzebna przede wszystkim ze względów ekonomicznych. Łatwe i szybkie przekazywanie danych stało się wówczas zagrożeniem dla prywatności jednostki. Nie miała ona kontroli do kogo dotrą jej indywidualne informacje i w jaki sposób zostaną one wykorzystane. Wymusiło to rozwój ochrony tych danych, konieczność stworzenia regula-

cji prawnych, które skupiałyby się na wytyczeniu granic, stref pozbawionych ingerencji instytucji, aby chronić prywatność jednostki.

Na przełomie XIX i XX wieku sformułowane zostały podstawowe prawa człowieka. Prawa pierwszej generacji to te najbardziej podstawowe, wynikające z ludzkiej natury i niezależne od sytuacji geopolitycznych – jest to właśnie prawo do prywatności. Akcentują one konieczność zachowania pewnego obszaru wolnego od ingerencji, mają zatem charakter obronny i służą wzmocnieniu autonomii jednostki. Do praw jednostki pierwszej generacji, obok prawa do ochrony prywatności, zaliczyć możemy także: wolność słowa, równość wobec prawa, nietykalność osobistą, ochronę życia ludzkiego [4]. Prawa pierwszej generacji są uznawane przez większość krajów i są regulowane przede wszystkim przez Powszechną Deklarację Praw Człowieka i Międzynarodowy Pakt Praw Człowieka, na bazie których powstawały poszczególne lokalne regulacje. W Polsce jest to np. prawo do ochrony prywatności mieszczące się w Konstytucji.

„art. 51. **Nikt nie może być obowiązany inaczej niż na podstawie ustawy do ujawniania informacji dotyczących jego osoby.** Władze publiczne nie mogą pozyskiwać, gromadzić i udostępniać innych informacji o obywatelach niż niezbędne w demokratycznym państwie prawnym” [5].

Konstytucja jest podstawowym źródłem prawa do prywatności. Poza nią zasady RODO precyzują: Rozporządzenie 2016/679 (RODO), Dyrektywa 2016/680, Ustawa z dnia 21 lutego 2019 r. o zmianie niektórych ustaw w związku z zapewnieniem stosowania RODO, Ustawa z 10 maja 2018 o ochronie danych osobowych.

Do najbardziej popularnych należy rozporządzenie z 2016 roku, potocznie nazywane przepisami RODO. Zostało ono przyjęte przez Unię Europejską, po to, żeby ujednoczyć przepisy obowiązujące na całym jej terenie. Zaczęło obowiązywać we wszystkich krajach członkowskich UE od dnia 25 maja 2018 roku. Dzięki temu, że przepisy te zostały wprowadzone przez Unię Europejską, nie trzeba było wdrażać przepisów w poszczególnych krajach członkowskich.

Głównymi ich celami jest ochrona osób fizycznych, ich danych, przestrzeganie ich praw i wolności. Celem było również uaktualnienie przepisów do sytuacji rynkowych, czyli wzrostu przepływów informacji pomiędzy państwami. Unia Europejska chciała dzięki nim uzyskać także wzrost wzajemnego zaufania. Przepisy RODO wprowadziły trojakią ochronę danych, tzn. ochronę na etapie przetwarzania danych – określone prawa, obowiązki osób odpowiedzialnych za przetwarzanie danych. Rozporządzenie RODO zapewniło państwom członkowskim Unii Europejskiej jednolitość prawa, która dotyczy przetwarzania danych. A trzecim – ostatnim – jest egzekwowanie, kontrola, wprowadzenie kar za nieprawidłowe stosowanie się do przepisów.

## 2. RODO a zadania IODO

Obowiązuje siedem głównych zasad dotyczących przetwarzania danych. Opisane są w rozporządzeniu 2016/679 (RODO) w artykule piątym [6]. Można wśród nich wyszczególnić:

- **Zasada rzetelności, zgodności z prawem i przejrzystości.**  
To zasada mówiąca o obowiązku informacyjnym wobec podmiotów, których dane są przetwarzane. Wskazuje na to, że dane powinny być zgodne z obowiązującym prawem, rzetelne i łatwo dostępne. Zasada przejrzystości zobowiązuje do tego, aby właściciel informacji został poinformowany, że dane są zbierane, w jaki sposób będą wykorzystywane i przetwarzane. Informacje te powinny być przekazane w sposób zrozumiały i prosty dla odbiorcy.
- **Zasada ograniczenia celu.**  
Zasada ta wskazuje na poinformowanie o celu, w jakim dane będą przetwarzane. Cel musi być wyraźnie sformułowany, a także zgodny z prawem, a właściciel danych musi zostać o nim poinformowany. Jeżeli administrator chciałby przetwarzać dane w innym celu niż wskazany, ma obowiązek poinformować osobę, której te dane dotyczą. Dozwolone jest przetwarzanie danych w celach archiwalnych w interesie publicznym, statystycznych, badań naukowych lub historycznych.
- **Zasada minimalizacji danych.**  
Zasada minimalizacji odnosi się do ograniczenia przetwarzania danych do informacji jedynie koniecznych do osiągnięcia zamierzonego celu. Zasada wskazuje także, aby przetwarzanie danych osobowych wykorzystywać jedynie w sytuacjach, gdy jest to konieczne. Jeśli do wykonania analizy nie konieczne są informacje personalne dotyczące konkretnych osób (np. imię i nazwisko, adres zamieszkania), to nie należy ich pobierać.
- **Zasada prawidłowości.**  
Należy podjąć odpowiednie kroki techniczne i organizacyjne, aby przechowywane dane były prawidłowe. Zasada wskazuje na konieczność uaktualniania danych oraz usuwania błędnych.
- **Zasada ograniczenia przechowywania.**  
Zasada ta łączy się z zasadą minimalizmu. Skupia się ona na czasie przechowywania danych – powinien być on niezbędny do realizacji określonego celu. Podobnie jak w zasadzie minimalizacji, występuje wyjątek opisany w artykule 89 ust. 1., mówiący, że dane mogą być przechowywane po osiągnięciu celu w celach archiwalnych w interesie publicznym, badań naukowych, badań historycznych, badań statystycznych.
- **Zasada integralności i poufności.**  
Integralność danych jest to ochrona przed dekompletacją danych. Dane podczas przechowywania i udostępniania są narażone na utratę spójności.

Należy podjąć wszelkie działania organizacyjne i wykorzystać możliwości techniczne, które będą zabezpieczały dane przed ich uszkodzeniem.

- Zasada rozliczalności.

Administrator jest odpowiedzialny za przestrzeganie przepisów. Jego zadaniem danych osobowych jest udokumentowanie i wykazanie działań zgodnych z prawem.

Zasady dotyczące przetwarzania danych osobowych opisane w powyższym rozporządzeniu pokazują, że administrator może przetwarzać dane tylko, gdy ma do tego podstawy prawne, w określonych celach i w określony sposób. Administrator ma obowiązek przestrzegania zasad związanych z RODO, za uchybienia w tej kwestii może on ponieść karę pieniężną.

Inspektor ochrony danych osobowych to osoba, która jest organem doradczym we wdrażaniu przepisów o ochronie danych osobowych w określonej organizacji. Do obowiązków inspektora należy m.in. sprawdzenie stanu przestrzegania tych przepisów, informowanie o zakresie poufności danych oraz pomoc w przygotowaniu analizy ryzyka dla ochrony tych danych. Inspektor ochrony danych osobowych jest pierwszym punktem kontaktowym dla podmiotu nadzorującego wdrażanie standardów zgodności z RODO. Dlatego jednym z najważniejszych zadań inspektora jest pełnienie roli pośrednika między odpowiedzialnym podmiotem przetwarzającym a organem, który monitoruje i reguluje zarządzanie nim [7].

Inspektor ochrony danych ma odpowiedzialne działanie w przedsiębiorstwie. Jego decyzje powinny być trafne i zgodne z obowiązującym prawem. Może być pracownikiem firmy albo działać z zewnątrz. Inspektor powinien zostać zgłoszony do Prezesa Urzędu Ochrony Danych Osobowych.

### **3. Student jako inspektor ochrony danych**

Aktualnie członkowie koła mają w planach przystąpienie do szkolenia na inspektora ochrony danych osobowych, w celu zwiększenia świadomości na temat ochrony danych osobowych i uzyskania tytułu IODO, zwiększającego konkurencyjność na rynku pracy po skończonych studiach. W tym celu dokonano dogłębnej analizy dokumentacji prawnej dotyczącej legalizacji szkoleń i egzaminów. Przeprowadzono szereg rozmów telefonicznych mających za zadanie porównanie zakresu szkolenia z wytycznymi Ministerstwa. Wielu studentów wysłało zapytania e-mail o oferty cenowe szkoleń. Ważnym elementem prowadzonego sondażu było pytanie o uznawanie certyfikatu otrzymanego po zdaniu egzaminu.

Pierwszym krokiem, jaki podjęliśmy, był wybór firmy, która miała świadczyć dla nas usługi szkoleniowe. Po wstępnym zapoznaniu się z ofertami firm, wskazaliśmy najlepszych kandydatów. Chcieliśmy, aby certyfikat wystawiany przez wyłonioną firmę był jak najbardziej uznawany na rynku. Obecnie nie ma możliwości uzyskania certyfikatu od Prezesa Urzędu Ochrony Danych Osobowych. Nie ma też jednostek akredytowanych przez UODO. Kompetencje IODO określa się na podstawie kwalifikacji zawodowych, a przede wszystkim posiadanej wiedzy

praktycznej i teoretycznej. Dlatego zdecydowaliśmy się na firmę, która zaferowała szeroki zakres wiedzy teoretycznej, a także naukę umiejętności praktycznych wraz z materiałami dydaktycznymi, jakie pomocne będą w dalszym szkoleniu studentów. Dodatkowo szkolenie to ma być przeprowadzone przez specjalistów, działających w dziedzinie ochrony danych osobowych od wielu lat.

