

POLITECHNIKA WARSZAWSKA

DYSCYPLINA NAUKOWA INFORMATYKA TECHNICZNA I

TELEKOMUNIKACJA

DZIEDZINA NAUK INŻYNIERYJNO-TECHNICZNYCH

Rozprawa doktorska

mgr inż. Tomasz Wojciech Nowak

**Wyznaczanie izolacji w sieciach plastrowych z wykorzystaniem
skierowanych liczb rozmytych**

Promotor

prof. dr hab. inż. Zbigniew Kotulski

WARSZAWA 2023

*Dziękuję prof. Z. Kotulskiemu za wiele inspiracji,
kochanej żonie Sylwii za cierpliwość,
a pracę dedykuję Wiktorkowi, Laurce i A.W.*

STRESZCZENIE

W niniejszej rozprawie doktorskiej przedstawiono problematykę wyznaczania izolacji w sieciach plastrowych z wykorzystaniem skierowanych liczb rozmytych. Pierwsza część pracy poświęcona została przedstawieniu stanu wiedzy o sieciach plastrowych, potrzebie rozpatrywania poziomu izolacji oraz o sposobach zapewnienia izolacji w takich sieciach. Przedstawiono szeroki zakres technik zapewniania izolacji, zarówno programowych, jak i sprzętowych. Zaprezentowano wysokopoziomowy model sieci plastrowej, wraz z najważniejszymi pojęciami związanymi z sieciami plastrowymi. Opisano powody, dla których istotne jest określanie poziomu izolacji i zarządzanie nim, z perspektywy dostawcy usług sieciowych, operatora telekomunikacyjnego oraz użytkownika.

W dalszej części pracy zaprezentowano algorytm wyznaczania izolacji wykorzystujący do obliczeń liczby rzeczywiste, który następnie rozszerzany jest do postaci wykorzystującej skierowane liczby rozmyte. W celu rozwiązania postawionego problemu, przedstawiono sposób budowy skierowanych liczb rozmytych na podstawie zebranego ciągu danych pomiarowych oraz metodę poszukiwania skierowanej liczby właściwej najbliższej zadanej liczbie niewłaściwej, aby uzyskać obiekt matematyczny o pożądanym cechach. Kolejna część pracy przedstawia model probabilistyczny wyznaczania izolacji, który podobnie jak model wykorzystujący skierowane liczby rozmyte, bazuje na algorytmie zdefiniowanym dla liczb rzeczywistych.

Niniejsza dysertacja zawiera ponadto opis izolacji usług, który stanowi trzeci sposób modelowania izolacji zawarty w tej pracy, o węższym zakresie stosowalności niż dwa poprzednie. Jest to model podchodzący do zagadnienia izolacji w inny sposób niż model probabilistyczny i wykorzystujący skierowane liczby rozmyte, a jednocześnie prostszy w opisie.

Słowa kluczowe: sieci plastrowe, izolacja w sieciach teleinformatycznych, bezpieczeństwo sieci teleinformatycznych, skierowane liczby rozmyte, sieci teleinformatyczne, 5G

ABSTRACT

The dissertation presents the problem of determining isolation levels in sliced networks using ordered fuzzy numbers. The first part of the work describes the state of the art about network slicing, the need to consider the isolation level, and ways to ensure isolation in such networks. A wide range of isolation techniques, both software, and hardware, are presented. A high-level model of network slicing is described, along with the most important concepts related to sliced networks. It indicates reasons why it is important to determine and manage the isolation level from the perspective of the network service provider, the network operator, and the end user.

In the further part of the document, an algorithm for determining isolation using real numbers for calculations is presented, which is then enhanced with ordered fuzzy numbers. To solve the problem, a way to construct ordered fuzzy numbers based on the collected sequence of measurement data is presented. Author shows a method of searching for a proper ordered fuzzy number closest to a given improper fuzzy number to obtain a mathematical object with the desired features. The next part of the document presents a probabilistic model of isolation level, which, like the model using ordered fuzzy numbers, is based on an algorithm defined for real numbers.

This dissertation also contains a description of service isolation, which is the third way of modelling isolation contained in this work, applicable to a narrower range of use cases. It is a model that describes the isolation in a different way than the probabilistic model, or ordered fuzzy numbers-based model, and at the same time simpler to describe.

Keywords: network slicing, isolation in ICT networks, security of ICT networks, ordered fuzzy numbers, ICT networks, 5G

SPIS TREŚCI

1. Wstęp	11
1.1. Współczesne sieci mobilne	12
1.2. Teza pracy	18
1.3. Struktura pracy	19
2. Sieć plastrowa	21
2.1. Motywacja	21
2.2. Definicje sieci plastrowej	22
2.3. Podział horyzontalny i wertykalny	23
2.4. Pełne połączenie plastrowe	24
2.5. Zarządzanie czasem życia plastrów	26
3. Izolacja w sieci plastrowej	29
3.1. Motywacja	29
3.2. Bezpieczeństwo plastrów w pełnym połączeniu plastrowym	31
3.3. Typy izolacji	35
3.4. Podstawy modelowania izolacji	36
3.5. Istotność zagadnienia izolacji	40
3.6. Techniki izolacji	43
3.6.1. Izolacja w językach programowania	44
3.6.2. Izolacja wykorzystująca środowisko testowe	45
3.6.3. Izolacja wykorzystująca wirtualizację	45
3.6.4. Izolacja sprzętowa i izolacja poziomu jądra systemu operacyjnego	45
3.6.5. Separacja domen i procesów	46
3.6.6. Izolacja w środowisku chmurowym	48

3.6.7.	Techniki izolacji w sieci	49
3.6.8.	Izolacja w sieciach RAN i CN	51
4.	Algorytm wyznaczania izolacji	53
4.1.	Hipergrafowy model sieci	53
4.2.	Parametry i właściwości	55
4.3.	Podstawy matematyczne	58
4.4.	Zagadnienie porządku w kontekście izolacji	64
4.5.	Kroki algorytmu	65
5.	Skierowane liczby rozmyte	67
5.1.	Ogólne informacje o Skierowanych Liczbach Rozmytych	67
5.2.	Pojęcia i definicje matematyczne	67
6.	Model rozmyty	75
6.1.	Konstruowanie SLR na podstawie wyników pomiarów	75
6.2.	Problem niewłaściwych SLR	78
6.2.1.	Przybliżenie za pomocą właściwych SLR	78
6.2.2.	Warunki Karush-Kuhn-Tuckera	84
6.2.3.	Redukcja problemu QP do problemu LP	87
6.3.	Algorytm wyznaczania izolacji z wykorzystaniem SLR	94
6.4.	Porządek w modelu rozmytym	94
7.	Przykład obliczeniowy	103
8.	Model probabilistyczny	127
8.1.	Założenia modelu	127
8.2.	Model	127
8.2.1.	Funkcja minimum i maksimum	128
8.2.2.	Funkcja maksimum	133
8.2.3.	Estymator wartości oczekiwanej	135
8.2.4.	Wyznaczanie zbiorów	135
8.3.	Algorytm wyznaczania izolacji w modelu probabilistycznym	137

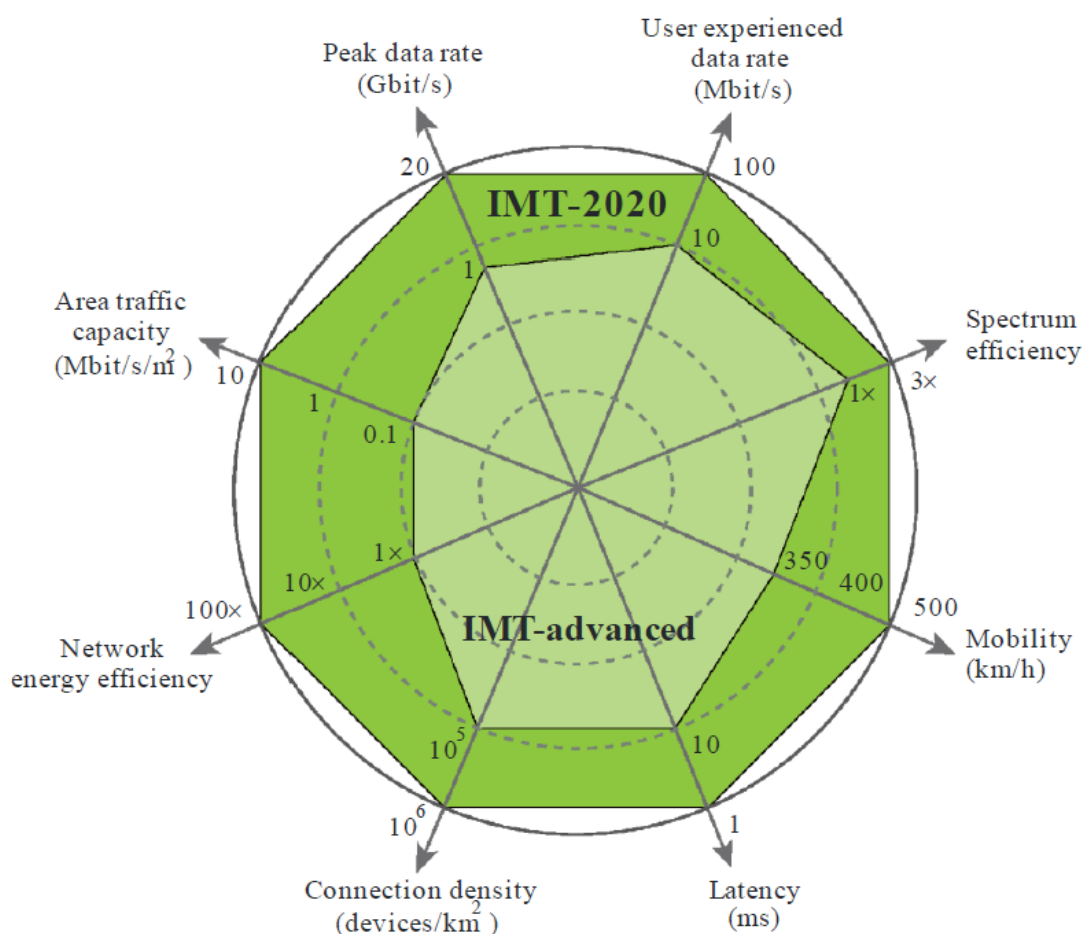
8.4. Przykład obliczeniowy w modelu probabilistycznym	137
9. Izolacja usług	143
9.1. Przedstawienie zagadnienia	143
9.2. Model matematyczny izolacji usług	144
9.3. Wnioski	146
10. Podsumowanie i wnioski	147
10.1. Podsumowanie wyników badań	147
10.2. Wnioski i kierunki dalszych badań	150
10.3. Przegląd opublikowanych wyników badań	151
BIBLIOGRAFIA	157
Dodatek A. Uzupełnienie do modelu probabilistycznego	177

1. Wstęp

Nowoczesne sieci telekomunikacyjne stoją przed bardzo skomplikowanymi wyzwaniami technicznymi. Cele stawiane sieciom, w tym sieciom mobilnym, już dawno wykroczyły poza tylko i wyłącznie aspekt realizacji połączeń telefonicznych, adaptując do portfolio świadczonych usług telekomunikacyjnych takie usługi jak komunikacja tekstowa (SMS - *Short Message Service*), graficzna (MMS - *Multimedia Messaging Service*), czy wreszcie dostęp do Internetu. Ta ostatnia usługa stała się kamieniem milowym w świecie telekomunikacji mobilnej - pozwoliła nie tylko na świadczenie ściśle zdefiniowanych usług, ale dała również możliwość skorzystania z bardzo szerokiej gamy usług internetowych, które co do zasady nie są znane bliżej operatorowi. Jest to pewnego rodzaju meta-usługa, której celem jest świadczenie dostępu do innych usług. W dzisiejszym Internecie powstało więcej tego typu meta-usług, które są wykorzystywane jako platformy usługowe; przykładową rodziną takich usług są sieci społecznościowe jak Facebook czy Instagram, które wyewoluowały z platform komunikacyjnych do bardzo rozbudowanych systemów, wykorzystywanych szeroko w *e-commerce* (sprzedaży internetowej), marketingu, czy też w przemyśle rozrywkowym.

Na przestrzeni lat sposób dostępu do Internetu za pośrednictwem sieci bezprzewodowych silnie ewoluował. Proces ten był sprzężony ze wzrostem intensywności użytkowania Internetu mobilnego, na którą się złożyły spadki cen usługi transferu danych oraz wprowadzenie na szeroką skalę urządzeń mobilnych, zdolnych do zaawansowanej interakcji z usługami internetowymi. Na skutek rozwoju usług internetowych i urządzeń mobilnych dostęp do Internetu (usług internetowych) stał się główną usługą świadczoną przez Operatora Telekomunikacyjnego (NO - ang. *Network Operator*). Ta strategiczna zmiana miała odzwierciedlenie w ewolucji architektury sieci mobilnej, szczególnie widoczna była w przypadku sieci 4G (ang. *4th Generation*), która traktowała dostęp do Internetu jako natywną usługę, a nie tylko pewien dodatek do usługi telefonii. Kolejnym krokiem miało być zwiększenie jakości świadczonej usługi, w ramach projektowanej sieci 5G (*5th generation*).