Po dokonaniu wyboru firmy, w której przystąpimy do szkolenia, będziemy w stanie określić środki, jakich potrzebujemy na przeszkolenie co najmniej dwóch studentów oraz opiekun koła naukowego, którzy mogliby przekazać zdobytą tam wiedzę naszym studentom. Budżet potrzeby na realizację szkolenia jest uzależniony nie tylko od wybranej firmy i cen szkoleń przeprowadzanych przez nią, ale także od liczby osób, które będą przeszkolone. Na tej podstawie została określona potrzebna pula środków. Kolejny krok do realizacji projektu to zdobycie funduszy. W tym celu opracowano biznesplan i przedstawiono go potencjalnym sponsorom. Jest on wciąż modyfikowany ze względu na zmieniające się ceny oraz uzyskiwane upusty cenowe.

Kolejnym krokiem realizacji projektu będzie zdobycie środków i zapis na konkretny termin szkolenia oraz jego odbycie i uzyskanie certyfikatu Inspektora Ochrony Danych Osobowych.

Następnym etapem przedsięwzięcia jest wyznaczenie terminu oraz sali na szkolenie, które zostanie przeprowadzone przez wcześniej przeszkolonych studentów, pod okiem opiekuna koła "InfoNaBi". Przeprowadzona zostanie wstępna selekcja osób chcących uzyskać certyfikat IODO. W pierwszej kolejności możliwość wzięcia udziału będą mieć studenci z kierunku inżynieria bezpieczeństwa, a gdy zostaną wolne miejsca – osoby z pozostałych kierunków. Znaczenie będzie miała również kolejność zapisów. Po zakończeniu szkolenia, na Politechnice Lubelskiej odbędzie się egzamin. Studenci, którzy go zaliczą, otrzymają certyfikat potwierdzający odbycie szkolenia na Inspektorów Ochrony Danych Osobowych wystawiony przez Politechnikę Lubelską i formowany przez władze Wydziału Podstaw Techniki. Egzamin zostanie przeprowadzony zgodnie z wymogami ministerialnymi i pod nadzorem opiekuna koła naukowego. Będzie miał formę testu, zarówno z pytaniami zamkniętymi, jak i otwartymi, z zagadnień omawianych na szkoleniu.

Cele projektu:

- zwrócenie uwagi studentów na istotę ochrony danych osobowych,
- poszerzenie horyzontów,
- zwiększenie możliwości zdobycia przyszłego zatrudnienia,
- zachęcenie studentów do rozwoju w kierunku pracy związanej z ochroną danych osobowych,
- zwiększenie świadomości o RODO,
- pozytywny wpływ na wizerunek Wydziału Podstaw Techniki,
- uatrakcyjnienie kierunku studiów IB oraz ETI,
- wzrost umiejętności i kompetencji studenta.

## 4. Podsumowanie

RODO w obecnych czasach staje się coraz bardziej powszechnym elementem funkcjonowania firm i przedsiębiorstw. Każda organizacja, której działanie wymaga przetwarzania danych osobowych, bez względu na jej wielkość, jest zobowiązana do przestrzegania przepisów ochrony danych osobowych. Nie tylko IODO powinni interesować się tematem RODO, ponieważ dane wszystkich ludzi są przetwarzane. Warto zwrócić uwagę, na co zgadzamy się podczas zawierania różnych umów i wyrażać zgody na przetwarzanie danych z większą świadomością i uwagą.

Politechnika Lubelska wraz z Kołem Naukowym “InfoNaBi” wychodzi na przeciw oczekiwaniom pracodawców i przygotowuje projekt, który ma na celu zapoznanie studentów z przepisami i wymaganiami dotyczącymi ochrony danych. Zwiększy on świadomość studentów w zakresie RODO, a także da im możliwość do dalszego rozwoju zawodowego i zdobywania doświadczenia.

## Literatura

- [1] Strona Koła Naukowego InfoNaBi – informatyka w nauce i biznesie, <https://ti.pollub.pl/informacje/statut> (dostęp: 09.03.2023).
- [2] Fleszer D., *Zakres przetwarzania danych osobowych w działalności gospodarczej*, Wolters Kluwer, Warszawa 2008.
- [3] Sibiga G., *Postępowanie w sprawach ochrony danych osobowych*, Dom Wydawniczy ABC, Warszawa 2002.
- [4] Kuczma P., *Prawa człowieka w zarysie*, Wydawnictwo Dolnośląskiej Wyższej Szkoły Przedsiębiorczości i Techniki, Polkowice 2012.
- [5] Konstytucja RP art. 51.
- [6] Rozporządzenie 2016/679 (RODO) art. 5 ust. 1.
- [7] Borowska A., *Na czym polega ochrona danych osobowych?*, <https://poradnikprzedsiębiorcy.pl/-na-czym-polega-ochrona-danych-osobowych> (dostęp: 09.03.2023).
- [8] Rozporządzenie 2016/679 (RODO) art. 39.
- [9] Góral A., *Inspektor ochrony danych – kim jest i czym się zajmuje i ile zarabia?*, <https://www.aplikuj.pl/porady-dla-pracownikow/2636/inspektor-ochrony-danych-kim-jest-czym-sie-zajmuje-i-ile-zarabia> (dostęp: 09.03.2023).



## **Student as a personal data protection officer**

### **Abstract**

The chapter presents the topic of personal data protection, the history of its creation and the tasks of the personal data protection officer. Legal regulations and rules regarding the processing of personal data have also been described. The activities of the “InfoNaBi” student science club and the new project organized by this student science club were presented.

*Keywords: privacy law, history of personal data protection, privacy, personal data protection officer, General Data Protection Regulation.*

## **Modelowanie procesu eksploatacji pojazdów wojskowych w aspekcie gotowości<sup>2</sup>**

### **Streszczenie**

Modelowanie gotowości technicznej ma szczególne znaczenie w systemach charakteryzujących się losowością występowania zadań. W analizie niezawodności coraz częściej wykorzystywane są języki programowania i narzędzia umożliwiające nie tylko prawidłowe funkcjonowanie badanego procesu, ale również stanowiące wsparcie w procesie podejmowania decyzji poprzez szybką analizę i interpretację napływających strumieni danych. Celem niniejszych badań było wyznaczenie, przy wykorzystaniu oprogramowania R, miary niezawodności, jaką jest współczynnik gotowości technicznej floty pojazdów ciężarowych realizujących zadania transportowe na terytorium Rzeczypospolitej Polskiej. W oparciu o przeprowadzoną analizę literatury przytoczono najważniejsze definicje dotyczące teorii niezawodności, na podstawie których opracowano założenia i skonstruowano matematyczny model rzeczywistego procesu eksploatacji środków transportu. W wyniku przeprowadzonych obliczeń i analiz, stwierdzono, że współczynnik gotowości technicznej równy 96,63(%) świadczy o prawidłowej organizacji procesu eksploatacji rozpatrywanej floty pojazdów.

*Słowa kluczowe: system eksploatacji, niezawodność, gotowość.*

### **Wstęp**

W literaturze często podejmowana jest tematyka związana z modelowaniem systemów eksploatacji środków transportu. Wykorzystanie odpowiednich metod matematycznych, w celu sterowania procesem eksploatacji, umożliwia podejmowanie racjonalnych decyzji, przy jednoczesnym zapewnieniu prawidłowej realizacji przydzielonych zadań. Jedną z miar opisujących funkcjonowanie systemu eksploatacji środków transportu jest jego gotowość do realizacji określonych zadań przewozowych. Metody modelowania procesów eksploatacji obiektów technicznych stanowią przedmiot zainteresowania m.in. teorii niezawodności, badań operacyjnych, teorii podejmowania decyzji, modelowania symulacyjnego, teorii odnowy czy też metod analizy statystycznej.

---

<sup>1</sup> Mgr inż. Joanna Szkutnik-Rogoż, Wojskowa Akademia Techniczna, joanna.szkutnik@wat.edu.pl.

<sup>2</sup> Praca została sfinansowana przez Wojskową Akademię Techniczną w ramach projektu nr UGB 799.

Teoria niezawodności oraz teoria odnowy są często wykorzystywane do rozwiązywania problemów związanych nie tylko z procesem użytkowania, ale również z zapewnieniem zdatności obiektów technicznych [1]. Zastosowanie praktyczne teorii niezawodności rozpatrywano w [2]. W artykule [3] zaproponowano strategię utrzymania ruchu statycznego, dynamicznego i limitów awaryjności. Analiza niezawodności systemów zrównoważonych stała się w ostatnich latach ważnym tematem w dziedzinie badań niezawodności. Model matematyczny i wyniki przedstawione w pracy [4] mogą być wykorzystane m.in. w niezawodności, zarządzaniu łańcuchem dostaw oraz planowaniu transportu. Algorytm umożliwiający wyznaczenie funkcji niezawodności został zaproponowany w artykule [5]. Opracowana metoda może być wykorzystana do szczegółowego badania stacjonarnych i niestacjonarnych charakterystyk niezawodnościowych. Innowacyjny algorytm aproksymacyjny przeznaczony do oceny niezawodności modelu sieci opartego na ścieżce ruchu został zaproponowany w [6]. Przykłady liczbowe pokazały, że proponowany algorytm może być zaimplementowany dla innych miar niezawodności, takich jak częstotliwość awarii, średni czas między awariami i średni czas naprawy.

W analizie niezawodności bardzo istotne znaczenie ma transformata Laplace'a. W [7] wprowadzono dwa przedziałowe wskaźniki dostępności dla naprawialnych systemów komputerowych. Wzory obliczeniowe zostały przedstawione w postaci macierzowej za pomocą przekształcenia Laplace'a. W pracy [8] przeanalizowano modele stochastyczne opisujące wielokrotną dostępność systemu z okresami eksploatacji i naprawy, które tworzą naprzemienny proces odnowy z wykładniczymi czasami do awarii i naprawy, natomiast równania całkowite wyprowadzono za pomocą transformaty Laplace'a. Teoria naprzemiennych procesów odnowy z wykorzystaniem transformaty Laplace'a do badania funkcji relaksacji została zaproponowana w [9]. W artykule [10] zastosowano transformatę Laplace'a mającą na celu wyznaczenie jawnej postaci funkcji niezawodności oraz średni czas do uszkodzenia systemu. W [11] wykorzystano transformatę Laplace'a do opracowania dwóch miar wydajności systemu, takich jak niezawodność systemu i średni czas do awarii systemu.

Rozwiązywanie problemów decyzyjnych, dotyczących oceny działania złożonych systemów eksploatacji różnych obiektów technicznych za pomocą metod analitycznych, często jest utrudnione lub niemożliwe do zrealizowania. W związku z tym, w celu rozwiązywania tego typu problemów, coraz częściej wykorzystywane są techniki symulacyjne. Podstawą budowy modeli symulacyjnych jest wyznaczenie rozkładów prawdopodobieństwa rozpatrywanych zmiennych losowych, opisujących badany system lub proces. Wyniki przeprowadzonych badań symulacyjnych często stanowią podstawę podejmowania racjonalnych decyzji dotyczących sterowania procesem eksploatacji [12]. Problematykę związaną z modelowaniem gotowości w systemach transportowych rozpatrywano w artykułach [13, 14]. W pracy [15] zaproponowano algorytm oparty na minimalnych ścieżkach oceny niezawodności systemu transportu kolejowego. Adapta-

cyjny kontroler sygnału do zarządzania opóźnieniami ruchu i niezawodnością usług autobusów miejskich z adaptowanymi acyklicznymi planami czasowymi został przedstawiony w [16]. Teoria wartości ekstremalnych do modelowania rzadkich zdarzeń, które pojawiają się jako opóźnienia w dostawach w transporcie drogowym, została wykorzystana w [17]. Wyniki badań potwierdzających, że dane automatycznej lokalizacji pojazdów mogą być używane do monitorowania niezawodności autobusów w czasie rzeczywistym omówiono w [18]. Zapewnienie wymaganego poziomu gotowości systemu eksploatacji środków transportu jest możliwe dzięki właściwie dobranej strategii sterowania procesem, wynikającej z podejmowania decyzji optymalnych ze względu na przyjęte kryteria oceny [19]. Niezawodność i jakość obsługi stanowią jedne z najważniejszych czynników zwiększających wykorzystanie transportu publicznego. Model oceny jakości usług jako trzypoziomowa hierarchiczna struktura oceny systemów transportu publicznego podczas pandemii została zaproponowana w [20]. Do oceny gotowości zarówno systemów, jak i obiektów często wykorzystywane są modele Markowa i semi-Markowa.

Celem niniejszych badań było wyznaczenie przy wykorzystaniu oprogramowania R miary niezawodności, jaką jest współczynnik gotowości technicznej floty pojazdów ciężarowych realizujących zadania transportowe na terytorium Rzeczypospolitej Polskiej. W oparciu o przeprowadzoną analizę literatury przytoczono najważniejsze definicje dotyczące teorii niezawodności, na podstawie których opracowano założenia i skonstruowano matematyczny model rzeczywistego procesu eksploatacji środków transportu.

## 1. Założenia do modelu matematycznego

Matematyczny opis, umożliwiający wyznaczenie odpowiednich wskaźników niezawodności rozpatrywanych obiektów technicznych przy jednoczesnym uwzględnieniu losowych zmian ich stanów fizycznych, warunków eksploatacji oraz stawianych zadań, jest nazywany modelem niezawodnościowym (stochastycznym). Ważną klasą procesów stochastycznych są procesy Markowa i semi-Markowa, stanowiące dogodny aparat matematyczny umożliwiający opis i badanie procesów losowych, np. procesu eksploatacji obiektów. Łańcuch Markowa charakteryzuje się określoną liczbą dopuszczalnych stanów i regułami przejścia pomiędzy nimi. Zbiór wszystkich możliwych wartości procesu stochastycznego stanowi przestrzeń stanów procesu lub przestrzeń fazową procesu [13].

**Definicja 1.** Procesem stochastycznym  $\{X_t: t \in T\}$  nazywa się taką funkcję  $(t, \omega) \rightarrow X_t(\omega)$ ,  $\omega \in \Omega, t \in T$ , że  $X_t(\omega)$  jest zmienną losową dla każdego  $t \in T$ .

Zbiór wszystkich możliwych wartości procesu stochastycznego  $S = \{X_t(\omega): t \in T, \omega \in \Omega\}$  stanowi przestrzeń stanów procesu lub przestrzeń fazową procesu, natomiast stanami procesu są elementy zbioru  $S$ .

**Definicja 2.** Proces stochastyczny  $\{X_t: t \in T\}$  o wartościach w przeliczalnym zbiorze stanów  $S$  nazywany jest łańcuchem Markowa wtedy i tylko wtedy, gdy

dla  $s_0, s_1, \dots, s_{n+1} \in S$  oraz dla  $t_0, t_1, \dots, t_{n+1} \in \mathbb{R}_+$  takich, że  $0 \leq t_0 < t_1 < \dots < t_{n+1}$  spełniona jest własność Markowa:

$$\begin{aligned} & P(X_{t_{n+1}} = s_{n+1} | X_{t_n} = s_n, X_{t_{n-1}} = s_{n-1}, \dots, X_{t_1} = s_1, X_{t_0} = s_0) = \\ & = P(X_{t_{n+1}} = s_{n+1} | X_{t_n} = s_n), \end{aligned} \quad (1)$$

jeśli tylko  $P(X_{t_n} = s_n, X_{t_{n-1}} = s_{n-1}, \dots, X_{t_1} = s_1, X_{t_0} = s_0) > 0$ .

Warunek ten nazywany jest własnością braku pamięci dla ciągu zmiennych losowych  $X_t$ . Oznacza to, że informacja o stanie bezpośrednio poprzedzającym zawiera w sobie całą informację, jaką można uzyskać, analizując poprzednie stany.

**Definicja 3.** Łańcuch Markowa jest jednorodny, jeśli dla dowolnych stanów  $s_0, s_1, \dots, s_{n+1} \in S$  oraz dla  $t_n, t_{n+1} \in \mathbb{R}_+$ , takich, że  $0 \leq t_n < t_{n+1}$  zachodzi równość:

$$P(X_{t_{n+1}} = s_{n+1} | X_{t_n} = s_n) = P(X_{t_{n+1}-t_n} = s_{n+1} | X_0 = s_n) = p_{ij}(t_{n+1} - t_n) \quad (2)$$

Przyjmując, że  $t = t_{n+1} - t_n$ , wówczas:  $p_{ij}(t) = P(X_t = s_j | X_0 = s_i)$ ,  $s_i, s_j \in S, t \geq 0$ .

Liczba  $p_{ij}(t)$  oznacza prawdopodobieństwo przejścia ze stanu  $s_i$  do stanu  $s_j$  po czasie  $t$ .

**Definicja 4.** Macierz  $P = [p_{ij}]_{s_i, s_j \in S}$  nazywana jest macierzą przejścia na  $S$ , gdy  $p_{ij}(t) \geq 0$  dla  $t \geq 0$  oraz dla każdego  $i \in S$  zachodzi warunek:  $\sum_{j \in S} p_{ij} = 1$ .

Taka macierz nazywana jest również macierzą stochastyczną. W macierzy stochastycznej suma wszystkich elementów każdego wiersza wynosi 1. Dla  $n$  stanów  $s_1, \dots, s_j \in S$  macierz prawdopodobieństw przejść  $p_{ij}(t)$  dla  $i, j = 1, 2, 3, \dots, n$  można zapisać następująco:

$$P(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & p_{13}(t) & \dots & p_{1n}(t) \\ p_{21}(t) & p_{22}(t) & p_{23}(t) & \dots & p_{2n}(t) \\ p_{31}(t) & p_{32}(t) & p_{33}(t) & \dots & p_{3n}(t) \\ \dots & \dots & \dots & \dots & \dots \\ p_{n1}(t) & p_{n2}(t) & p_{n3}(t) & \dots & p_{nn}(t) \end{bmatrix}. \quad (3)$$

Korzystając ze wzoru na prawdopodobieństwo całkowite, dla każdego  $m$  takiego, że  $1 \leq m < t$  można zapisać równanie:

$$p_{ij}(m+t) = \sum_{k \in S} p_{kj}(m)p_{ik}(t), \quad (4)$$

nazywane równaniem Chapmana–Kołmogorowa. Z równania Chapmana–Kołmogorowa wynika, że  $P(m+t) = P(m)P(t)$ , co oznacza, że znając rozkład początkowy  $p(0) = [p_1(0), p_2(0), p_3(0), \dots]$  oraz macierz prawdopodobieństw

przejść możliwe jest wyznaczenie rozkładu  $p(n) = [p_1(n), p_2(n), p_3(n), \dots]$  po  $n$  krokach zgodnie z zależnością:

$$p(n) = p(0)P^n. \quad (5)$$

Granice  $\lim_{n \rightarrow \infty} p(n) = \Pi$ , gdzie  $\Pi = (\Pi_1, \Pi_2, \Pi_3, \dots)$  (o ile istnieje) nazywa się rozkładem granicznym łańcucha Markowa, natomiast łańcuch Markowa, dla którego istnieje rozkład graniczny niezależny od rozkładu początkowego, nazywa się łańcuchem ergodycznym. Zastosowanie łańcuchów Markowa umożliwia wyznaczenie prawdopodobieństwa znajdowania się obiektu w dowolnej chwili w jednym ze stanów procesu, jednakże niezbędne jest spełnienie założenia dotyczącego wykładniczego rozkładu prawdopodobieństwa czasu przebywania w poszczególnych stanach. W przypadku gdy warunek ten nie jest spełniony możliwe jest zastosowanie procesu semi-Markowa, dla którego rozkład czasu przebywania w poszczególnych stanach może być dowolny.