Enhancement of key capabilities from IMT-Advanced to IMT-2020

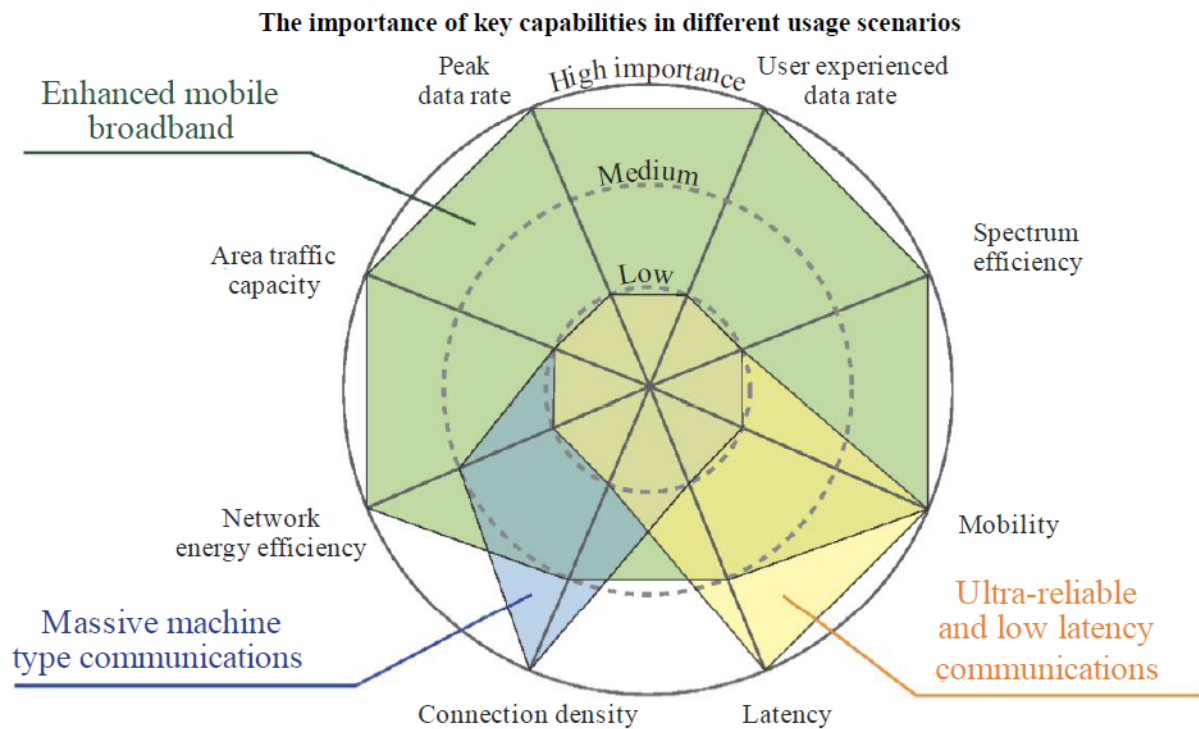


M.2083-03

Rysunek 1.1: Wymagania stawiane sieci 5G (IMT-2020 w nomenklaturze ETSI - European Telecommunications Standards Institute i ITU - *International Telecommunication Union*), w odniesieniu do wymagań wobec sieci 4G (IMT-advanced według ETSI i ITU); rysunek pochodzi z [100], zob. też [24], [36].

1.1. WSPÓŁCZESNE SIECI MOBILNE

Filozofią przyświecającą powstaniu sieci 5G jest w uproszczeniu 'więcej wszystkiego', co ma się przekładać na lepsze wyniki takich metryk jak QoE (*Quality of Experience*) czy QoS (*Quality of Service*), przy jednoczesnym zapewnieniu odpowiedniego poziomu bezpieczeństwa (*Quality of Security*) [59]). Założenia architektoniczne co do wymagań techniczno-biznesowych są przedstawione na rysunkach 1.1 i 1.2. Polepszenie jednocześnie tak wielu istotnych parametrów technicznych jest zadaniem bardzo trudnym - zmiany nie ograniczają się tutaj do wyprodukowania nowych urządzeń mobilnych (UE - *User Equipment*), ale dotyczą zarówno warstwę radiową



M.2083-04

Rysunek 1.2: Ważność kluczowych zdolności sieci w kontekście trzech podstawowych kategorii komunikacji. rysunek pochodzi z [100], zob. też [24], [36].

(RAN - *Radio Access Network*), jak i sieć szkieletową Operatora (CN - *Core Network*). Aby osiągnąć założone poziomy parametrów jakościowych, do wdrożenia sieci 5G zaproponowano użycie nowatorskich rozwiązań teleinformatycznych, m.in.

- SDN (*Software Defined Network*),
- NFV/VNF (*Network Function Virtualization/Virtual Network Functions*),
- koncept sieci plastrowych (ang. *Network Slicing*),
- MEC (*Mobile Edge Computing/Multi-access Edge Computing*),
- pikokomórki (*picocells*) w paśmie powyżej 6 GHz.

Partnerzy biznesowi Operatora Telekomunikacyjnego mogą mieć własne, specyficzne dla swojej branży i usług, wymagania i potrzeby związane z siecią 5G. Z perspektywy Operatora, partnerzy mogą należeć do istotnych sektorów gospodarki, nazywanych również sektorami wertykalnymi (ang. *verticals*). Przykładowymi sektorami wertykalnymi są [92]:

- Transport & Logistyka (*Transport & Logistics*),
- Ochrona Zdrowia (*Healthcare*),

- Handel (*Retail*),
- Sektor Finansowy (BFSI - *Banking, Financial Services, and Insurance*),
- Infrastruktura Krytyczna (*Critical Infrastructure*),
- Inteligentne Miasto (*Smart City*),
- Rolnictwo (*Agriculture*),
- Telekomunikacja (*Telecommunications*),
- Rozrywka (*Entertainment*),
- Rząd i Samorząd (*Authorities*),
- Edukacja (*Education*),
- Przemysł (*Manufacturing*).

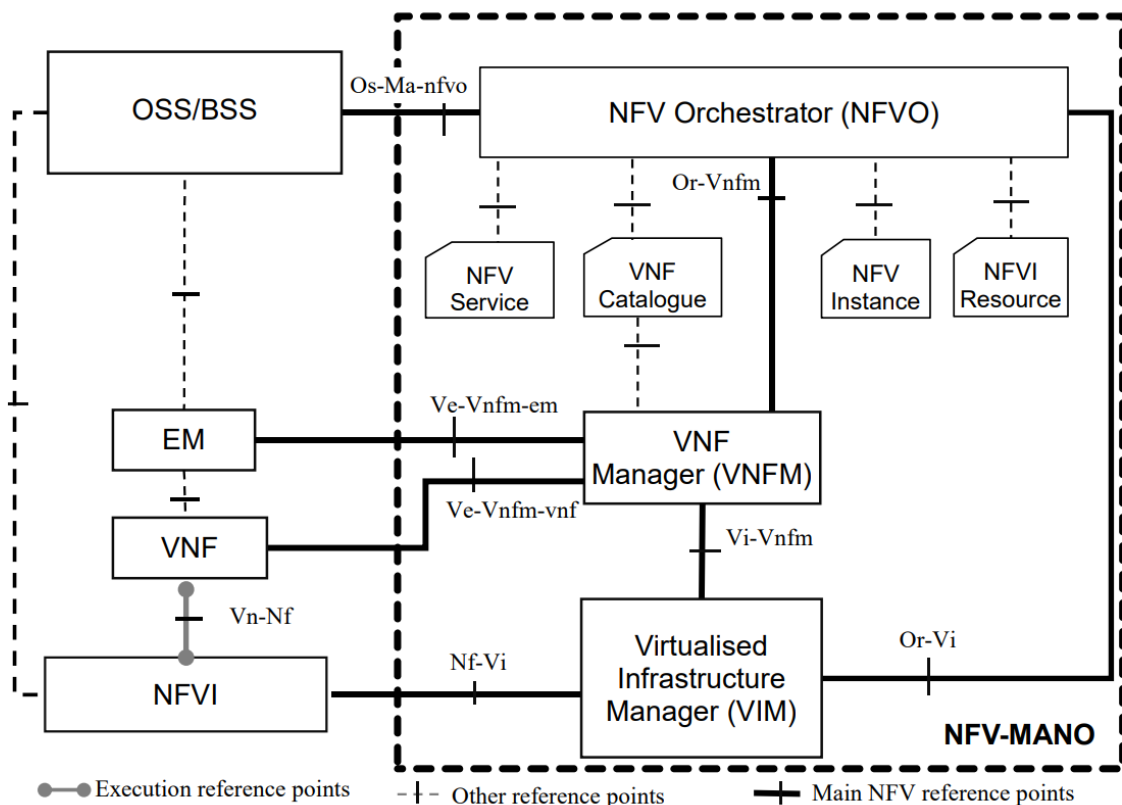
W pracach [85, 92] opisano szeroko przykłady zastosowania sieci 5G i 5G MEC w kontekście wymienionych sektorów wertykalnych. Niektóre z nich to:

- komunikacja typu *Mobile-to-Bank* (M2B) wspomagająca proces płatności,
- zdalne obserwowanie (monitorowanie) stanu zdrowia pacjentów,
- bezprzewodowe i zdalne operacje chirurgiczne,
- komunikacja w sytuacjach zarządzania kryzysowego,
- obserwacja i monitoring miejski,
- komunikacja związana z ruchem samochodowym (V2X - ang. *Vehicle-to-everything*),
- autonomiczne pojazdy i komunikacja autonomicznego pojazdu ze światem zewnętrznym.

Prace [37, 81, 87, 112] podniosły zagadnienie sformułowania zbioru wymagań jakościowych wobec sieci 5G, które są bardzo trudne lub bardzo drogie do zrealizowania w obrębie całej sieci, przez cały czas. Przykładowo, są to wymagania w łącznej postaci:

- przepływność binarna przynajmniej 300 Mbit/s,
- opóźnienie rzędu kilku milisekund,
- sieć dostępna dla 200 000 urządzeń na km²,
- 99.999% dostępności usług z takim poziomem wymagań jakościowych.

Wymienione wymagania są w różnym stopniu stawiane przez poszczególnych uczestników sektorów wertykalnych, ponieważ ich ważność dla realizacji poszczególnych scenariuszy biznesowych nie jest jednakowa. Przykładowo, dla komunikacji związanej z samochodami bardzo istotne jest niskie opóźnienie i wysoka dostępność usługi, co z kolei nie jest kluczowe dla scenariuszy związanych z rolnictwem. Łatwiej jest zapewnić opisany poziom jakościowy dostępu

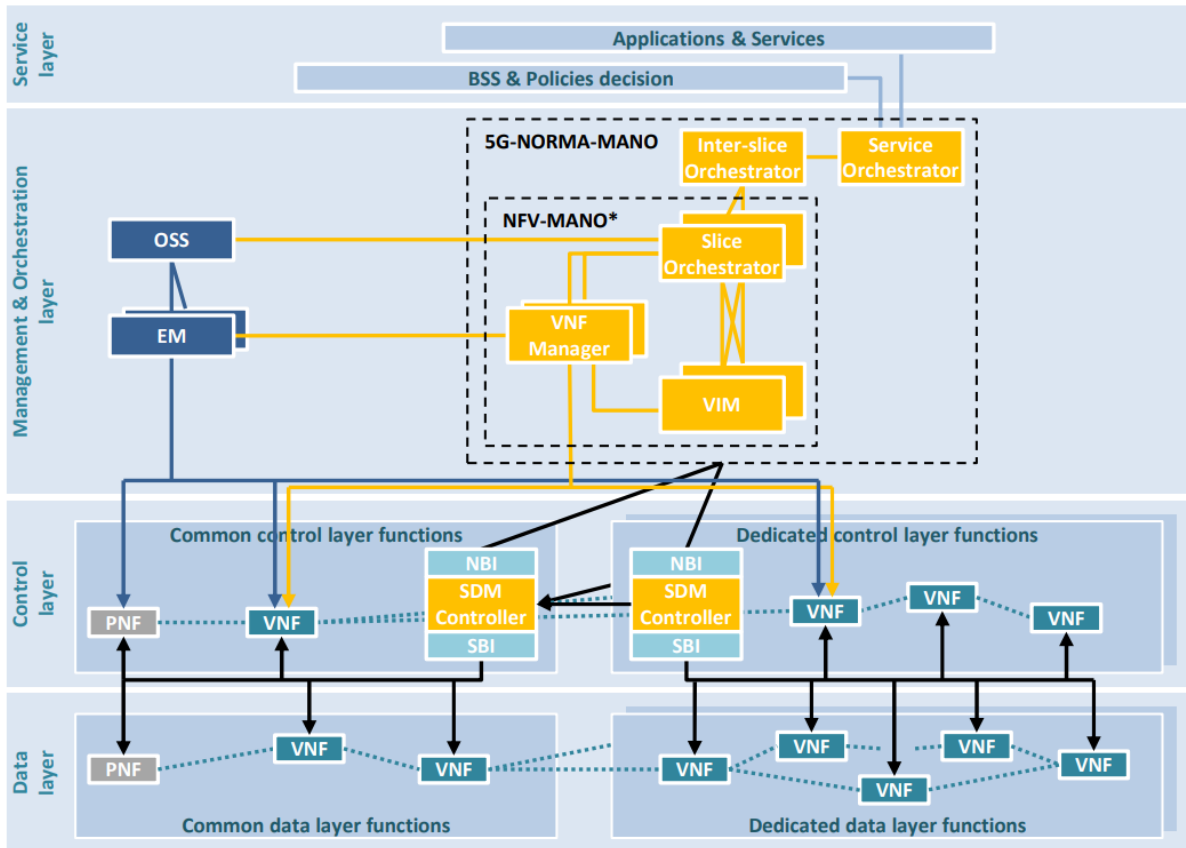


Rysunek 1.3: Model NFV-MANO (*Network Function Virtualization-Management and Orchestration*) zdefiniowany w [26], patrz też [25].

do usług, spełniając tylko wybrane wymagania, w ramach danego logicznego fragmentu sieci, który będzie służyć do realizacji pewnych usług sieciowych, podobnych do siebie pod kątem wymagań jakościowych wobec sieci [11, 87, 102]. Idea ta jest realizowana poprzez sieć plastrów, opisaną szerzej w rozdziale 2.