## 2. Procesy Markowa

Badaniom poddano 45 pojazdów ciężarowych tego samego typu realizujących zadania transportowe na terytorium Rzeczypospolitej Polskiej. Rozpatrywano samochody ciężarowe średniej ładowności wysokiej mobilności o masie własnej 7350 (kg) i ładowności 8000 (kg). Wszystkie pojazdy pochodziły z jednej partii produkcyjnej, co pozwoliło uznać próbę za jednorodną. Podstawę opracowania bazy danych stanowiła dokumentacja eksploatacyjna w formie rozkazów wyjazdu oraz kart usług technicznych. Dane liczbowe niezbędne do przeprowadzenia obliczeń obejmowały okres od 01.01.2017 r. do 31.12.2017 r. W modelu określono 9 stanów eksploatacyjnych oraz wskazano dozwolone przejścia międzystanowe. Przestrzeń fazowa obejmowała następujące stany eksploatacyjne pojazdów ciężarowych:

- $s_1$  – zadanie,
- $s_2$  – tankowanie,
- $s_3$  – garażowanie,
- $s_4$  – obsługa dzienna,
- $s_5$  – dzień techniczny,
- $s_6$  – naprawa bieżąca,
- $s_7$  – obsługa okresowa-1,
- $s_8$  – obsługa okresowa-2,
- $s_9$  – obsługa roczna.

Stany  $s_1, s_2, s_3, s_4, s_5$  zaliczono do stanów gotowości, natomiast stany  $s_6, s_7, s_8, s_9$  zaliczono do stanów niegotowości. Przyjęto założenie upraszczające, że czas przebywania w stanie  $s_i$  przed przejściem do stanu  $s_j$  jest zmienną losową o rozkładzie wykładniczym. W tabeli 1 zobrazowano macierz przejść międzystanowych ze stanu  $s_i$  do  $s_j$ , gdzie:

- $i$  – numer wiersza macierzy oznaczający numer stanu przed przejściem,
- $j$  – numer kolumny macierzy oznaczający numer stanu po przejściu,
- 0 – przejście zabronione,
- 1 – przejście dozwolone.

**Tabela 1. Macierz dozwolonych i zabronionych przejść międzystanowych pojazdów ciężarowych**

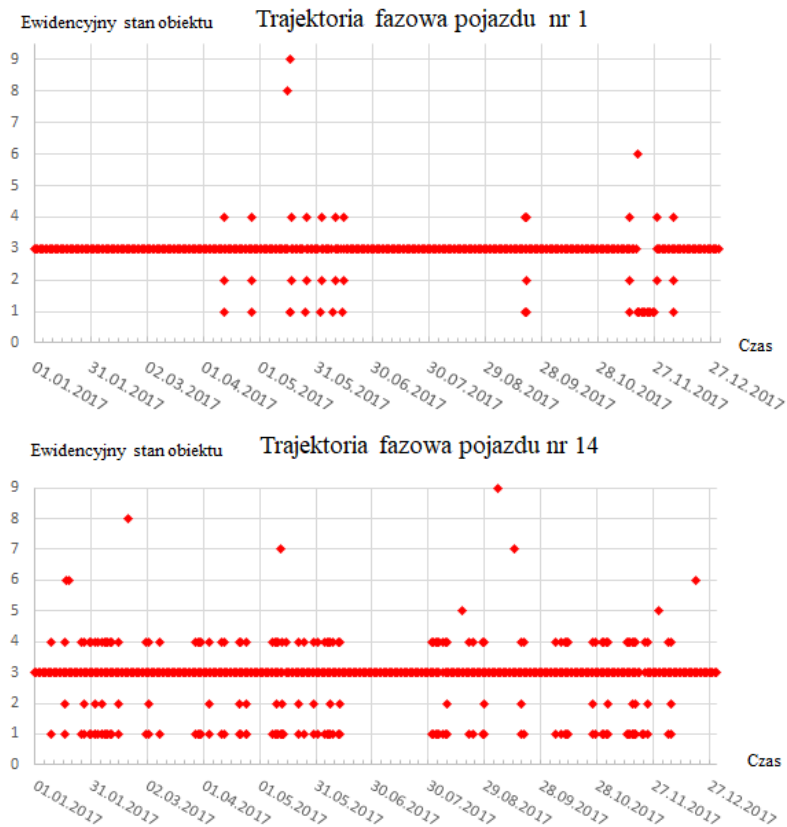
$\downarrow s_i$ $\rightarrow s_j$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$	$s_7$	$s_8$	$s_9$
$s_1$	0	1	0	1	1	1	1	1	1
$s_2$	1	0	0	1	0	0	0	0	0
$s_3$	1	0	0	0	1	1	1	1	1
$s_4$	0	0	1	0	0	0	0	0	0
$s_5$	1	0	1	1	0	0	0	0	0
$s_6$	1	1	1	1	0	0	0	0	0
$s_7$	1	1	1	0	0	0	0	0	0
$s_8$	1	1	1	0	0	0	0	0	0
$s_9$	1	0	1	1	0	0	0	0	0

Źródło: opracowanie własne.

W przypadku gdy przedmiotem badań jest zbiór obiektów tego samego typu, dla których możliwe jest wyznaczenie prawdopodobieństwa przebywania w określonym stanie, jako model należy zastosować  $n$ -wymiarowy proces losowy. Realizacjom takiego procesu losowego odpowiadają poszczególne trajektorie w przestrzeni fazowej. Źródłowe trajektorie fazowe w wersji punktowej dla dwóch losowo wybranych pojazdów zobrazowano na rysunku 1.

Matematycznym opisem procesu stochastycznego, z czasem dyskretnym o skończonej liczbie stanów, jest kwadratowa macierz stochastyczna o stopniu równym liczbie stanów procesu. Dla procesów eksploatacyjnych zakłada się brak przejść z powrotem do stanu poprzedniego, co narzuca zerowe wartości elementów diagonalnych. W tabeli 2 zestawiono liczbę wszystkich przejść międzystanowych pojazdów ciężarowych w rozpatrywanym okresie badawczym.

**Rysunek 1. Wizualizacja trajektorii fazowej dla losowo wybranych pojazdów**



Źródło: opracowanie własne.

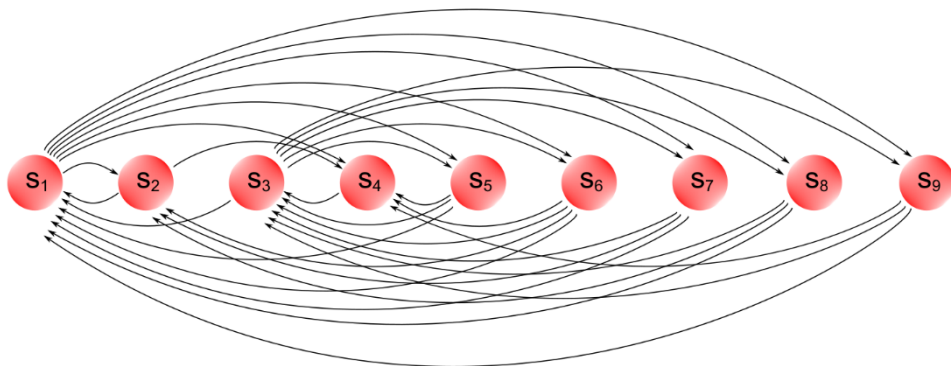
**Tabela 2. Macierz przejść międzystanowych pojazdów ciężarowych**

$\downarrow s_i \rightarrow s_j$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$	$s_7$	$s_8$	$s_9$
$s_1$	0	533	0	1110	13	18	14	1	17
$s_2$	286	0	0	251	0	0	0	0	0
$s_3$	1385	0	0	0	20	38	26	19	48
$s_4$	0	0	1376	0	0	0	0	0	0
$s_5$	10	0	19	4	0	0	0	0	0
$s_6$	8	1	36	11	0	0	0	0	0
$s_7$	8	2	30	0	0	0	0	0	0
$s_8$	2	1	17	0	0	0	0	0	0
$s_9$	7	0	57	1	0	0	0	0	0

Źródło: opracowanie własne.



Obrazem macierzy przejść dozwolonych jest skierowany graf przejść międzystanowych zobrazowany na rysunku 2.



**Rysunek 2. Graf procesu eksploatacji pojazdów**

Źródło: opracowanie własne.

Estymacja prawdopodobieństw przejść, na podstawie danych empirycznych dotyczących rozpatrywanego procesu eksploatacji, stanowi pierwszy etap budowy modelu matematycznego. Estymatorami  $\hat{p}_{ij}$  prawdopodobieństw przejść  $p_{ij}$  ze stanu  $s_i$  do  $s_j$  są częstości przejść  $W_{ij}$  opisane równaniem:

$$\hat{p}_{ij} = W_{ij} = \frac{N_{ij}}{N_i^{tr}}, \quad (6)$$

$$N_i^{tr} = \sum_{j \neq i} N_{ij}, \quad (7)$$

$$\sum_{j \neq i} W_{ij} = 1, \quad (8)$$

gdzie:

$\hat{p}_{ij}$  – częstość przejść ze stanu  $s_i$  do  $s_{j \neq i}$ ,

$N_{ij}$  – losowa liczba obserwacji stanu  $s_i$  przed  $s_j$ ,

$N_i^{tr}$  – liczba wszystkich przejść (wyjść) ze stanu  $s_i$ .

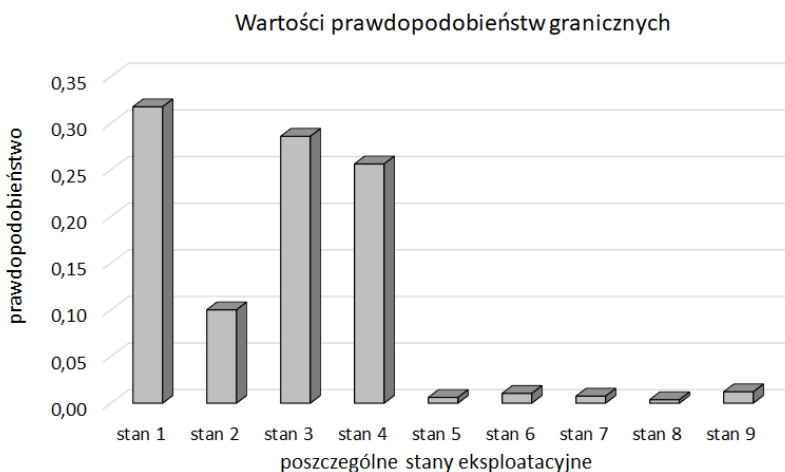
W celu wyznaczenia prawdopodobieństw granicznych dla procesu Markowa zastosowano oprogramowanie R. W wyniku przeprowadzonych obliczeń, otrzymane wartości prawdopodobieństw granicznych zestawiono w tabeli 3 oraz na rysunku 3.

**Tabela 3. Wartości prawdopodobieństw granicznych dla rozpatrywanych stanów eksploatacyjnych**

Stan	Prawdopodobieństwo graniczne
$S_1$	0,317691
$S_2$	0,100000
$S_3$	0,286034
$S_4$	0,256425
$S_5$	0,006145
$S_6$	0,010428
$S_7$	0,007449
$S_8$	0,003724
$S_9$	0,012104

Źródło: opracowanie własne.

**Rysunek 3. Wartości prawdopodobieństw granicznych przebywania w poszczególnych stanach eksploatacyjnych**



Źródło: opracowanie własne.

Celem oceny gotowości do realizacji stawianych zadań jest najczęściej wyznaczenie prognozy, czy dany obiekt lub system będzie znajdował się w stanie zdatności w przyszłości i czy będzie gotowy do natychmiastowego zrealizowania zadań zgodnie z przeznaczeniem. W literaturze nie ma obiektywnego podziału obiektów technicznych na elementy i systemy, w związku z czym podstawą ustalenia, czy dany obiekt będzie rozpatrywany jako system, czy też system będzie rozpatrywany jako zbiór obiektów, jest skonstruowany model matematyczny. Analiza niezawodności umożliwia określenie związków pomiędzy niezawodnością poszczególnych elementów systemu a niezawodnością całego systemu.

Liczbowa miarę niezawodności obiektów/systemów stanowią wskaźniki niezawodności, które mogą być wyznaczone zarówno metodami analitycznymi, jak i symulacyjnymi. Wartość wybranego wskaźnika niezawodności rozpatrywanego obiektu/systemu jest odpowiednią funkcją charakterystyk probabilistycznych modelu niezawodnościowego. Gotowość oznacza zdolność do natychmiastowej realizacji stawianych zadań we właściwym czasie, miejscu oraz przy ustalonych warunkach eksploatacji. W skonstruowanym modelu procesu eksploatacji floty pojazdów, gotowość można wyrazić poprzez iloraz sumy prawdopodobieństw granicznych stanów zdatności w stosunku do sumy prawdopodobieństw granicznych stanów zdatności i niezdatności:

$$k_g = \frac{\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4 + \Pi_5}{\sum_{i=1}^9 \Pi_i}. \quad (9)$$

Współczynnik gotowości technicznej można interpretować jako udział czasu przebywania obiektów technicznych w stanie zdatności w całym okresie eksploatacji. Dla rozpatrywanej floty pojazdów współczynnik gotowości technicznej wyniósł 96,63(%), co świadczy o prawidłowej realizacji procesu eksploatacji.

### 3. Podsumowanie

Metody modelowania procesów eksploatacji obiektów technicznych mają bardzo szeroki zakres zastosowania i stanowią przedmiot zainteresowania badań m.in. z zakresu teorii niezawodności, teorii podejmowania decyzji, badań operacyjnych, metod analizy statystycznej czy też metod modelowania symulacyjnego. W analizie niezawodności obiektów technicznych coraz częściej wykorzystywane jest oprogramowanie komputerowe oraz języki programowania. Oprogramowanie komputerowe zapewnia nie tylko prawidłowe działanie analizowanego systemu, ale może także stanowić wsparcie w procesie podejmowania decyzji poprzez szybką interpretację napływających strumieni danych. Korzystając z dostępnych bibliotek oraz pakietów dostępnych np. w oprogramowaniu R możliwe jest znaczne skrócenie czasu niezbędnego do przeprowadzenia obliczeń. Celem niniejszych badań było wyznaczenie, przy wykorzystaniu oprogramowania R, miary niezawodności, jaką jest współczynnik gotowości technicznej floty pojazdów ciężarowych realizujących zadania transportowe na terytorium Rzeczypospolitej Polskiej. W wyniku przeprowadzonych obliczeń stwierdzono, że otrzymany współczynnik gotowości technicznej równy 96,63(%) świadczy o prawidłowej organizacji procesu eksploatacji. Analiza wartości współczynnika gotowości technicznej umożliwia racjonalne sterowanie procesem eksploatacji środków transportu przy jednoczesnym zapewnieniu realizacji stawianych zadań. Zwiększenie gotowości technicznej systemu jest możliwe dzięki wprowadzeniu redundancji, czyli dodatkowych środków lub możliwości, które są niezbędne do realizacji stawianych zadań. Należy jednak pamiętać, że redundancja powinna być racjonalnie zaplanowana, gdyż nie każdy nadmiar powoduje zwiększenie nieza-

wodności systemu, ponadto nadmiarowość często jest przyczyną dodatkowych kosztów. Celem racjonalnego sterowania procesem eksploatacji obiektów technicznych jest opracowanie strategii zapewniającej minimalizację oczekiwanych kosztów eksploatacji obiektów, przy jednoczesnym zapewnieniu odpowiedniego (wymaganego) poziomu ich gotowości. Planowanie redundancji systemu jest zadaniem optymalizacyjnym, w którym bardzo istotna jest znajomość wskaźników gotowości i zdolności rozpatrywanych obiektów. Słusznym kolejnym kierunkiem w dalszych badaniach będzie opracowanie optymalnej strategii sterowania procesem eksploatacji, w oparciu o metodę Monte Carlo, umożliwiającą modelowanie procesów, których przebieg zależy od czynników losowych.

## **Finansowanie**

Praca została sfinansowana przez Wojskową Akademię Techniczną w ramach projektu nr UGB 799.

## **Literatura**

- [1] Ziółkowski J., Małachowski J., Oszczypała M., Szkutnik-Rogoż J., Lęgas A., *Modelling of the Military Helicopter Operation Process in Terms of Readiness*, "Defence Science Journal" 2021, vol. 71, no. 5, pp. 602–611.
- [2] Bougofa M., Taleb-Berrouane M., Bouafia A., Baziz A., Kharzi R., Bellaouar A., *Dynamic availability analysis using dynamic Bayesian and evidential networks*, "Process Safety and Environmental Protection" 2021, vol. 153, pp. 486–499.
- [3] Nguyen D., Dijoux Y., Fouladirad M., *Analytical properties of an imperfect repair model and application in preventive maintenance scheduling*, "European Journal of Operational Research" 2017, vol. 256, no. 2, pp. 439–453.
- [4] Fang C., Cui L., *Reliability evaluation for balanced systems with auto-balancing mechanisms*, "Reliability Engineering and System Safety" 2021, vol. 213, pp. 1–13.
- [5] Rykov V., Kozyrev D., Filimonov A., Ivanova N., *On reliability function of a k-out-of-n system with general repair time distribution*, "Probability in the Engineering and Informational Sciences" 2020, vol. 35, no. 4, pp. 885–902.
- [6] Hayashi M., Abe T., *Efficient reliability approximation method for traffic-path-based network*, "IEEE Transactions on Reliability" 2011, vol. 60, no. 2, pp. 460–469.
- [7] Cui L., Chen J., Wu B., *New interval availability indexes for Markov repairable systems*, "Reliability Engineering and System Safety" 2017, vol. 168, pp. 12–17.