W sieci 5G położono duży nacisk na przeorganizowanie modelu zarządzania i wprowadzenia orkiestracji zasobami sieciowymi. Jeden z podstawowych modeli zarządzania i orkiestracji został zaproponowany przez ETSI i jest znany jako *ETSI NFV-MANO architectural framework* [25, 26]; został on zilustrowany na rysunku 1.3. Założeniem tego modelu było wsparcie wszystkich sieci wykorzystujących wirtualizację sieci, w tym także sieci 5G. Ze względu na wysoki poziom ogólności, nie jest to model wyczerpujący zagadnienie w przypadku sieci 5G i wobec tego wymaga uszczegółowienia pod kątem tej sieci.

Możliwe jest połączenie modelu ETSI NFV-MANO z podejściem SDN [80, 83]. W obecnej chwili tego typu rozwiązania wydają się naturalne i wpisują się w trendy występujące w świecie oprogramowania, aby używać oprogramowania do zarządzania infrastrukturą (systemy



Rysunek 1.4: Model NFV-MANO w uzupełnieniu 5G NORMA, zdefiniowany w [3].

typu Docker [20], Kubernetes [116], Infrastructure-as-Code [76]). Możliwe jest również zastosowanie tego modelu do budowy bardziej skomplikowanych systemów zarządzania w oparciu na hierarchii zarządzania lub orkiestracji [4]. Pozwala to na budowę systemu wielodomenowego, wieloinstancyjnego czy w oparciu na wielu dostawcach (*multi-domain, multi-tenant, multi-vendor*). W takiej rozbudowanej strukturze można również implementować koncepcję łańcucha plasterów (ang. *slice chaining*) [23, 58], które pozwala połączyć wiele sieci plasterowych między sobą aby stworzyć warunki wystarczające do realizacji usługi.

W pracy [114] poruszono aspekty zarządzania i orkiestracji w środowisku sieci wielodomenowych. W takich sieciach każda domena charakteryzuje się własnym zestawem wymagań i ograniczeń w zakresie bezpieczeństwa, technologii, aspektów prawnych. Aby zrealizować złożoną usługę, należy uzgodnić wspólne warunki dla wszystkich domen biorących udział w realizacji danej usługi. Praca [114] proponuje podejście wykorzystujące kryteria SLA (*Service Level Agreement*) do orkiestracji zasobami sieciowymi. W literaturze proponowane jest również inne podejście oparte na choreografii (np. w [16]), gdzie decyzja nie jest podejmowana w

scentralizowany sposób, jak w orkiestracji, ale w ramach węzłów sieciowych, uwzględniając ich własne, szeroko rozumiane interesy.

Model ETSI NFV-MANO nie przewiduje wprost wsparcia dla podziału sieci na fragmenty służące do realizacji różnych usług; realizacja takiego wymagania funkcjonalnego wymaga uzupełnienia modelu w pewnych aspektach. W pracy [66] zaproponowano rozszerzenie modelu o dodatkową usługę realizującą to wymaganie. Inne podejście zaproponowano w pracy [3], gdzie orkiestracja została rozszerzona o funkcje wspierające taki podział, w ramach nowych, dedykowanych orkiestratorów (ang. *Slice Orchestrator* i *Inter-Slice Orchestrator*). koncepcja ta została przedstawiona na rysunku 1.4.

Jednym z istotnych wymagań sieci 5G jest zapewnienie bardzo małych opóźnień, zarówno w jednym kierunku, jak i dwukierunkowo (RTT - *Round Trip Time*). W kontekście tej sieci oznacza to opóźnienia rzędu pojedynczych milisekund, zasadniczo w komunikacji bezprzewodowej. Odpowiedzią na takie wymaganie jest przeniesienie usług sieciowych bliżej klienta końcowego, aby zredukować fizyczny i topologiczny dystans między UE a serwerem świadczącym usługę. Istnieje kilka sposobów zrealizowania takiej architektury, m.in.:

- MEC,
- *Dev Computing*,
- *Fog Computing*,
- CDN (*Content Delivery Network*) - rozwiązanie znane z Internetu, które udostępnia popularne materiały dla użytkowników jak filmy wideo, *streaming*, popularne biblioteki programistyczne (np. skrypty napisane w języku JavaScript wykorzystywane na stronach WWW). Przeniesienie usługi bliżej klienta końcowego niesie ze sobą również pewne wyzwania techniczne, w tym związane z bezpieczeństwem takiej usługi. Zmiana miejsca realizacji usługi pociąga za sobą konieczność uwierzytelnienia i autoryzowania użytkownika w nowym miejscu [55, 84, 85, 108], Dodatkowo, niektóre usługi mogą wymagać bardzo szybkiego uwierzytelnienia i autoryzacji [85, 92], co stanowi kolejny przyczynek do zajęcia się podziałem sieci na części, aby mieć możliwość realizacji takich usług, nawet w ograniczonym zakresie. Podział na dostawców usług, dostawców sieci dostępowej i dostawców usługi MEC generuje wiele nowych wyzwań w dziedzinie uwierzytelnienia i autoryzacji, zostały one szerzej opisane i skatalogowane w pracy [54].

1.2. TEZA PRACY

W niniejszej pracy stawiamy następującą tezę:

Zarządzanie izolacją (i jej opis ilościowy) jest ważnym elementem zapewniania bezpieczeństwa usług w sieciach plastrowych. Możliwe jest zdefiniowanie pojęcia izolacji w sieciach plastrowych, które można opisać ilościowo za pomocą skierowanych liczb rozmytych. Tak zdefiniowany poziom izolacji można porównywać w ustalonym porządku.

W celu wykazania prawdziwości tezy, przeprowadzono następujące zadania badawcze.

- Jakościowa analiza zagadnienia bezpieczeństwa plastrów E2E (*End-To-End*) i ich izolacji w sieciach 5G i 5G MEC w celu ustalenia potrzeb ich opisu ilościowego - podrozdziały 2.4., 3.2. i 3.5. oraz uwzględnienie różnorodności wymagań dla wertikali - podrozdział 3.2. i praca [92].
- Badanie zagadnienia izolacji w systemach informacyjnych - analiza technik wykorzystywanych do zapewniania izolacji i ich usystematyzowanie - podrozdział 3.6..
- Stworzenie modelu izolacji z wykorzystaniem liczb rzeczywistych - publikacje [56, 57], przytoczony w rozdziale 4.
- Analiza aparatu matematycznego związanego ze zbiorami rozmytymi, liczbami rozmytymi, skierowanymi liczbami rozmytymi, pod kątem możliwości użycia w modelu zaproponowanym dla liczb rzeczywistych.
- Opracowanie sposobu tworzenia skierowanych liczb rozmytych na podstawie danych pomiarowych - publikacja [89], wyniki zadania badawczego w postaci zaproponowanego algorytmu konstruowania skierowanych liczby rozmytych zostały wykorzystane w podrozdziale 6.1..
- Opracowanie sposobu poszukiwania skierowanej liczby rozmytej najbliższej zadanej (nie-właściwej) skierowanej liczbie rozmytej - podrozdział 6.2..
- Walidacja opracowanych algorytmów z wykorzystaniem przykładu obliczeniowego wspomaganego komputerowo - rozdział 7.
- Stworzenie modelu probabilistycznego opisu poziomemu izolacji zaprezentowanego w rozdziale 8.
- Opracowanie modelu izolacji usług w sieciach plastrowych, będącym przedmiotem publikacji [88] i rozdziału 9.

1.3. STRUKTURA PRACY

Niniejsza praca ma następującą strukturę: w rozdziale 1 przedstawiono wstęp do niniejszej rozprawy doktorskiej, wraz z opisem tezy pracy i przedstawieniem głównych zadań badawczych. Rozdział 2 opisuje koncept sieci plastrowej, podając najważniejsze pojęcia związane z sieciami plastrowymi. Rozdział 3 rozszerza analizę sieci plastrowych o pojęcie izolacji. Przedstawione są powody, dla których warto podejmować temat izolacji, opisane są typy izolacji i przedstawione przykładowe techniki zapewniania izolacji. Rozdział 4 poświęcony jest przedstawieniu algorytmu wyznaczania izolacji z wykorzystaniem liczb rzeczywistych. Podejście to stanowi bazę do rozważań poczynionych w dalszej części pracy. Przedmiotem rozważań podjętych w rozdziale 5 jest przedstawienie skierowanych liczb rozmytych, ich podstawowych cech i właściwości matematycznych. W rozdziale 6 opisano metodę wyznaczania izolacji wykorzystując w tym celu skierowane liczby rozmyte. Zaprezentowano sposób tworzenia skierowanych liczb rozmytych na podstawie wyników pomiarowych i określania skierowania tak utworzonej liczby rozmytej. Przedstawiono metodę poszukiwania skierowanej liczby rozmytej właściwej dla zadanej skierowanej liczby rozmytej niewłaściwej za pomocą technik programowania kwadratowego i liniowego, minimalizując zadaną metrykę odległości pomiędzy liczbami. W rozdziale 7 zaprezentowano przykładowe obliczenia z wykorzystaniem przedstawionego modelu rozmytego. Rozdział 8 poświęcony jest zdefiniowaniu modelu probabilistycznego izolacji. W rozdziale 9 przedstawiono zagadnienie izolacji usług, wraz z modelem ją opisującym. Podsumowanie pracy i wnioski zostały ujęte w rozdziale 10. Praca zawiera dodatek A uzupełniający model probabilistyczny o zbiór lematów i twierdzeń z nim powiązanych.

2. Sieć plastrowa

2.1. MOTYWACJA

Sieć plastrowa, znana w anglojęzycznej literaturze jako technika *Network Slicing*, jest jedną z kilku technologii rozpatrywanych do implementacji sieci typu 5G oraz nowszych - powstają już prace wiążące ideę sieci plastrowej z sieciami 6G [78, 122]. Kluczową różnicą między siecią 5G a poprzednimi sieciami komórkowymi, jest odejście od koncepcji jednej sieci dla wszystkich usług internetowych, czy też realizowanie wszystkich usług z jednakowym poziomem QoS/QoE. Dotychczas istniały pewne mechanizmy pozwalające różnicować ruch sieciowy i stosować odpowiednie polityki obsługi takiego ruchu, np. z wykorzystaniem pola TOS (*Type Of Service*) w ramach protokołu IPv4 (*Internet Protocol, version 4*), natomiast nie były one szeroko stosowane i respektowane w ramach Internetu. Brak szerokiego poparcia dla tego typu technik, uniemożliwiał realizowanie za pomocą Internetu pewnych klas usług, jak np. skorzystanie z numeru alarmowego za pomocą technologii VoIP (*Voice over IP*). Mimo rozpowszechnienia się platform tego typu (np. Skype, Messenger, Microsoft Teams, Zoom, IBM Webex), dostawcy usług nie mogli zagwarantować dostępności do telefonu alarmowego na poziomie dostępności oferowanej przez operatorów sieciowych, czy to w ramach sieci stacjonarnej (PSTN - *Public Switched Telephone Network*), czy komórkowej. Podział logiczny sieci, wprowadzany w sieci plastrowej, ma umożliwić rozwiązanie części z tych problemów.

W niniejszej pracy będziemy skupiać się na sieciach plastrowych w ramach sieci 5G, chociaż ogólny koncept może być zastosowany do innych sieci telekomunikacyjnych. W szczególności, może on być dalej utrzymywany w ramach sieci *Beyond 5G* i 6G, nad którymi już są czynione prace koncepcyjne (np. [39, 82]). Historycznie, sieć plastrowa została opisana w pracy [96], gdzie autorzy opisali sieć *PlanetLab*, która wśród założonych wymagań architektonicznych miała implementować wsparcie dla podziału na plastry [59] - rozumiane w tym projekcie jako część zasobów sprzętowych (czas procesora, przepływność binarna łącza, pamięć trwała).

W ramach sieci plastrowej występuje podział logiczny sieci na plastry, które są skojarzone z zasobami. Zasoby mogą być fizyczne, bądź też zwirtualizowane. Kluczowym zagadnieniem jest stworzenie takiego środowiska technicznego i podziału na plastry, w ramach którego plastry są odizolowane od siebie wzajemnie oraz od środowiska zewnętrznego. Pozwala to na świadczenie usług z określonymi (zwykle wysokimi) wymaganiami SLA, o określonym poziomie jakości QoE i QoS. W przypadku niewystarczającej izolacji plastrów, warunki techniczne mogą nie być spełnione, co niesie za sobą konsekwencję w postaci niezrealizowania wymagań jakościowych przez usługodawcę, a w konsekwencji niezadowolenie klienta końcowego z otrzymanej usługi. Sieć 5G może być wykorzystywana przez wiele różnych podmiotów z całej gospodarki, wobec czego zadowolenie klienta końcowego ze świadczonych usług za pomocą tej sieci staje się zagadnieniem bardzo istotnym. Kluczowi klienci, o największym potencjale ekonomicznym to wspomniani we wstępie członkowie sektorów wertykalnych.

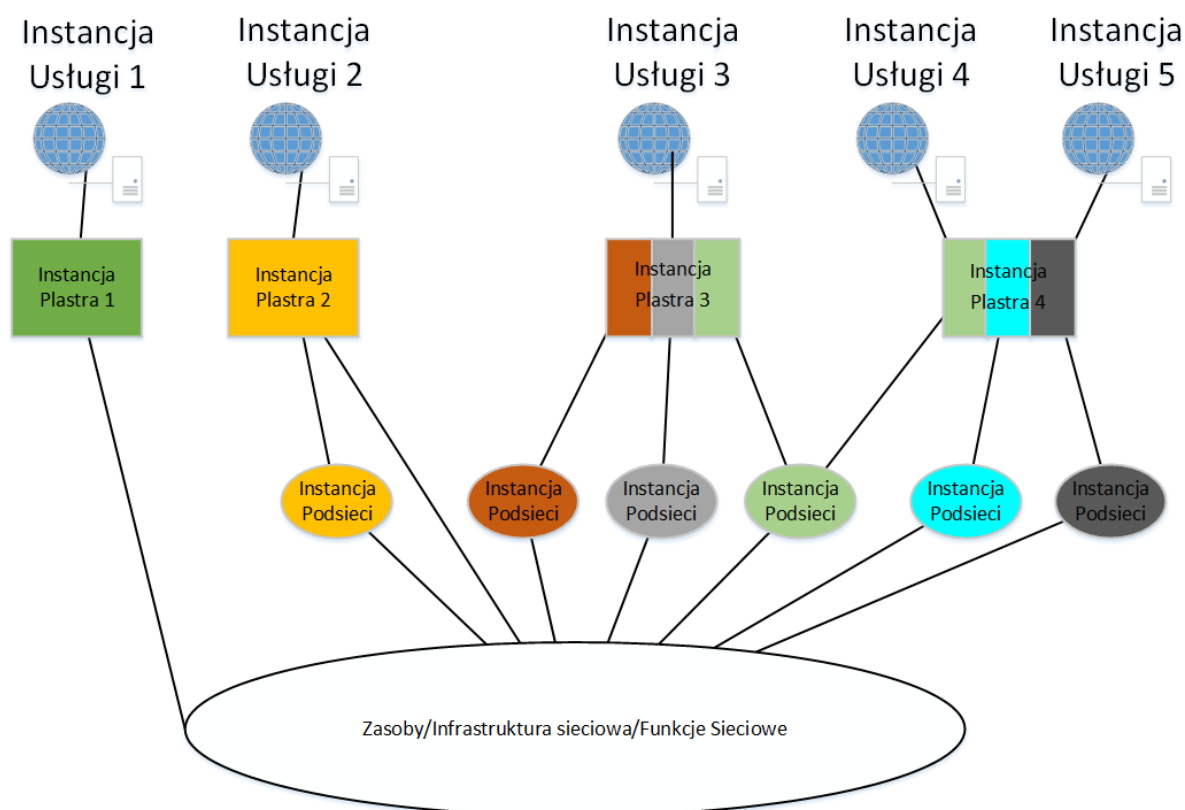
2.2. DEFINICJE SIECI PLASTROWEJ

Jednym z modeli sieci plastrowej jest model zaproponowany przez NGMN (*Next Generation Mobile Networks*) [14] i opisany szerzej w [59], zaprezentowany na rysunku 2.1). Składa się on z trzech warstw:

- Warstwa Instancji Usług (*Service Instance Layer*),
- Warstwa Instancji Plastrów (*Network Slice Instance Layer*),
- Warstwa Zasobów (*Resource Layer*).

Warstwa Instancji Usług zawiera usługi wspierane przez sieć plastrową. Każda usługa jest realizowana poprzez jedną bądź wiele Instancji Usług w ramach tej warstwy. Usługi są dostarczane i świadczone przez zewnętrznych partnerów bądź też przez operatora.

Warstwa Instancji Plastrów jest zbiorem funkcji sieciowych (zwirtualizowanych bądź też fizycznych) wykorzystujących zasoby sieciowe z Warstwy Zasobów do ich realizacji. Funkcje sieciowe są dostarczane z zachowaniem parametrów sieciowych, z poziomem jakościowym zdefiniowanym w wymaganiach Instancji Plastra, która ma za zadanie spełnić wymagania jakościowe żądane przez Instancje Usług, korzystające z danej Instancji Plastra. Poszczególne Instancje Plastrów mogą być między sobą odizolowane z wykorzystaniem różnych technik, jak np. izolacja fizyczna bądź logiczna. Wiele Instancji Usług może współdzielić jedną Instancję



Rysunek 2.1: Koncepcja sieci plastrowej według NGMN - rysunek pochodzi z pracy [91], stworzony na podstawie [14].

Plastra, a sama Warstwa Instancji Plastrów zawiera jednocześnie wiele Instancji Plastrów. O Instancji Plastra można myśleć, jak o wydzielonym, logicznym fragmencie sieci.

Warstwa Zasobów składa się z zasobów fizycznych i logicznych wykorzystywanych do stworzenia Instancji Plastra.

Inne podejście do sieci plastrowych zaprezentowano w pracy Furuhashi i Nakao [31], gdzie zdefiniowano plaster jako jednostkę zasobów sieciowych i obliczeniowych (czas procesora, pamięć trwała i RAM - *Random Access Memory*, przepływność binarna interfejsów sieciowych, itd.), która została przydzielona do przetwarzania zbioru pakietów. Jest to podejście podobne do podejścia założonego w projekcie PlanetLab [96].

2.3. PODZIAŁ HORYZONTALNY I WERTYKALNY

W literaturze na przykład w pracach [17, 59, 64] opisano dwa podejścia do definiowania plastrów - podział horyzontalny i wertykalny na plastry. Porównanie obydwu podejść jest przedstawione w tabeli 2.1. Obydwa podejścia mogą występować równoległe w ramach jednej sieci [59].

Tablica 2.1: Porównanie koncepcji plastrów horyzontalnych i plastrów wertykalnych

Podział horyzontalny	Podział wertykalny
Podział horyzontalny pozwalają na współdzielenie fizycznych zasobów pomiędzy logicznymi węzłami sieci (urządzeniami sieciowymi) [77]. W podziale na horyzontalnym funkcjonalność jest odseparowana od fizycznego urządzenia, które wykonuje dane funkcjonalności.	Podział wertykalny polega na współdzieleniu zasobów pomiędzy usługami i aplikacjami.[77]
Podział na plastry horyzontalne pozwala na dynamiczne dodawanie węzłów aby obsłużyć nowe plastry lub też zwiększyć możliwości obsługi obecnych plastrów - wspierana jest skalowalność ruchu oraz zwinność rozwiązania poprzez łatwiejsze wdrażanie funkcji sieciowych w węzłach.	Definiowane są plastry w ramach pewnej domeny sieciowej, które następnie są łączone z plastrami przynależnymi do sąsiednich domen. Ruch sieciowy jest klasyfikowany, aby przydzielić go do odpowiedniego plastra [77].

2.4. PEŁNE POŁĄCZENIE PLASTROWE

Z praktycznego punktu widzenia użytkownik chciałby móc skorzystać z sieci plasterowej na wszystkich odcinkach łącza pomiędzy UE użytkownika a usługą, z której zamierza on skorzystać. W literaturze angielskiej używany jest termin *E2E (End-to-End) Network Slicing* w stosunku do tego zagadnienia [1, 59, 64]; na potrzeby tej pracy można go przetłumaczyć jako **pełne połączenie plasterowe**. Możemy wydzielić trzy podstawowe obszary podziału tego połączenia na plastry [59]:

— obszar bezprzewodowy,

- obszar związany z obsługą sieci bezprzewodowej po stronie Operatora (RAN),
- sieć szkieletową (CN).

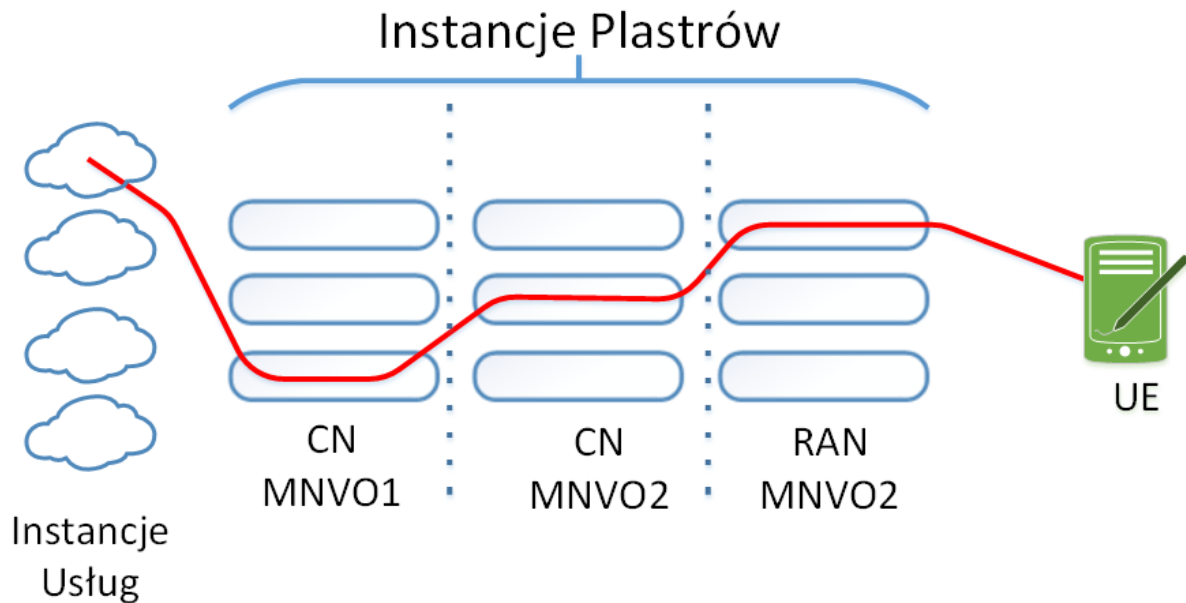
Podział jest tutaj umowny i w konkretnych realizacjach może być bardziej rozbudowany.

Obszar bezprzewodowy charakteryzuje się naturalnym współdzieleniem pasma częstotliwościowego pomiędzy różnymi użytkownikami i urządzeniami. Sygnały nadawane w zadanym paśmie częstotliwości mogą być generowane przez urządzenia zupełnie niezwiązane z telekomunikacją mobilną, na przykład kuchenki mikrofalowe, czy też związane z innymi systemami łączności (DECT - *Digital Enhanced Cordless Telephony*, telewizja analogowa i cyfrowa, radio, Wi-Fi - *Wireless-Fidelity*, itd.). Oznacza to, że naturalnie występują tutaj zewnętrzne zakłócenia, jakie mogą wpływać na komunikację wewnątrz danego plastra. Wsparcie dla podziału na plastry w ramach tego odcinka wymaga rozpatrzenia wzajemnego wpływu nie tylko pomiędzy plastrami, ale i również między plastrzem a *światem zewnętrznym*. Przykładowymi technikami pomocnymi w realizacji tych wymagań jest wdrożenie separacji kanałów częstotliwościowych (wykorzystanie FDMA - *Frequency-Division Multiple Access*), bądź też wdrożenie odpowiedniego kodowania, które pozwoli poprawnie odebrać nadawany sygnał w sytuacji znaczącego poziomu szumu tła (niskie SNR - *Signal to Noise Ratio*; przykładowo zastosowanie rozwiązania podobnego do CDMA - *Code-Division Multiple Access*).

W przypadku obszaru RAN, najważniejsze jest wydzielenie dwóch warstw - warstwy sterowania (*Control Plane*) i warstwy danych użytkownika (*User/Data Plane*). Izolację powinno się osobno analizować wobec obydwu warstw, ponieważ izolacja może być niewystarczająca dla tylko jednej z warstw, aby możliwe było przeprowadzenie skutecznego ataku na realizowaną usługę. W ogólności nie zakładamy konkretnej technologii warstwy fizycznej i łącza danych do realizacji usług, mogą to być przykładowo rozwiązania wykorzystujące RAN 4G, Wi-Fi, WiMax i inne, a nawet tradycyjne przewodowe sieci LAN (*Local Area Network*).

Ostatnim obszarem jest odcinek wewnątrz sieci Operatora, który prowadzi do serwera świadczącego usługę. Podział na plastry w ramach tego odcinka może być wspierany przez takie rozwiązania sieciowe jak SDN lub NFV [64].

Jednym ze sposobów na zrealizowanie pełnego połączenia plastrowego jest stworzenie łańcucha plastrów (ang. *Slice Chaining*), które pozwalają zestawić połączenie pomiędzy dwoma punktami końcowymi; pomysł został opisany w [23, 58, 91], a graficznie przedstawiony jest na



Rysunek 2.2: Koncepcja łańcucha plastrów - rysunek pochodzi z pracy [91].

rysunku 2.2. Oczekiwane jest zestawienie takiego połączenia, które będzie spełniać wymagania jakościowe w każdym plastrze z osobna, oraz co nie mniej istotne - sam łańcuch też je będzie spełniać. Spełnienie warunków jakościowych przez wszystkie ogniwa (plaster) łańcucha z osobna nie pociąga za sobą automatycznie spełnienia takich warunków przez łańcuch. Przykładowo, opóźnienia czy jitter muszą być rozpatrywane na poziomie całego łańcucha.

2.5. ZARZĄDZANIE CZASEM ŻYCIA PŁASTRÓW

Z perspektywy Operatora Telekomunikacyjnego, klientów i usługodawców ważne jest, aby w momencie realizacji usługi, istniały odpowiednie plastry, które mają być wykorzystywane do realizacji tej usługi. Muszą wobec tego istnieć odpowiednie procesy i protokoły ustalania oraz wyboru plastra, mechanizmy kontroli dostępu, a więc i również zapewniania uwierzytelnienia i autoryzacji na poczet tych operacji. W niniejszej pracy nie będziemy się skupiać na tym zagadnieniu, zakładając jednocześnie, że plastry zostały już uprzednio utworzone - można to nazwać analizą statyczną. Dopełnieniem do niej jest analiza dynamiczna, w której bada się wpływ zmiany stanu sieci plasterkowej na inne plastry i na izolację plastrów. Elementy analizy dynamicznej zostały opisane w pracy [17].

Zarządzanie plastrami, szczególnie w pełnym połączeniu plastrowym, jest zadaniem realizowanym przez mechanizm orkiestracji (*orchestration*). Przykładowo, firma Ericsson wyszczególnia cztery główne zagadnienia związane z realizacją orkiestracji w ich implementacji (*E2E network slicing orchestration*) [43].