- [8] Finkelstein M., Zarudnij V., *Laplace-transforms and fast-repair approximations for multiple availability and its generalizations*, “IEEE Transactions on Reliability” 2002, vol. 51, no. 2, pp. 168–176.
- [9] Miyaguchi T., Uneyama T., Akimoto T., *Brownian motion with alternately fluctuating diffusivity: Stretched-exponential and power-law relaxation*, “Physical Review E” 2019, vol. 100, no. 1, pp. 1–13
- [10] Yang D. Y., Wu C. H., *Evaluation of the availability and reliability of a standby repairable system incorporating imperfect switchovers and working breakdowns*, “Reliability Engineering and System Safety” 2021, vol. 207, pp. 1–16.
- [11] Yen T. C., Wang K. H., Wu C. H., *Reliability-based measure of a retrieval machine repair problem with working breakdowns under the F-policy*, “Computers and Industrial Engineering” 2020, vol. 150, pp. 1–11.
- [12] Durán O., Aguilar J., Capaldo A., Arata A., *Fleet resilience: evaluating maintenance strategies in critical equipment*, “Applied Sciences” 2021, vol. 11, no. 1, pp. 1–17. DOI:
- [13] Żurek J., Zieja M., Ziółkowski J., *The analysis of the helicopter technical readiness by means of the Markov processes*, “Proceedings IRF2018: 6th International Conference Integrity-Reliability-Failure”, Lisbon, 2018, pp. 1–14.
- [14] Ahmed S., Dey K., *Resilience modeling concepts in transportation systems: a comprehensive review based on mode, and modeling techniques*, “Journal of Infrastructure Preservation and Resilience” 2020, vol. 1, no. 8, pp. 1–20.
- [15] Yeh C. T., Lin Y. K., Yeng L. C. L., Huang P. T., *Reliability evaluation of a multistate railway transportation network from the perspective of a travel agent*, “Reliability Engineering and System Safety” 2021, vol. 214, pp. 1–34.
- [16] Chow A., Su Z., Liang E., Zhong R., *Adaptive signal control for bus service reliability with connected vehicle technology via reinforcement learning*, “Transportation Research Part C: Emerging Technologies” 2021, vol. 129, pp. 1–21.
- [17] Fałdziński M., Osińska M., Zalewski W., *Extreme value theory in application to delivery delays*, “Entropy” 2021, vol. 23, no. 7, pp. 1–14.
- [18] Barabino B., DI Francesco M., *Diagnosis of Irregularity Sources by Automatic Vehicle Location Data*, “IEEE Intelligent Transportation Systems Magazine” 2021, vol. 13, no. 2, pp. 152–165.
- [19] Małachowski J., Ziółkowski J., Lęgas A., Oszczypała M., Szkutnik-Rogoż J., *Application of the Bloch-Schmigalla Method to Optimize the Organization of the Process of Repairing Unmanned Ground Vehicles*, “Advances in Science and Technology. Research. Journal” 2020, vol. 14, no. 3, pp. 39–48.

- [20] Tumsekcali E., Ayyildiz E., Taskin A., *Interval valued intuitionistic fuzzy AHP-WASPAS based public transportation service quality evaluation by a new extension of SERVQUAL Model: P-SERVQUAL 4.0*, "Expert Systems with Applications" 2021, vol. 186, pp. 1–17.

## **Modelling of the Military Vehicles Operation Process in Terms of Readiness**

### **Abstract**

Modelling technical readiness is particularly important in systems characterised by randomness in the occurrence of tasks. Programming languages and tools are increasingly used in reliability analysis, enabling not only the proper functioning of the tested process, but also supporting the decision-making process enabling quick analysis and interpretation of incoming data streams. The aim of the article was to determine a reliability measure, which was the technical readiness coefficient of the fleet of trucks carrying out transport tasks on the territory of the Republic of Poland, using the R programming language. The literature on the most important theorems and definitions concerning the theory of reliability was reviewed to develop assumptions and a mathematical model of the exploitation process of using means of transport. As a result of the calculations and analyzes carried out, it was found that the technical readiness coefficient equal to 96.63(%) proves that the exploitation process of the available fleet of vehicles was correctly organised.

*Keywords: operation system, reliability, readiness.*



**Joanna Tomasik<sup>1</sup>**

## **Analiza wpływu liczby przywiązanych ligandów na budowę włochatej nanocząstki przy wykorzystaniu symulacji komputerowych**

### **Streszczenie**

Włochate nanocząstki, dzięki swoim właściwościom wynikającym z niewielkich rozmiarów oraz możliwości przyłączania polimerowych ligandów o różnych właściwościach, zyskują coraz większą popularność w świecie medycyny i kosmetologii. Znajdują również zastosowanie w kolejnych obszarach przemysłu. Coraz szersze stosowanie nanostruktur oraz tworzenie nowych nanocząstek o niespotykanej wcześniej budowie jest szansą na rozwój różnych dziedzin nauki, jednak nie możemy zapominać o płynących ze specyficznej budowy nanocząstek zagrożeniach dla człowieka i środowiska. Symulacje komputerowe metodą dynamiki molekularnej pozwalają na obserwację układu w czasie, dzięki czemu możemy analizować, w jaki sposób układ dążył do równowagi. Przeprowadzanie symulacji pozwala zaoszczędzić czas oraz pieniądze. Wyniki symulacji pokazały, że w zależności od liczby przywiązanych do rdzenia polimerowych ligandów, finalna struktura włochatej nanocząstki zmienia się. Dla małej liczby przywiązanych ligandów, nanocząstka przyjmuje formę cząstek Janusa, natomiast dla większej liczby – strukturę core-shell. Na podstawie przeprowadzonych symulacji można stwierdzić, że zmieniając podstawowe parametry włochatych nanocząstek, takie jak liczba przywiązanych łańcuchów, możemy tworzyć nanocząstki o pożądanym właściwościach.

*Słowa kluczowe: nanocząstki, włochate nanocząstki, symulacje komputerowe, dynamika molekularna.*

### **Wstęp**

Nanocząstki są to cząstki, których przynajmniej jeden z wymiarów ma poniżej 100 nm. Charakteryzują się one wysokim stosunkiem powierzchni do objętości. Włochate nanocząstki (HNPs, ang. *Hairy Nanoparticles*) składają się z rdzenia oraz przyłączonych do niego ligandów, na przykład polimerowych. Struktura, jaką przyjmą nanocząstki, wpływa na właściwości fizyczne oraz mechaniczne, co przekłada się na ich interakcje w roztworze [1]. Jedną z możliwych struktur, którą przybierają włochate nanocząstki, są cząstki Janusa. Jest to bardzo interesująca

---

<sup>1</sup> Joanna Tomasik, Wydział Podstaw Techniki, Politechnika Lubelska, s101049@pollub.edu.pl.



struktura, ponieważ składa się z dwóch lub więcej części o zróżnicowanych właściwościach fizycznych i chemicznych [2].

Włochate nanocząstki są stosowane w medycynie, fotowoltaice, sensorach oraz katalizie. Najważniejszym kierunkiem zastosowań włochatych nanocząstek jest medycyna. Kluczowymi właściwościami, dającymi możliwość rozwoju w tej dziedzinie, jest wysoki stosunek powierzchni do objętości oraz możliwość przenikania przez błony biologiczne [3]. Obecnie czynnikiem, który hamuje rozwój nanotechnologii w medycynie, jest obawa przed skutkami ubocznymi, w tym cytotoksycznym działaniem na zdrowe i poprawnie funkcjonujące komórki [4, 5]. Dlatego leczenie, w którym wykorzystuje się nanocząstki, najczęściej stosuje się podczas terapii nowotworów, kiedy to inne metody leczenia nie dają szans na odzyskanie zdrowia. Efekt toksyczny nanocząstek może być spowodowany unikalnymi właściwościami fizycznymi i chemicznymi, między innymi: wielkością, kształtem, powierzchnią właściwą czy obecnością grup aktywnych na powierzchni [5]. Aktualne dane dotyczące toksyczności nanocząstek sugerują, że potrzebujemy więcej badań na ich temat, aby lepiej poznać ich mechanizmy toksyczności [6].

## 1. Model i metoda symulacji

Obiektem badań była pojedyncza nanocząstka zbudowana z kulistego rdzenia (ang. *core*, indeksy  $c$ ) o średnicy  $\sigma_c$ , pokrytego łańcuchami polimerowymi złożonymi z sześciu ( $M = 6$ ) stygnie połączonych kulistych segmentów (ang. *segments*, indeksy „s”) o średnicy  $\sigma_s$ . Pierwszy segment łańcucha był na stałe związany z rdzeniem w odległości  $\sigma_{cs} = 0,5(\sigma_c + \sigma_s)$ , poprzez potencjał harmoniczny:  $u_{cs} = k_{cs}(r - \sigma_{cs})^2$ . Łańcuchy polimerowe mogły swobodnie przemieszczać się po powierzchni rdzenia [7]. Spójność łańcuchów zapewniał potencjał harmoniczny:  $u_{ss} = k_{ss}(r - \sigma_s)^2$ . Oddziaływania w układzie modelowano przesuniętym i obciążonym potencjałem Lennarda–Jonesa 12/6 przedstawia równanie 1.1.

$$u^{(ij)} = \begin{cases} 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r} \right)^{12} - \left( \frac{\sigma_{ij}}{r} \right)^6 \right] + \Delta_{(r)}^{(ij)} & r < r_{cut}^{(ij)} \\ 0 & \text{w pozostałych stałych przypadkach} \end{cases} \quad (1)$$

gdzie:

$$\Delta u_{(r)}^{(ij)} = \frac{-(r - r_{cut}^{(ij)}) \partial u^{(ij)}(r_{cut}^{(ij)})}{\partial r},$$

$r_{cut}^{(ij)}$  oznacza zasięg działania potencjału,  $\sigma_{ij} = 0,5(\sigma_i + \sigma_j)$  ( $i, j = c, s$ ) oraz  $\varepsilon_{ij}$  jest parametrem charakteryzującym siłę oddziaływań pomiędzy elementami  $i$  oraz  $j$ .