- Przydział zasobów przez orkiestrator w oparciu o zapotrzebowanie i dostępne zasoby sieciowe (*transport-aware slicing*).
- Użycie abstrakcyjnego opisu sieci przez orkiestrator, który pozwala odseparować warstwy usług od konkretnych implementacji rozwiązań technicznych, jakie są wykorzystywane przez Operatora Telekomunikacyjnego do realizacji usług *Transport-aware network abstraction*.
- Zarządzanie dostępem do plastrów i zasobów sieci z wykorzystaniem sztucznej inteligencji *Artificial intelligence (AI) based admission control (AC)*.
- Optymalizacja dostarczanego QoS w dynamicznym środowisku sieciowym z wykorzystaniem sztucznej inteligencji *AI-based network dynamics for QoS tuning*.

Zaproponowane przez Ericssona rozwiązania mogą być realizowane na poczet różnych sektorów wertykalnych, w szczególności praca [43] opisuje przypadek użycia tego rozwiązania w nowoczesnej inteligentnej fabryce (*smart factory*).

Orkiestracja może być realizowana w sposób rozproszony, szczególnie gdy ścieżka E2E przechodzi przez sieci należące do różnych operatorów. Przykładowo, praca [5] opisuje architekturę zarządzania plastrami w federacji sieci za pomocą wielu orkiestratorów domenowych (*Domain Specific Orchestrator*), które są sterowane za pomocą globalnego orkiestratora. Podobny, hierarchiczny sposób orkiestracji został przedstawiony w pracy [105], w którym za pomocą *hyperstratora* koordynowane są podległe mu rozproszone orkiestratory, wyspecjalizowane w zarządzaniu konkretnymi fragmentami sieci.

W ramach cyklu życia instancji plastra wchodzi nie tylko faza przygotowawcza i kreacyjna danego plastra, ale i również jego utrzymanie i proces wycofywania go z użycia (*decommissioning*) [17]. Cykl życia instancji plastra jest powiązany z cyklem życia Usług Sieciowych, NFV i NFVI (*Network Function Virtualization Infrastructure*) według modelu ETSI NFV-MANO (patrz [25]).

3. Izolacja w sieci plastrowej

Fundamentalną kwestią dla niniejszej pracy jest zdefiniowanie izolacji w sieci plastrowej. Zagadnienie to zostało jakościowo opisane w pracy [59], której najistotniejsze wyniki przytaczamy tutaj.

3.1. MOTYWACJA

Istnieje kilka powodów, dla których warto rozpatrywać zagadnienie izolacji w sieciach plastrowych. Pierwszym z nich jest chęć opisu i zbadania jak zachowuje się *logiczna* granica między częściami sieci (częściami systemu informacyjnego), w zależności od sposobu jej realizacji. Drugim z powodów jest zaspokojenie potrzeb interesariuszy systemu informacyjnego, w szczególności:

- potrzeby operatorów sieciowych w zakresie zapewnienia bezpiecznego i innowacyjnego środowiska sieciowego, które będzie stanowić fundament dla usług nowej generacji,
- potrzeby usługodawców, w szczególności w sieciach MEC, związane z bezpiecznym świadczeniem usług i przenoszeniem usług pomiędzy różnymi komórkami sieci MEC,
- potrzeby klientów - oczekiwanie lepszej jakości usług, bezpieczniejszych usług.

Wymienione potrzeby wymagają pewnego zarządzania poziomem izolacji, dobór technik i środków zarządzania zależny jest od zaawansowania technologicznego danej organizacji bądź klienta (szczególnie klienta korporacyjnego). Zarządzanie izolacją może być w pewien sposób rozproszone i ulegać centralizacji w ramach organizacji na skutek zmian architektonicznych w zakresie bezpieczeństwa informacji, w szczególności na skutek zmiany architektury korporacyjnej.

Omówmy teraz historię technik realizacji izolacji, przedstawioną w [56]. Jedną z pierwszych, a zarazem ważniejszych technik jest tunelowanie ruchu sieciowego. Rozwiązanie to pozwala na ukrycie treści komunikacji przed nieuprawnionymi aktorami; możliwe jest również wzbogacenie komunikacji o inne cechy jak mechanizm podpisu cyfrowego, wzajemnie

uwierzytelnienie stron, detekcji błędów, itd. Obecnie tego typu technologia wyewoluowała w rozwiązaniu VPN (ang. *Virtual Private Network*). Tunele VPN mogą być zestawiane w różnych warstwach modelu ISO/OSI RM (*ISO Open Systems Interconnection Reference Model*) - w zasadzie już od warstwy drugiej istnieje możliwość logicznego oddzielenia strumieni danych od reszty ruchu sieciowego. Kolejnym krokiem było zaproponowanie sieci programowalnych, które pozwalają oddzielić warstwę sprzętową od implementacji poszczególnych funkcji sieciowych realizujących potrzeby biznesowe Operatora Telekomunikacyjnego.

W ramach nowoczesnych sieci telekomunikacyjnych jak 5G, Beyond 5G, 6G, F5G izolacja pomiędzy plastrami będzie zapewniana przy użyciu wielu technologii, sprzętowych oraz programistycznych, wykraczając poza proste zbudowanie tunelu VPN. Niesie to za sobą konieczność zintegrowania rozwiązań i ich łącznej analizy pod kątem luk bezpieczeństwa, uwzględniając również bezpieczeństwo fizyczne danego systemu.

W temacie izolacji, szczególnie w ramach prac dotyczących sieci 5G, były prowadzone projekty badawczo-rozwojowe, przykładowo [56]:

- Projekt CHARISMA [15], w ramach którego powstał pomysł sieci ICN (*Information-Centric Network*),
- Projekt 5G-ENSURE [2],
- Projekt SliceNet [113].

Tematyka izolacji wydajnościowej w kontekście zwirtualizowanych kontenerów została przedstawiona w pracy [126] - podjęto tam próbę wyznaczenia izolacji kontenerów za pomocą ich odporności na degradację wydajności na skutek zmniejszenia dostępnych zasobów. Do badania eksperymentalnie izolacji w środowisku zwirtualizowanym można wykorzystać narzędzie *Isolation Benchmark Suite* [44], które jest dedykowanym zestawem testów obciążeniowych, mających na celu zbadanie zachowania systemu w przypadku bardzo wysokiego obciążenia, czyli w sytuacji, gdy izolacja (w rozumieniu wydajnościowym) może zostać złamana.

3.2. BEZPIECZEŃSTWO PŁASTRÓW W PEŁNYM POŁĄCZENIU PŁASTROWYM

W niniejszym podrozdziale skupimy się na zagadnieniu bezpieczeństwa plastrów sieci w szerokim kontekście, poruszając również zagadnienia bezpieczeństwa systemów informacyjnych, w których z pewnych powodów warto (bądź jest konieczność) wdrażać izolację.

Świadomość poziomu izolacji między konkretnymi systemami w infrastrukturze IT (*Information Technology*) jest ważna, aby prawidłowo projektować, wdrażać i utrzymywać systemy informacyjne. Rozpatrzmy następującą hipotetyczną sytuację: usługodawca uruchomił usługę w dwóch plastrach udostępnianych przez Operatora Telekomunikacyjnego - testowym i tzw. produkcyjnym. Obydwie usługi wykorzystują serwer CDN udostępniany przez Operatora, za którego używanie Operator pobiera opłatę. Usługodawca udostępnił numer karty kredytowej swojemu Operatorowi w celu zabezpieczenia płatności oraz ustanowił limit wydatków w skali miesiąca z powodu bezpieczeństwa finansowego przedsiębiorstwa. Podczas testów obciążeniowych na środowisku testowym okazało się, że miesięczny limit wydatków został przekroczony, co spowodowało niedostępność usługi świadczonej przez CDN zarówno dla środowiska testowego jak i produkcyjnego. Z tego powodu usługa korzystająca z plastra produkcyjnego nie mogła w założonym czasie uzyskać potrzebnych jej danych i w związku z tym usługodawca nie był w stanie zrealizować wymagań stawianych w kontrakcie SLA.

W przedstawionym przykładzie nastąpiło sprzężenie pomiędzy testową i produkcyjną instancją usługi. Doszło do tego, pomimo możliwości zastosowania wysokiego poziomu odseparowania plastrów pomiędzy sobą - mogły być wdrożone na oddzielnych fizycznych maszynach i w rozłącznych podsięciach. Wspólnym zasobem okazał się rachunek karty kredytowej i limit wydatków w ramach subskrypcji usługi CDN. Ze względu na fakt, że nawet takie pozornie nieistotne z perspektywy technicznej zagadnienia mogą wpływać na dostępność i bezpieczeństwo usług, bardzo trudno jest ocenić całościowo poziom bezpieczeństwa danego rozwiązania; jest to jedna z fundamentalnych trudności w przypadku migracji kontekstu bezpieczeństwa pomiędzy instancjami usług; zagadnienie migracji kontekstu bezpieczeństwa zostało opisane szerzej w [85].

W kontekście izolacji warto wspomnieć zagadnienie przetwarzania informacji (danych) niejawnych. Dane niejawne, w rozumieniu *Ustawy z dnia 5 sierpnia 2010 r. o ochronie informacji niejawnych* [117] wymagają specjalnego sposobu ich przetwarzania, gdzie za *przetwarzanie* uważa się w szczególności [6]: wytwarzanie, modyfikowanie, kopiowanie, klasyfikowanie, gromadzenie, przechowywanie, przekazywanie lub udostępnianie informacji niejawnych. Zagadnienie tworzenia systemów teleinformatycznych przeznaczonych do przetwarzania informacji niejawnych, w szczególności z myślą o akredytacji bezpieczeństwa teleinformatycznych systemów wykracza poza zakres tej pracy. Monografia [62] opisuje sposób konstruowania istotnego elementu takiego systemu, jakim jest bezpieczny kryptograficznie generator losowych ciągów binarnych.

Nawet pobieżna lektura przytoczonej Ustawy prowadzi do wniosku, że regulator (ustawodawca) również przewidział konieczność zarządzania informacjami o różnych poziomach niejawności, a wobec tego i budowy systemów do przetwarzania tego typu informacji. Wraz ze wzrostem poziomu tajności przetwarzanych danych, potrzeba ich ochrony rośnie, w szczególności tracona jest możliwość przetwarzania takich danych w ogólnodostępnych systemach, czy też administrowania takimi systemami przez osoby bez odpowiednich poświadczeń bezpieczeństwa. Koncepcja posiadania kancelarii tajnej w przedsiębiorstwie do przetwarzania danych niejawnych jest praktyczną implementacją fizycznej separacji procesów informacyjnych, w swoim zamyśle bardzo podobna do fizycznej separacji (izolacji) plastrów w sieciach plastrowych poprzez fizyczne wydzielenie części zasobów sieciowych tylko do użytku danego plastra. Opis sposobu zarządzania materiałami oznaczonymi klauzulami tajności został przedstawiony w rozporządzeniu [103]; co ciekawe, jest to w swoim głównym zamierzeniu system etykietowania materiałów odpowiednimi klauzulami. Organizacje należące do takich sektorów wertykalnych jak Infrastruktura Krytyczna czy Rząd i Samorząd [92] (włączając podległe służby cywilne i wojskowe) są użytkownikami tego typu rozwiązań i w związku z tym są zainteresowane procesem izolowania pewnych danych i informacji.

Przy okazji omawiania systemów informacyjnych o bardzo wysokim poziomie izolacji warto spojrzeć na konsekwencje takiego podejścia. Rozpatrzmy system informacyjny, który został idealnie odizolowany od świata zewnętrznego. Załóżmy, że chcemy w takim systemie zaszyfrować pewne dane, lub z jakiegoś powodu potrzebujemy wygenerować zestaw bitów losowych, czy to na potrzeby kryptograficzne, czy z przeznaczeniem do pozakryptograficznego

zastosowania. Zupełne odseparowanie takiego systemu uniemożliwia pobieranie informacji z zewnątrz oraz czerpania korzyści z wpływów zewnętrznych na niektóre wybrane elementy systemu, jak na przykład sprzętowe generatory ciągów losowych [62, 90].

Innym zagadnieniem związanym z intencjonalnym zmniejszaniem izolacji jest zastosowanie różnych form pułapek na atakującego, które określa się w literaturze jako *honeypots* i *honeynets* [12], w przypadku dużych zbiorów takich systemów, mogących nawet tworzyć dedykowane podsieci. Celem takich pułapek jest zachęcenie atakującego do skupienia swojej uwagi na danym systemie (sieci), aby uniknąć ataku na realne systemy (sieci) oraz zdobycie informacji na temat sposobu działania intruza lub wykrytego złośliwego oprogramowania.

Brak izolacji w pewnym ograniczonym zakresie może być wykorzystywany do stworzenia kanału komunikacyjnego pomiędzy różnymi węzłami sieci. Przykładowo, sygnalizacja RTS/CTS (*Ready To Send / Clear To Send*) [41] intencjonalnie wykorzystuje wspólne medium do przeprowadzania pewnego rodzaju synchronizacji działań pomiędzy węzłami sieci.

Z analizy wyzwań stawianych bezpieczeństwu sieci plastrów w pracach [7, 35] wynikają następujące zagadnienia związane z izolacją, przedstawione zbiorczo w [59].

- Współdzielenie zasobów i dedykowanie zasobów wraz z problemem wysycania zasobów przez jeden lub wiele plastrów, kosztem innych plastrów.
- Dedykowanie zasobów dla zadanych protokołów/strumieni informacyjnych w ramach plastra.
- Uniemożliwianie przepływu informacji pomiędzy plastrami.
- Umożliwianie przepływu informacji pomiędzy plastrami w ściśle zdefiniowany i kontrolowany sposób.
- Ochrona przed włamaniem do plastra, aby skutecznie wpłynąć na strumienie lub jeziora informacyjne w innym plastrze.
- Zarządzanie i sygnalizacja związana z zapewnianiem izolacji, komunikacja pomiędzy orkiestratorem a komponentami tworzącymi plaster, oraz komunikacja pomiędzy tymi komponentami.
- Zapewnianie niezawodności fizycznych elementów użytych do konstrukcji plastra.
- Bezpieczna komunikacja pomiędzy (wieloma) systemami zarządzającymi plastrami.
- Izolacja uwzględniająca poziom emisji informacji (np. entropii) do środowiska zewnętrznego, z uwzględnieniem tzw. *side-channel attacks* i ochrony przed nimi.

- Izolacja w środowisku hybrydowym zbudowanym z fizycznych i zwirtualizowanych funkcji sieciowych (NFV i VNF).
- Izolacja plastrów w przypadku, gdy urządzenie końcowe (UE) jest jednocześnie klientem wielu plastrów.

W pracy [7] opisano sposoby jak przeciwdziałać wymienionym wyzwaniom dla bezpieczeństwa systemów telekomunikacyjnych.

Poruszmy teraz kwestię wymagań stawianych bezpieczeństwu plastrów w pełnym połączeniu plastrowym przez sektory wertykalne. W pracy [92] dokonano szerokiego przeglądu wymagań dotyczących bezpieczeństwa w sieciach 5G MEC użytkowanych przez takie podmioty, a które to sieci mogą implementować podział na plastry w celu realizacji wymagań jakościowych związanych z izolacją i bezpieczeństwem. Można tutaj wyszczególnić następujące kategorie przypadków użycia lub też całych sektorów wertykalnych.

- Przypadki użycia i sektory o bardzo wysokich wymaganiach dotyczących bezpieczeństwa danych i systemów informacyjnych. Może to wynikać z charakteru działalności danej organizacji - przykładowo sektor Ochrony Zdrowia ze względu na przetwarzane dane wrażliwe dotyczące zdrowia pacjentów musi szczególnie chronić te dane. W stosunku do zapewnienia ciągłości usług zdrowotnych takich jak zdalna obserwacja stanu zdrowia pacjenta, automatyczne leczenie (*smarter medication*), zdalne operacje chirurgiczne [92] wymagane jest zapewnienie niskich opóźnień w transmisji, wysokiej dostępności (*High Availability*), wysokiej niezawodności, a niekiedy i też sporej przepływności binarnej łącza. Wymienione czynniki składają się na szeroko rozumianą dostępność usługi, co z kolei jest jednym z zagadnień zapewniania bezpieczeństwa i izolacji - niezależnie od sytuacji sieciowej, usługa powinna być zrealizowana. Innym przykładem są przypadki użycia z sektora Transportu i Logistyki, głównie związane z komunikacją pomiędzy pojazdem a światem zewnętrznym (V2X) czy autonomicznymi pojazdami [92] - gdzie wymagania dotyczące dostępności do usługi są bardzo wygórowane, dodatkowo utrudnione przez wysoki stopień mobilności pojazdów i konieczność zachowania niskich opóźnień, ze względu na prędkość poruszania się pojazdów (autonomicznych i sąsiednich); jest to przykład ruchu z kategorii URLLC (*Ultra-Reliable and Low-Latency Communication*). Również sektor BFSI ma wysokie wymagania związane z dostępnością do usług czy poziomem izolacji, chociaż jest mniej wymagający w innych kwestiach [92]. Ostatnim ciekawym przykładem jest sektor wertykalny

Infrastruktura Krytyczna, który zawiera z założenia takie organizacje (lub ich istotne komponenty), jakie są ważne z perspektywy społeczeństwa i ciągłości działania państwa. Praca [92] wymienia wiele sektorów należących do Infrastruktury Krytycznej, o bardzo szerokim spektrum - od żywności, przez systemy transportowe, opiekę zdrowotną, energetykę, system finansowy, po zarządzanie odpadami i sam sektor IT. Przez to, wrażliwość na problemy związane z bezpieczeństwem informacji i niedostateczną izolacją informacji i procesów informacyjnych jest różna, chociaż w zdecydowanej większości z nich jest to ważny obszar z perspektywy ciągłości działania przedsiębiorstwa, nawet jeżeli jego głównym produktem (usługą) nie są produkty i usługi informacyjne.

- Przypadki użycia i sektory o średnim i niskim poziomie wymagań bezpieczeństwa - niektóre usługi ze względu na specyfikę nie mają wygórowanych wymagań SLA, w szczególności związanych z bezpieczeństwem. Przykładowo, sektor Handlu może dostarczać usługi związane z transmisją video do urzędzeń klientów, co wymaga pewnej przepływności binarnej na stosunkowo niewielkim obszarze [92] - ale bezpośrednie konsekwencje związane z niezrealizowaniem usługi są niewielkie w porównaniu z innymi sektorami. Podobnie sytuacja wygląda w sektorze Rolnictwa [92] - gęstość urzędzeń podłączonych do sieci może być znacząca, natomiast inne wymagania sieciowe nie są aż tak restrykcyjne - jest to ruch należący do kategorii mMTC (*massive Machine Type Communication*).

3.3. TYPY IZOLACJI

W sieciach telekomunikacyjnych szeroko rozpowszechnione są techniki multipleksacyjne, które pozwalają na współdzielenie i alokację dostępnych zasobów pomiędzy kanałami telekomunikacyjnymi. Można wyszczególnić trzy podstawowe typy izolacji, jakie dostarczają takie metody [59]:

- pełna izolacja - gdy każdy kanał korzysta z zasobów dedykowanych tylko dla niego, np. szczeliny czasowe w TDMA (*Time-Division Multiple Access*).
- częściowa izolacja - gdy kanały mogą współdzielić część dostępnych zasobów, np. kanały w sieci Wi-Fi, które na siebie zachodzą w znacznym stopniu w dziedzinie częstotliwości.
- brak izolacji - kiedy wszystkie kanały współdzielą część dostępnych zasobów, np. kanały w OFDMA (*Orthogonal Frequency-Division Multiple Access*).

W niektórych przypadkach możliwe jest zapewnienie izolacji poprzez odpowiednie przyporządkowanie kanałów do zasobów, aby uniknąć współdzielenia zasobów między kanałami. W niektórych przypadkach, mimo dostępnej dużej bazy zasobów, np. kanałów w sieci IEEE 802.11b (Wi-Fi), można równocześnie bezproblemowo korzystać tylko z części z nich, ponieważ sąsiednie kanały współdzielą pasma częstotliwości, co powoduje utrudnioną komunikację [59]. W tym miejscu warto również wspomnieć o istotnym czynniku wpływającym na izolację - obecność innych urządzeń współdzielących dany zasób, ale nie będący pod kontrolą warstwy zarządzania danym zasobem, jednocześnie wpływających na system [59]. Przykładem takiego zjawiska są urządzenia działające w otwartym paśmie 2.4 GHz (np. kuchenki mikrofalowe), które zagłuszają kanały komunikacyjne w sieciach bezprzewodowych [59].

Izolacja między kanałami może być zmniejszona na skutek zjawisk fizycznych związanych z propagacją sygnału. Praca [59] wymienia tutaj zjawisko wielodrogowości [101, 106] (odbiornik odbiera sygnały nadane wcześniej lub nieprzeznaczone dla niego), dyspersji czasowej sygnału [106] (sygnał rozciąga się w czasie), tłumienia sygnału [101, 106], gdzie maleje SNR względem szumu tła.

3.4. PODSTAWY MODELOWANIA IZOLACJI

Nowoczesna sieć telekomunikacyjna może być zamodelowana za pomocą hipergrafu, który można zdefiniować następująco.

Definicja 3.1 (Hipergraf)

Hipergraf $H = (V, E)$ jest uporządkowaną parą zbioru wierzchołków V i hiperkrawędzi E . Hiperkrawędź definiujemy jako dowolny multizbiór wierzchołków z V o przynajmniej dwóch elementach, niekoniecznie różnych.

Definicja 3.2 (Multizbiór)

Niech A będzie dowolnym zbiorem skończonym oraz będzie zadana funkcja $Q_B(x) : A \rightarrow \mathbb{N}_{\geq 0}$. Multizbiorem B nazywamy zbiór $B = \{(x, Q_B(x)) : x \in A \wedge Q_B(x) > 0\}$.

Będziemy dla uproszczenia korzystać z notacji $\{a_1, a_1, a_2, a_3\}$, która oznacza, że element a_1 występuje w multizbiorze dokładnie dwa razy, a elementy a_2 i a_3 po jednym razie. Przykładowo, gdy $V = \{v_1, v_2, v_3\}$, to dopuszczalne są hiperkrawędzie $\{v_1, v_1\}$, $\{v_1, v_2, v_3\}$,

$\{v_1, v_2, v_2\}$. Oznacza to, że hiperkrawędź może przechodzić przez węzeł wielokrotnie oraz obejmować więcej niż dwa różne węzły. Przykładem takiej hiperkrawędzi w praktyce jest ośrodek w jakim następuje propagacja sygnału bezprzewodowego. Nieoczywiste staje się zagadnienie skierowania hiperkrawędzi; w ramach niniejszej pracy będziemy przyjmować formułę rozgłoszeniową - węzeł należący do hiperkrawędzi rozgłasza dane (informacje) do wszystkich węzłów incydentnych z daną hiperkrawędzią. Szerszy opis hipergrafowego modelu sieci znajduje się w dalszej części pracy w podrozdziale 4.1.. Wprowadźmy teraz następujące definicje formalne.

Definicja 3.3 (Strumień danych)

Proces przesyłania danych między węzłami hipergrafu sieci plastrowej nazywamy strumieniem danych. Węzeł źródłowy może być tożsamy z węzłem docelowym w ramach tego procesu.

Definicja 3.4 (Jezioro danych)

Węzeł hipergrafu pozwalający wytworzyć strumień danych nazywamy jeziorem danych.

Według tej definicji, jeziorem danych jest element sieciowy wyposażony w formę pamięci, którą można odczytać (dysk twardy, RAM, pamięć podręczna). Jest nim również łącze między elementami sieciowymi - zarówno przewodowe, jak i bezprzewodowe. Ze względu na propagację fali elektromagnetycznej (radiowej, świetlnej) czy też propagację sygnału elektrycznego, dane mogą być odczytane z takiego łącza, ale przez bardzo krótki czas, w zasadzie jednokrotnie. Dla uproszczenia, pomijamy tutaj zjawiska związane z wielodrogowością propagacji sygnału.

Definicja 3.5 (Izolacja)

Izolacją nazywamy zdolność do ochrony jezior i strumieni danych.

Zagrożenia dla izolacji jezior i strumieni danych mogą pochodzić z: wnętrza plastra, od innego plastra, oraz ze świata zewnętrznego. Omówmy teraz pokrótce charakter każdego z głównych źródeł zagrożeń.

Wewnątrz plastra może występować wiele strumieni danych, które wzajemnie na siebie oddziałują. Jedną z podstawowych form przenoszenia tego oddziaływania jest zjawisko wykorzystywania wspólnych zasobów przez procesy informacyjne. Można tutaj podać kilka przykładów takich oddziaływań.

— Strumienie informacyjne przepływają przez wspólny serwer. Współdzielą one skończoną ilość pamięci RAM, czasu procesora, przepływności binarnej kart sieciowych, bufor zapisu

i odczytu z nośnika pamięci trwałej (dysk twardy, dysk SSD - *Solid State Drive*), pamięć cache procesora, otwarte gniazda (*sockets*) do obsługi ruchu TCP (*Transmission Control Protocol*), UDP (*User Datagram Protocol*) przez system operacyjny i aplikacje. Bardzo intensywne wykorzystywanie jednego (lub więcej) z tych zasobów może wpłynąć na inne strumienie informacyjne, np. wydłużając czas przetwarzania komunikatu przesyłanego w ramach strumienia. To proste zjawisko może być również zaobserwowane w sytuacji wykorzystania kontenerów (np. z wykorzystaniem oprogramowania Docker, Kubernetes) czy też maszyn wirtualnych - gdy zasób współdzielony między kontenerami, maszynami wirtualnymi - czyli został zwirtualizowany.

- W przypadku sieci lokalnej o topologii szyny (klasyczna sieć LAN, w pewnym sensie również sieć Wi-Fi), dane wysyłane przez jedną stację uniemożliwiają innym stacjom nadawanie. Zajęte współdzielone medium wydłuża czas potrzebny na propagację komunikatu między aplikacjami. Ten problem odnosi się nie tylko do warstwy fizycznej czy też łącza danych (wg. modelu ISO/OSI RM), ale dotyka też innych współdzielonych zasobów na zasadzie wyłączności dostępu, np. chronionych sekcją krytyczną za pomocą muteksów lub semaforów.

W ramach tego samego plastra można również wpłynąć na strumień danych za pomocą konwencjonalnych ataków hakerskich. Najważniejsze z nich to:

- generyczny hacking - strumień danych może zostać zakłócony przez jednego lub więcej użytkowników sieci - na przykład poprzez atak na serwer usług. Może również nastąpić niepowołany dostęp do jeziora danych, bez zakłócania strumieni danych, ale z naruszeniem założonych polityk bezpieczeństwa (*explicite* lub *implicite*) - przykładowo, atakujący może wykraść dane z serwera usług posługując się spreparowanym żądaniem wykorzystującym podatność bezpieczeństwa taką jak *SQL-injection* (SQL - *Structured Query Language*).
- (D)DoS (*Distributed Denial of Service*) - intencjonalne spowodowanie, że pewne kluczowe zasoby wykorzystywane do obsługi strumieni danych zostaną wyczerpane. W przeciwieństwie do klasycznego wyczerpania zasobów opisanego powyżej, w tym wypadku intencją jest atak, nie zaś zwykle wykorzystywanie usługi.
- *Spoofing* - atak polegający na podszywaniu się pod jeden z fragmentów systemu informacyjnego, może być realizowany na wiele sposobów, np. poprzez podszywanie się pod hosta

z innym adresem IP, MAC (*Medium Access Control*) lub pod usługę taką jak WWW (*World Wide Web*) czy DNS (*Domain Name System*) - patrz [45].