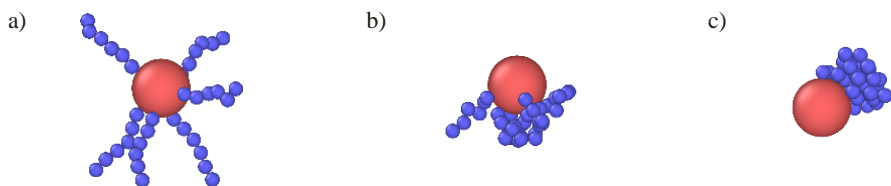
$R_{cut}$  dla oddziaływań przyciągających wynosiło  $r_{cut}^{(ij)} = 3,5 \sigma_{ij}$ , dla oddziaływań odpychających  $r_{cut}^{(ij)} = \sigma_{ij}$ . Przyjmowano, że układ znajduje się w równowadze, jeśli jego energia całkowita osiągała stałą wartość, przy której pojawiały się niewielkie fluktuacje wokół jej wartości średniej. Podczas równowagowania analizowano czasową ewolucję układu.

Symulacje zostały wykonane metodą dynamiki molekularnej przy użyciu pakietu symulacyjnego LAMMPS. Badano zachowanie pojedynczej włochatej nanocząstki, gdzie łańcuchy polimerowe mogły przemieszczać się swobodnie po powierzchni rdzenia [8].

## 2. Wyniki symulacji

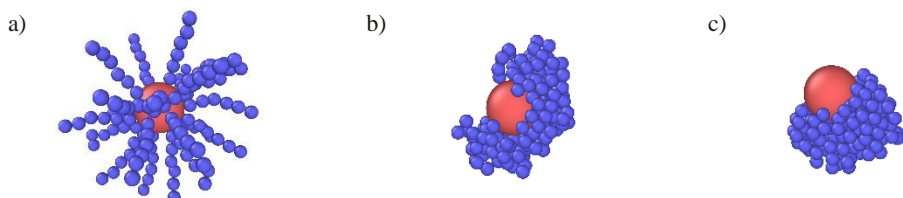
Symulacje przeprowadzono dla pięciu różnych liczb przywiązanych do powierzchni ligandów, badane nanocząstki posiadały  $f = 6, 12, 24, 36, 48$  przywiązanych ligandów. Rysunki 1a–1c przedstawiają nanocząstki z  $f = 6$  przywiązanych ligandów, rysunki 2a–2c ukazują nanocząstki z  $f = 24$  przywiązanych ligandów, natomiast rysunki 3a–3c nanocząstki z  $f = 48$  przywiązanych ligandów. Na rysunkach 1a, 2a oraz 3a widzimy włochate nanocząstki po  $10^2$  kroków symulacyjnych – konfiguracje startowe. Na rysunkach 1b, 2b oraz 3b przedstawiono nanocząstki po  $10^4$  kroków symulacyjnych. Zaś na rysunkach 1c, 2c oraz 3c ukazano nanocząstki w konfiguracji równowagowej, po  $10^6$  kroków symulacyjnych. Czerwona, największa kula w centrum układu jest rdzeniem, natomiast niebieskie, mniejsze kule są segmentami przywiązanych ligandów.

Na rysunkach 1a–3a można zauważyć, że w konfiguracji startowej przywiązane łańcuchy rozłożyły się równomiernie na powierzchni rdzenia oraz są w większości całkowicie rozprostowane. Obserwując ewolucję czasową, po  $10^4$  kroków symulacyjnych, możemy zauważyć, że dla włochatych nanocząstek pokrytych mniejszą ( $f = 6, 24$ ) liczbą łańcuchów, przemieszczają się one gromadząc się blisko siebie, przez co zwiększa się asymetria struktury (rys. 1b, 2b). Natomiast dla nanocząstki pokrytej 48 łańcuchami w dalszym ciągu obserwuje się ich równomierne rozmieszczenie (rys. 3b). We wszystkich przypadkach następuje zwijanie się łańcuchów (rys. 1b, 2b, 3b). Po  $10^6$  kroków symulacyjnych obserwuje się utworzenie dwóch typów struktur. Dla nanocząstek pokrytych liczbą łańcuchów  $f = 6$  oraz  $f = 24$  widoczne jest zwinięcie łańcuchów oraz zgrupowanie po jednej stronie oraz odsłonięcie rdzenia (rys. 1c, 2c). Utworzona w ten sposób struktura jest nazywana cząstkami Janusa. Natomiast dla nanocząstki o  $f = 48$  przywiązanych ligandów obserwuje się duże zwinięcie łańcuchów, które całkowicie pokrywają powierzchnię rdzenia, w tym przypadku następuje utworzenie struktury *core-shell* (z ang. rdzeń-powłoka) (rys. 3c).



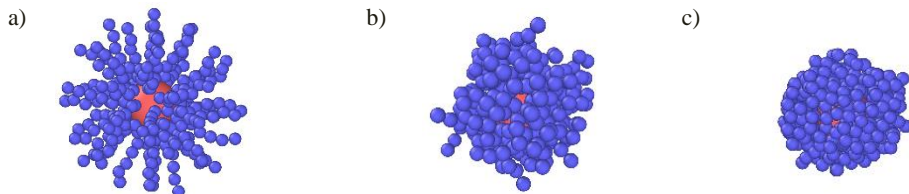
**Rysunek 1.** Przykładowe konfiguracje nanocząstki dla  $f = 6$  przywiązanych ligandów – a) po  $10^2$  krokach symulacyjnych; b) po  $10^4$  krokach symulacyjnych; c) po  $10^6$  krokach symulacyjnych

Źródło: opracowanie własne.



**Rysunek 2.** Przykładowe konfiguracje nanocząstki dla  $f = 24$  przywiązanych ligandów – a) po  $10^2$  krokach symulacyjnych; b) po  $10^4$  krokach symulacyjnych; c) po  $10^6$  krokach symulacyjnych

Źródło: opracowanie własne.

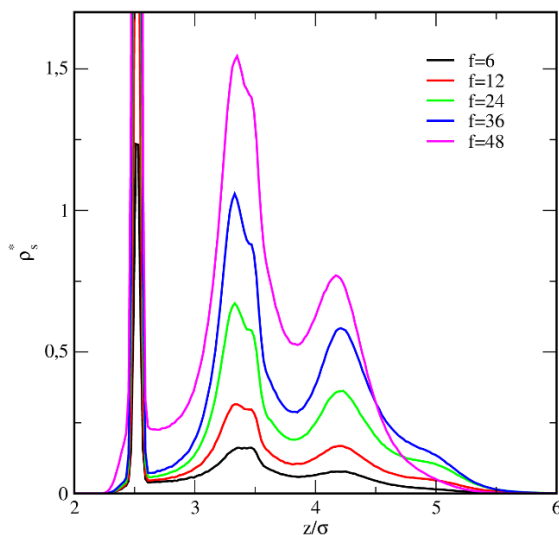


**Rysunek 3.** Przykładowe konfiguracje nanocząstki dla  $f = 48$  przywiązanych ligandów – a) po  $10^2$  krokach symulacyjnych; b) po  $10^4$  krokach symulacyjnych; c) po  $10^6$  krokach symulacyjnych

Źródło: opracowanie własne.

Aby dokładnie przeanalizować strukturę powłoki utworzonej przez polimerowe ligandy, obliczono jednowymiarowe profile gęstości segmentów ligandów  $\rho_s^*$  (rys. 4). Dla wszystkich badanych układów widzimy wąski, ostry pik przy  $\frac{z}{\sigma} = 2,52$ . Pik ten odpowiada segmentom, które przywiązane są bezpośrednio do powierzchni rdzenia. Następnie dla nanocząstek o liczbie przywiązanych ligandów  $f = 6, 12, 24, 36$  przy  $\frac{z}{\sigma} = 3,34$  oraz  $\frac{z}{\sigma} = 4,22$  obserwujemy dwa szerokie piki

odpowiadające kolejnym warstwom segmentów. Dla  $\frac{z}{\sigma} = 4,97$  obserwujemy przebiegię świadczące o tworzeniu się kolejnej warstwy segmentów. Dla nanocząstki o liczbie przywiązanych ligandów  $f = 48$ , poza pierwszym ostrym pikiem, widzimy również 2 szersze. Maksimum pierwszego znajdują się w tym samym miejscu, co dla pozostałych wartości  $f$ , natomiast drugi jest nieznacznie przesunięty w lewo ( $\frac{z}{\sigma} = 4,18$ ). Maksymalny zasięg profilu jest podobny dla wszystkich liczb przywiązanych ligandów, świadczy to o większym zwijaniu się łańcuchów dla nanocząstek o większej liczbie przywiązanych ligandów, a w konsekwencji niższą wysokością polimerowej powłoki.



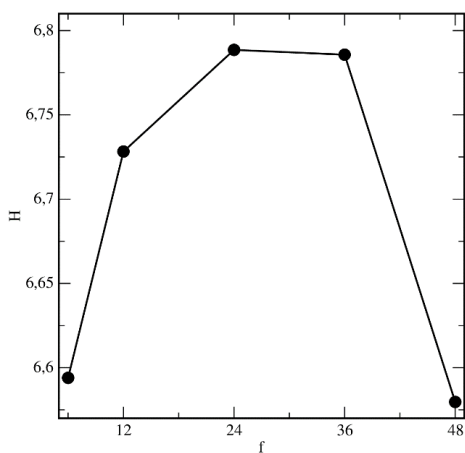
Rysunek 4. Profile gęstości segmentów przywiązanych ligandów

Źródło: opracowanie własne.

Na rysunku 5 przedstawiono, że początkowo wysokość korony polimerowej rośnie wraz ze wzrostem liczbie przywiązanych ligandów. Dla  $f = 24$  obserwujemy pewną stabilizację. Natomiast dla nanocząstki o największej liczbie przywiązanych ligandów ( $f = 48$ ) obserwujemy gwałtowny spadek. Oznacza to, że łańcuchy „kładą się” na powierzchni rdzenia, formując „polimerowy płaszcz” wokół rdzenia (rys. 3c).

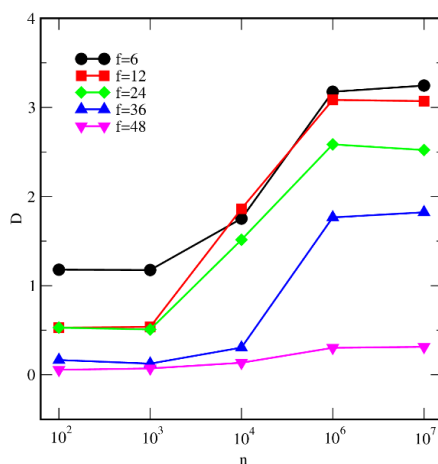
Na rysunku 6 przedstawiono ewolucję czasową masowego momentu dipolowego ( $D$ ). Masowy moment dipolowy jest odległością pomiędzy środkiem rdzenia nanocząstki a środkiem masy chmury segmentów przywiązanych ligandów. Dla cząstek idealnie symetrycznych wartość ta wynosi zero, natomiast im bardziej asymetryczna cząstka, tym ta wartość rośnie. Wzrost masowego momentu dipolowego jest spowodowany gromadzeniem się ligandów po jednej stronie rdzenia,

równocześnie odsłaniając jego drugą część. To powoduje wzrost asymetryczności włochatej nanocząstki. Na wykresie widzimy, że początkowo ( $10^2$  kroków symulacyjnych) dla wszystkich badanych liczb przywiązanych ligandów wartość masowego momentu dipolowego przyjmuje stosunkowo niskie wartości. Po  $10^3$  kroków symulacyjnych wartości te w zasadzie się nie zmieniają. Informuje nas to o tym, że pomimo zwijania się łańcuchów w tym momencie symulacji ligandy nie kumulowały się jeszcze po jednej stronie rdzenia. Po  $10^4$  kroków symulacyjnych obserwujemy wzrost wartości masowego momentu dipolowego dla wszystkich badanych układów poza włochatą nanocząstką o największej liczbie przywiązanych ligandów. Dla nanocząstek o 6 oraz 24 przywiązanych ligandach obserwujemy przesuwanie się ligandów oraz odsłanianie rdzenia (rys. 1b, 2b). Po  $10^6$  kroków symulacyjnych wartość masowego momentu dipolowego jeszcze bardziej rośnie, co świadczy o tym, że włochate nanocząstki stają się coraz bardziej niesymetryczne (rys 1c, 2c). W przypadku włochatej nanocząstki o  $f=48$  nie obserwujemy tak gwałtownej zmiany, a wzrost jest nieznaczny zarówno po  $10^4$  kroków symulacyjnych, jak i  $10^6$ .



**Rysunek 5. Zależność wysokości korony polimerowej (H) od liczby przywiązanych ligandów (f)**

Źródło: opracowanie własne.



**Rysunek 6. Ewolucja czasowa masowego momentu dipolowego (D)**

Źródło: opracowanie własne.

Obserwując końcowe wartości masowego momentu dipolowego możemy zauważyć, że jest on najwyższy (najbardziej asymetryczna włochata nanocząstka) dla włochatej nanocząstki o najmniejszej liczbie przywiązanych ligandów spośród badanych układów. Dodatkowo wraz ze wzrostem liczby przywiązanych ligandów, masowy moment dipolowy zmniejsza się. Dla układów o  $f=6, 12, 24, 36$  obserwujemy dużą asymetrię, wyższy masowy moment dipolowy oraz strukturę

cząstek Janusa. Natomiast dla nanocząstki z  $f = 48$  przywiązanych ligandów masowy moment dipolowy jest bliski zeru, a więc jest praktycznie symetryczna, a włochata nanocząstka przyjmuje strukturę core-shell.

### 3. Wnioski

Zbadano pięć rodzajów włochatych nanocząstek, pokrytych krótkimi ( $M = 6$ ), homogenicznymi, idealnie giętkimi łańcuchami polimerowymi, które mogły przemieszczać się po powierzchni rdzenia. Nanocząstki różniły się liczbą przywiązanych do rdzenia ligandów ( $f = 6, 12, 24, 36, 48$ ). Badania zostały przeprowadzone za pomocą symulacji komputerowych metodą dynamiki molekularnej przy użyciu pakietu symulacyjnego LAMMPS.

Z przeprowadzonych symulacji, obliczonych profili gęstości, zależności wysokości korony polimerowej od liczby przywiązanych ligandów oraz po analizie ewolucji czasowej masowego momentu dipolowego wynika, że w zależności od liczby przyłączonych łańcuchów finalna struktura różni się. Nanocząstki o mniejszej liczbie homogenicznych łańcuchów, w odpowiednich warunkach, przyjmują strukturę cząstek Janusa. Dla cząstek o większej liczbie łańcuchów obserwujemy zwijanie się łańcuchów, które, kładąc się na powierzchni rdzenia, tworzą strukturę *core-shell* (rdzeń-płaszcz). Dla wszystkich liczb przywiązanych ligandów obserwujemy zwijanie się łańcuchów, jednak dla mniejszych liczb ( $f = 6, 12, 24, 36$ ) łańcuchy te dodatkowo kumulują się po jednej stronie rdzenia, odsłaniając go z drugiej strony. Tworzenie się różnych struktur zależy od wielu czynników, jednak zawsze układ dąży do najkorzystniejszej energetycznie konfiguracji. W tym przypadku zmiana tylko jednego parametru spowodowała zmiany energii, które pozwoliły na stworzenie się dwóch różnych struktur. Dalsze badania pozwolą na dokładniejsze poznanie włochatych nanocząstek oraz struktur, jakie mogą utworzyć.

Badania zostały przeprowadzone w ramach pracy licencjackiej na Uniwersytecie Marii Curie-Skłodowskiej w Katedrze Chemii Teoretycznej (wówczas Zakładzie Modelowania Procesów Fizykochemicznych) pod kierunkiem dra hab. Tomasza Staszewskiego.

### Literatura

- [1] Staszewski T., *Structural Changes in Hairy Nanoparticles – Insights from Molecular Simulations*, “The Journal of Physical Chemistry C” 2020, vol. 124, no. 49, pp. 27118–27129.
- [2] Su H., Hurd Price, Jing L., Tian Q., Liu J., Qian K., *Janus particles: design, preparation, and biomedical applications*, “Materials Today Bio” 2019, vol. 4, pp. 1–2.

- [3] Baran A., *Nanotechnology: legal and ethical issues*, "Engineering Management in Production and Services" 2016, vol. 8, no. 1, 2016, pp. 47–54.
- [4] Ajdary M., Moosavi M. A., Rahmati M., Falahati M., Mahboubi M., Mandegary A., Jangjoo S., Mohammadinejad R., Varma R. S., *Health Concerns of Various Nanoparticles: A Review of Their in Vitro and in Vivo Toxicity*, "Nanomaterials" 2018, vol. 8, pp. 2–4.
- [5] Sukhanova A., Bozrova S., Sokolov P., Berestovoy M., Karaulov A., Nabiev I., *Dependence of Nanoparticle Toxicity on Their Physical and Chemical Properties*, "Nanoscale Res Letters" 2018, vol. 13, pp. 2–4.
- [6] Najahi-Missaoui W., Arnold R. D., Cummings B. S., *Safe Nanoparticles: Are We There Yet?*, "International journal of molecular sciences" 2020, vol. 22, no. 1, pp. 1–3.
- [7] Borówko M., Rżysko W., Sokołowski S., Staszewski, *Self-assembly of hairy disks in two dimensions - insights from molecular simulations*, "Soft Matter" 2018, vol. 14, pp. 3115–3126.
- [8] J. Tomasiak, *Wpływ liczby łańcuchów na strukturę włochatej nanocząstki*, praca licencjacka, Uniwersytet Marii Curie-Skłodowskiej, 2020.

## The impact of a number of chains on the structure of hairy nanoparticle using computer simulations

### Abstract

Hairy nanoparticles, due to their properties resulting from their small size and the ability to attach polymer ligands with various properties, are gaining more and more popularity in the world of medicine and cosmetology. They are also used in other areas of industry. The increasing use of nanostructures and the creation of new nanoparticles with an unprecedented structure is an opportunity for the development of various fields of science, but we must not forget about the threats to humans and the environment that result from the specific structure of nanoparticles. Molecular dynamics computer simulations allow the observation of a system over time, which allows us to analyze how the system tended to equilibrium. Simulations save time and money. The simulation results showed that depending on the amount of polymer ligands bound to the core, the final structure of the hairy nanoparticle changes. For a small number of bound ligands, the nanoparticle takes the form of Janus particles, while for a larger number - a core-shell structure. Based on the simulations, it can be concluded that by changing the basic parameters of hairy nanoparticles, such as the number of attached chains, we can create nanoparticles with the desired properties.

*Keywords: nanoparticles, hairy nanoparticles, computer simulations, molecular dynamics.*