- Podłuchiwanie - atakujący zyskuje nieautoryzowany dostęp do strumienia danych, który jest przesyłany w systemie.
- *Cryptojacking* [61] - atak polegający na użyciu zasobów w celu wykonania obliczeń bez wiedzy i jawnej zgody użytkownika maszyny. Przykładem takich obliczeń jest kopanie kryptowalut, których beneficjentem jest atakujący. Atak wpływa na dostępne zasoby dla przetwarzania strumieni danych.

Rozważmy teraz zagadnienie zagrożeń pochodzących od innego plastra. Wiele plastrów może wykorzystywać wspólny zasób, tak samo jak strumień danych wewnątrz jednego plastra. Ze względu na wymagania bezpieczeństwa, sposób współdzielenia zasobu może być inny niż w przypadku usług w ramach jednego plastra.

Ostatnim zagadnieniem jest wpływ świata zewnętrznego (zagrożenia ze świata zewnętrznego) na strumień i jeziora danych realizowane w ramach sieci plastrowej. Przez świat zewnętrzny rozumiemy tutaj między innymi strumień i jeziora danych pracujące poza siecią plastrową (tzw. sieć *legacy*). Z punktu widzenia telekomunikacji bezprzewodowej istotne tutaj są również inne źródła sygnału w używanym paśmie, szczególnie gdy pasmo jest otwarte (np. 2.4 GHz wykorzystywane na potrzeby sieci Wi-Fi). Zaliczają się do tych źródeł również źródła zakłócające sygnał jak infrastruktura wojskowa czy też urządzenia wpływające na sposób realizacji połączenia przez UE, np. wykorzystywane przez służby specjalne. W przypadku koegzystencji sieci plastrowej z siecią *legacy*, możliwe jest również wyczerpanie zasobów przez sieć *legacy*, co uniemożliwi sieci plastrowej realizację usług. Wreszcie, pewien wpływ mają również pewne zjawiska atmosferyczne (burze, silny wiatr, gwałtowny deszcz), ze względu na wpływ na propagację sygnału elektromagnetycznego w otwartej przestrzeni. Istotnym czynnikiem jest również dostęp do zasilania z sieci energetycznej, w przypadku przerwy w dostawie prądu następuje poważne ryzyko utraty dostępu do usług realizowanych przez infrastrukturę fizyczną za pośrednictwem danego plastra.

3.5. ISTOTNOŚĆ ZAGADNIENIA IZOLACJI

Przejdźmy teraz do omówienia zagadnienia postawionego w tezie pracy - stwierdzenia, że *zarządzanie izolacją (i jej opis ilościowy) jest ważnym elementem zapewniania bezpieczeństwa usług w sieciach plastrowych*. Jak zostało wspomniane w podrozdziale 3.1., niektórzy interesariusze systemu (operatorzy sieciowi, usługodawcy) mogą być zainteresowani poziomem bezpieczeństwa sieci plastrowej lub konkretnego systemu teleinformatycznego. Przedstawmy kilka powodów dla których interesariusze mogą wyrażać takie zainteresowanie.

- Zobowiązania prawne takie jak wymagania ustawowe lub kontraktowe. Przykładem wymagań ustawowych jest rozporządzenie RODO (Rozporządzenie 2016/679 Parlamentu Europejskiego o ochronie danych osobowych) [28] wymuszające na usługodawcy i w pewnym stopniu na operatorze telekomunikacyjnym odpowiednie dbanie o dane osobowe i wrażliwe; do konkretnych sektorów wertykalnych mogą być również stosowane inne akty prawne, właściwe dla charakteru świadczonej usługi. W ramach zobowiązań kontraktowych możliwe jest uzgodnienie, że strona (strony) umowy będą przetwarzać dane w określony sposób, przetwarzanie danych musi być odpowiednio zabezpieczone (poufność i integralność danych), monitorowane czy podlegające audytowi. Stroną takiej umowy może być użytkownik końcowy (w szczególności korporacyjny), usługodawca lub operator telekomunikacyjny. Również w ramach zobowiązań kontraktowych usługa może mieć zdefiniowane wymagania SLA, co wymusza na usługodawcy i operatorze telekomunikacyjnym odpowiednie zarządzanie dostępnymi zasobami dla usługi i sposobem implementacji procesów informacyjnych.
- Chęć przeprowadzenia audytu systemu, własnego - w przypadku operatorów, partnera biznesowego - w przypadku usługodawcy lub użytkownika końcowego. Audyt może być cykliczny, powiązany z chęcią uzyskania lub utrzymania odpowiedniego certyfikatu jakości, lub na żądanie - w szczególności po poważnych incydentach bezpieczeństwa.
- Potrzeba monitorowania stanu bezpieczeństwa systemu lub sieci może również wynikać ze standardowych procedur w danym przedsiębiorstwie, które przewidują regularne przeglądy stanu bezpieczeństwa systemów. Przegląd może mieć charakter długookresowy, np.

coroczny przegląd zgodności używanych rozwiązań kryptograficznych z zaleceniami odpowiednich instytucji zajmujących się bezpieczeństwem i wewnętrznymi standardami korporacyjnymi. W przypadku przeglądów krótkookresowych może to być działanie o charakterze bieżącego monitoringu stanu systemu lub sieci, aby wykryć sytuacje niepożądane takie jak np. przeciążenie systemu, awarie, trwające ataki na system lub sieć. Takie monitorowanie jest potrzebne, ponieważ usługi w sieciach plastrowych są narażone na zakłócenia ich działania, na przykład poprzez oddziaływania wymienione wcześniej w tym rozdziale - które w uproszczeniu powodują niedostępność zasobów kluczowych do realizacji danej usługi, atakują system uniemożliwiając realizację usługi, bądź też atakują dane powiązane z usługą (np. podsłuchiwanie). Tego typu oddziaływania mogą mieć charakter chwilowy, nietrwały, przejściowy, co oznacza, że trzeba z dość dużą częstotliwością monitorować stan systemu i/lub sieci, aby odpowiednio szybko zareagować na negatywne oddziaływanie lub niekorzystny stan. Potrzeba zapewnienia bezpiecznego środowiska do realizacji usług nie jest zatem spełniona raz na zawsze; zapewnianie bezpieczeństwa dla usług to zadanie o charakterze procesu, które należy wykonywać cyklicznie, aby uzyskać zamierzony rezultat.

W związku z tym, że wymagania SLA mogą być ściśle sprecyzowane, samo zapewnienie jakościowe, że system lub sieć działa w pewien ogólny sposób, może nie być wystarczające. Przykładowo, informacja o tym, że połączenie jest szyfrowane, i jest możliwość zestawienia połączenia, nie oznacza, że szyfrowanie jest wykonane z użyciem bezpiecznego klucza o długości przynajmniej 128 bitów, a przepływność binarna łącza wynosi przynajmniej 1 Gbit/s. Określone parametry sieciowe mogą się zmieniać w czasie, niektóre z nich bardzo dynamicznie, jak np. dostępna przepływność binarna łącza, dostępna pamięć RAM, dostępne miejsce na dysku, dostępne połączenia do serwera bazodanowego w ramach ustalonego limitu semaforów, dostępne kanały łączności w komunikacji bezprzewodowej, itd. Inne z nich są wolnozmiennie, np. długość klucza (w bitach) należącego do certyfikatu X.509 użytego do uwierzytelnienia serwera na poczet nawiązania szyfrowanej sesji z użytkownikiem. Te wszystkie informacje są bardzo często zbierane i monitorowane przez organizacje, w celu bieżącego zarządzania systemami informacyjnymi w ramach codziennej pracy operacyjnej. Również w takim przypadku większą użyteczność przedstawiają dane ilościowe niż jakościowe. Operator sieci jest zainteresowany posiadaniem wiedzy na temat dostępnych zasobów, aby móc efektywnie zarządzać tymi zasobami w celu maksymalizacji zysku poprzez podejmowanie odpowiednich decyzji biznesowych

związanych z tymi zasobami. Oznacza to, że operator sieci jest zainteresowany (efektywnym) zarządzaniem tymi zasobami, a dostępność odpowiednich zasobów sieciowych jest kluczowa dla możliwości realizacji usług - odpowiada za dostępność (*availability*) usługi, która jest jednym z trzech elementów triady bezpieczeństwa CIA (*Confidentiality, Integrity, Availability*) [13].

Wiedza na temat stanu izolacji w sieci pozwala świadomie podejmować decyzje o tym, czy należy ten stan poprawić czy też nie, względnie osłabić, aby uzyskać inne korzyści, np. znacząco zmniejszyć koszt obsługi systemu. Zastosowanie ilościowego modelu izolacji pozwala na lepsze porównywanie dwóch rozwiązań (np. architektury obecnej i docelowej), czy też analizować stan systemu wraz z upływem czasu. Dzięki temu możliwe staje się wytyczanie celów dla architektur przejściowych i docelowych danego rozwiązania, w tym tworzenie celów spełniających postulat S.M.A.R.T. (*Specific, Measurable, Achievable, Relevant, Time-bound*), ponieważ stają się one mierzalne.

Założmy na chwilę, że usługodawca lub operator sieci plastrowej chciałby świadczyć usługi lub udostępniać platformę sieci plastrowej bez zarządzania izolacją w sposób ilościowy. Oznacza to, że nie jest obserwowany (monitorowany) stan systemu lub sieci pod kątem zdolności do ochrony jezior i strumieni danych. W skład tych zdolności wchodzi już użytkowana infrastruktura IT, wraz z jej bieżącą konfiguracją sprzętową i programową, która zapewnia zdolność do świadczenia usług w sieci plastrowej oraz do zapewnienia bezpieczeństwa tych usług zarówno w bieżącej chwili (np. firewall, systemy kontroli dostępu, segmentacja sieci, zarządzanie dostępem do kanałów bezprzewodowych) jak i w pewnej przyszłości (np. szyfrowanie ruchu sieciowego). Stan tych systemów i ich konfiguracja zwykle jest znana w organizacji, chociaż ze względu na rozproszenie może obejmować tylko pewien fragment całościowego systemu, który uczestniczy w realizacji danej usługi. Gdyby organizacja faktycznie nie zarządzała izolacją w sposób ilościowy (w żaden sposób), nie miałaby tej wiedzy. W związku z tym, jej możliwość reakcji na zagrożenia (krytyczny incydent lub incydent bezpieczeństwa) byłaby mocno ograniczona, ze względu na brak świadomości na temat bieżącego stanu systemu. Oznacza to, że usługa mogłaby być dłużej niedostępna, dane należące do strumienia danych mogłyby być przez dłuższy czas przedmiotem ataku lub niedostępności usługi.

Zdolność do ochrony jezior i strumieni danych (izolacja) jest przedstawiana przez wiele technik, opisanych szerzej w podrozdziale 3.6.. W szczególności, takie zdolności dostarczane

są przez duże przedsiębiorstwa IT w ramach systemów i platform udostępnianych przez te organizacje - przykłady zostały podane w 3.6.5. i 3.6.6.. Jest to kolejny argument za tym, że zarządzanie izolacją jest ważnym zagadnieniem, skoro powstają zaawansowane rozwiązania komercyjne udostępniające zdolności w tym obszarze.

Oczywiście, można podać przykłady usług, dla których zarządzanie izolacją nie jest ważnym zagadnieniem, a w szczególności opis ilościowy izolacji. Przykładem takich usług są usługi typu non-profit, często hobbystyczne, lub też związane z działalnością gospodarczą o małej skali, dla których nie jest to wiodąca przewaga konkurencyjna. Najczęściej takie usługi nie mają zakontraktowanych wymagań SLA w stosunku do użytkowników końcowych usługi i są świadczone w trybie *best effort*. Naturalnie, z perspektywy operatora telekomunikacyjnego takie usługi będą przedmiotem kontraktu między usługodawcą i operatorem telekomunikacyjnym, ale może to być kontrakt bez zdefiniowanego SLA. W przypadku bardziej zaawansowanych usług, taki zakres kontraktu może być nieakceptowalny dla obydwu stron, na przykład ze względu na wielkość obsługiwanego ruchu, liczbę zgłoszeń, poziom zabezpieczeń samej usługi lub dostępność usługi gwarantowanej przez operatora. W takim przypadku pojawia się potrzeba zarządzania bezpieczeństwem tej usługi (w tym dostępnością usługi dla użytkowników), a więc też zarządzania poziomem ochrony tej usługi (poziomem izolacji). Z perspektywy ekonomicznej, nadmierny poziom izolacji może być nieuzasadniony w niektórych przypadkach, co stanowi kolejną przesłankę do tego, aby zarządzać poziomem izolacji.

3.6. TECHNIKI IZOLACJI

Przedstawmy teraz opis technik izolacji, opierając się na pracy [56], w której zostały one w zbiorczy sposób opisane. Izolacja była rozpatrywana przez wielu autorów jako sposób na poprawienie (zapewnienie) bezpieczeństwa systemów informacyjnych lub ich części. Taksonomia technik izolacji została zaprezentowana w pracy [118], w której podzielono techniki izolacji na izolację komputerową (ang. *isolation in computers*) i izolację sieciową (ang. *isolation in networks*) [56]. Jest to podział intuicyjny, związany pośrednio z pojmowaniem sieci jako klasycznego grafu składającego się z krawędzi i węzłów. Podział ten nie obejmuje wprost technologii wprowadzonych wraz z rozwojem rozwiązań chmurowych takich jak SaaS, PaaS, IaaS (ang. *Software/Platform/Infrastructure-as-a-Service*) [8, 56]. W przypadku sieci SDN mamy również podział na warstwy danych, sterowania i aplikacji [56, 93], a każda z nich może być

rozpatrywana pod kątem izolacji. Możliwe jest również analizowanie fragmentów sieci takich jak RAN czy CN oraz połączenia koniec-koniec (E2E) [4, 56].

Przedstawmy teraz pewną klasyfikację technik izolacji, opartą na podziale wprowadzonym w [118], opisaną w [59] i szerzej w [56]. Są to następujące techniki:

- izolacja w językach programowania (*the language-based isolation*),
- izolacja wykorzystująca środowisko testowe (*the sandbox-based isolation*),
- izolacja wykorzystująca wirtualizację (*Virtual Machine (VM) based isolation*),
- izolacja poziomu jądra systemu operacyjnego (*Operating System - kernel based isolation*),
- izolacja sprzętowa (*the hardware-based isolation*).

Wymienione metody są z powodzeniem stosowane w praktyce, nie tylko w celu uzyskania założonego poziomu bezpieczeństwa i izolacji, ale i również w innych celach, np. optymalizacji zużycia zasobów w wypadku technik wykorzystujących mechanizm wirtualizacji.

3.6.1. Izolacja w językach programowania

Język programowania, kompilator lub interpreter kodu źródłowego może wprowadzać lub wymuszać mechanizmy, które mają właściwości izolacyjne, czyli podnoszą poziom izolacji w systemie informacyjnym. Języki programowania mogą dyktować programiście tworzenie kodu programu, który izoluje go od innych programów [56, 107].

W pracy [56] wydzielono dwie główne gałęzie technik izolowania za pomocą języków programowania:

- izolacja związana z kodem źródłowym - izolacja jest wprowadzana za pomocą semantyki języka, kompilatora lub środowiska uruchomieniowego. Dzięki temu program ma dostęp tylko do poprawnych (autoryzowanych) danych w pamięci operacyjnej [56, 97, 107],
- certyfikowane kompilatory - wykonywana jest weryfikacja kodu źródłowego, a w szczególności pochodzenie kodu. Kod pochodzący z zaufanego źródła przekształcany jest na program wykonywalny, który może zostać opatrzony odpowiednim certyfikatem [30, 56]. Dostępnych jest wiele metod potwierdzania integralności i autentyczności oprogramowania [60]; wykorzystywane w tym celu są np. podpisy cyfrowe, certyfikaty pochodzące z PKI (*Public Key Infrastructure*), algorytmy szyfrowania symetrycznego, kryptograficznie bezpieczne funkcje skrótu, dowody z wiedzą zerową [56].

3.6.2. Izolacja wykorzystująca środowisko testowe

Opiszmy teraz technikę, nazywaną również w literaturze jako *sandboxing* [38, 56, 119]. Idea tej techniki polega na wykonaniu (niezaufanego) kodu programu w taki sposób, że wykorzystywany jest tylko dedykowany obszar pamięci do zapisu danych, oraz możliwe są tylko instrukcje skoku wewnątrz dedykowanego sektora pamięci zawierającego kod wykonywalny. Realizacja praktyczna wymaga wstawienia dodatkowych instrukcji lub instrukcji weryfikujących zachowanie programu przed operacją zapisu lub instrukcją skoku [56]. Praktyczna implementacja wykorzystująca instrukcje procesora jest zależna od architektury procesora, na jakim jest wykonywany *sandboxing* [32, 56]. Możliwe jest również wykonywanie tej techniki na poziomie bibliotek programistycznych, dodając dodatkowe sprawdzenia (asercje) przed wykorzystaniem kodu biblioteki [56]. Trzecim sposobem realizacji *sandddboxingu* jest zarządzanie uprawnieniami do urządzeń wejścia/wyjścia takich jak dostęp do systemu plików, interfejsów sieciowych, urządzeń peryferyjnych, itp. [56, 67].

3.6.3. Izolacja wykorzystująca wirtualizację

Kolejną techniką izolacji przedstawioną w pracy [56] jest wykorzystanie środowiska maszyn wirtualnych, które pozwalają oddzielić między sobą poszczególne maszyny uruchomione na fizycznym hoście, oraz środowisko hosta pełniące rolę platformy do uruchamiania maszyn wirtualnych. W pracy [118] wyszczególniono sposoby realizacji maszyn wirtualnych i separacji (izolacji) między nimi - od poziomu procesu i systemu operacyjnego, aż do poziomu sprzętowego, gdzie wsparcie dla wirtualizacji jest natywnie dostarczane przez dostępny sprzęt. Zasadniczo za zapewnienie izolacji pomiędzy maszynami wirtualnymi odpowiada oprogramowanie hipernadzorcy [10, 56] (*hypervisor*), które można wykorzystywać dostępne wsparcie sprzętowe. Pokrewną techniką jest izolacja z użyciem konteneryzacji.

3.6.4. Izolacja sprzętowa i izolacja poziomu jądra systemu operacyjnego

Jednym z najsilniejszych sposobów izolacji jest izolacja sprzętowa [56]. Idea izolacji sprzętowej polega na fizycznym odseparowaniu procesów informacyjnych od siebie wzajemnie i potencjalnie również od świata zewnętrznego. W zależności od wymagań biznesowych stosowane są różne poziomy izolacji sprzętowej. Przytoczmy kilka przykładów.

- Użytkowanie osobnych maszyn umieszczonych w osobnych serwerowniach w osobnych regionach geograficznych, oddalonych od siebie na znaczną odległość, korzystających z niezależnych od siebie dostawców mediów (prąd, dostęp do sieci Internet). Przykładem takiego rozwiązania jest podział na *Availability Regions* [73] w chmurze Azure, które realizują replikację danych i aplikacji z opóźnieniem rzędu 2 ms [75].
- Wykorzystywanie dedykowanych maszyn w ramach dostępnej infrastruktury do realizowania potrzeb biznesowych. Możliwe jest współdzielenie niektórych elementów sieciowych w ramach takiej konfiguracji, szczególnie jeżeli usługa nie jest wrażliwa na tymczasowy spadek dostępnej przepływności binarnej łączy.
- wykorzystanie dedykowanych zasobów w ramach współdzielonej maszyny, np. aplikacja zapisuje dane na dedykowanym dysku i korzysta z dedykowanego procesora logicznego. Pozwala to na zagwarantowanie zasobów obliczeniowych dla procesu oraz mityguje ryzyko wyczerpania wolnego miejsca na dysku przez inne aplikacje, które wykorzystują ten sam dysk.
- Komponenty i techniki izolujące wykonywany proces informacyjny od świata zewnętrznego, np. SGX (ang. *Software Guard Extensions*) implementowane przez Intelą [70, 124], czy też TPM (ang. *Trusted Platform Module*) [42].

Izolacja poziomu jądra systemu operacyjnego pozwala na izolowanie procesów (aplikacji) uruchamianych w danym systemie. Bezpieczeństwo zależy tutaj ściśle od bezpieczeństwa samego systemu operacyjnego [56, 127]. Pokrewną techniką jest izolacja wykorzystująca mechanizm konteneryzacji, np. oprogramowanie Docker [20]. W założeniu aplikacja uruchomiona w kontenerze może oddziaływać ze światem zewnętrznym tylko za pomocą dedykowanych interfejsów i urządzeń (zwykle wirtualnych); w praktyce wirtualne (logiczne) komponenty są realizowane przez fizyczne urządzenia np. fizyczny procesor lub karta sieciowa.

3.6.5. Separacja domen i procesów

Techniki izolacji wykorzystywane i rozwijane są również w komercyjnym oprogramowaniu niezwiązanym bezpośrednio z sieciami telekomunikacyjnymi. Przykładem takiego rozwiązania jest platforma ServiceNow, która jest oprogramowaniem klasy ITSM (*Information Technology Service Management*), służące (w dużym uproszczeniu) do zarządzania usługami IT w ramach danego przedsiębiorstwa. W przypadku firm świadczących usługi z sektora ICT (*Information*

and Communication Technology), mogą to być świadczenia realizowane na rzecz kontrahentów, z wykorzystaniem jednego lub wielu dostawców rozwiązań IT - poczynając od infrastruktury, przez usługi wsparcia technicznego (*technical support*), po zaawansowane usługi konsultacyjne. Ze względu na mnogość klientów, wymagania kontraktowe i biznesowe powstaje potrzeba oddzielenia danych i procesów informacyjnych należących do poszczególnych klientów i/lub dostawców usług.

Aby zrealizować wspomniany podział, platforma ServiceNow udostępnia dwa rozwiązania: separację domen oraz separację procesów [110]. Separacja domen polega na rozszerzeniu modelu danych o informację, do jakiej domeny należy dany wiersz w tabeli bazodanowej (platforma ServiceNow korzysta z relacyjnych baz danych, np. MySQL/MariaDB). Domeny są między sobą powiązane - tworzą graf będący drzewem, w którym węzły odpowiadają domenom, a skierowane krawędzie odpowiadają relacji zawierania się jednej domeny w drugiej. Pozwala to uwspólniać pewne dane poprzez przechowywanie ich we współdzielonej domenie, np. w korzeniu drzewa. Wspólnienie danych oznacza intencjonalne wprowadzenie braku pełnej izolacji między poszczególnymi domenami, co niesie za sobą istotne konsekwencje z perspektywy utrzymania systemu - utrudnia rozwój funkcji wykorzystujących te dane, zmiany dedykowane jednemu klientowi oddziałują na innych klientów, powstaje naturalna możliwość udostępnienia danych w sposób niekontrolowany dla nieautoryzowanych użytkowników.

Drugim filarem obecnym na platformie ServiceNow jest separacja procesów. Pozwala ona wykonywać kod źródłowy (czyli realizować proces informacyjny) w taki sposób, że wykorzystywane jest kod źródłowy zdefiniowany w domenie związanej z danym rekordem (lub użytkownikiem, jaki wykonuje kod - w zależności od ustawienia w systemie). W przypadku braku konkretnej implementacji logiki biznesowej w danej domenie, wykonywany jest kod zdefiniowany w domenie nadrzędnej; w przypadku braku takiej implementacji, brane są kolejne domeny nadrzędne pod uwagę, aż do korzenia włącznie. Mechanizm jest podobny do mechanizmu funkcji wirtualnych znanych z obiektowych języków programowania.

Platforma ServiceNow pozwala rozwiązywać problemy związane z separacją strumieni i jezior danych za pomocą innych sposobów, przedstawionych na rysunku 3.1.



Rysunek 3.1: Przedstawienie kompleksowości potrzeb biznesowych w zakresie separacji strumieni danych i jezior danych na platformie ServiceNow. Szerokość poziomu piramidy symbolizuje częstotliwość z jaką należy wybierać poszczególne rozwiązania. Poziomy wyższe potrafią rozwiązywać bardziej skomplikowane problemy niż poziomy niższe. Rysunek utworzony na podstawie [109].

3.6.6. Izolacja w środowisku chmurowym

Kolejnym ważnym miejscem, gdzie może być rozpatrywana izolacja jest środowisko chmurowe, które jest ciekawą alternatywą dla rozwiązań utrzymywanych bezpośrednio przez organizację (tzw. *on-premise*). Pierwszym poziomem izolacji jest charakter chmury - może być ona publiczna, prywatna lub hybrydowa. Jednym z modeli biznesowych jest wykorzystywanie oferty złożonych usług dostawcy infrastruktury IT, który z kolei do realizacji swojej usługi IT wykorzystuje innego dostawcę infrastruktury (chmurowego). Rozwój narzędzi chmurowych wspomaga trend w podejściu wielochmurowym (ang. *multicloud*), aby uniknąć uzależnienia się od jednego dostawcy bądź też otrzymać bardzo zaawansowane rozwiązania zapewniające skalowalność i dostępność.

W ramach środowiska chmurowego mogą być dostępne różne centra danych, zgrupowane w regiony. Na rynku dostępne są rozwiązania (np. chmura Microsoft Azure) dostarczające infrastrukturę w sposób redundantny, pomiędzy różnymi regionami geograficznymi, z odpowiednimi strategiami aktualizacji oprogramowania i sprzętu, które gwarantują dostępność usługi w zasadzie bez przerwy. Operator chmurowy korzysta ze zjawiska opisywanego przez prawo

wiązki - większa wiązka (system chmurowy) jest łatwiej obciążalna i trudniej przeciązalna. Z perspektywy użytkownika chmury (dostawcy usługi wykorzystującej chmurę obliczeniową) zasoby można praktycznie dowolnie skalować w szerz (np. uruchamiać kolejne wirtualne maszyny, kontenery, instancje tzw. *serverless functions*) i z pewnymi ograniczeniami skalować w górę lub w dół, zwiększając lub zmniejszając zasoby sprzętowe danej maszyny wirtualnej.

Dostawcy usług chmurowych mogą w szczególnych przypadkach dokonać daleko idących działań, aby zapewnić odpowiedni poziom izolacji świadczonych usług. Stosunkowo popularnym wymaganiami biznesowymi stała się rezydencja danych na obszarze konkretnego państwa lub na terenie Unii Europejskiej - głównym akceleratorem było tutaj rozporządzenie RODO (GDPR - *General Data Protection Regulation*) [28] o ochronie danych osobowych; obecnie spodziewany jest podobny trend w przypadku dyrektywy NIS2 (*Network and Information Security 2 Directive*) [27]. W bardziej skomplikowanych przypadkach wymagana jest konkretna fizyczna lokalizacja osób utrzymujących platformę chmurową, z potencjalnym dostępem do danych. Tak daleko idące wymagania pojawiają się przy przetwarzaniu danych wrażliwych - danych medycznych, podatkowych, policyjnych, czy też danych szczególnie chronionych ze względu na specyfikę działania przedsiębiorstwa - danych objętych tajemnicą bankową, danych powiązanych z infrastrukturą krytyczną, itd. Warto tutaj wymienić przykład dedykowanej chmury Microsoft Azure dla rządu USA [71], który dostarcza usługi dla danych z nadaną klauzulą *Top Secret* [72], odpowiadającą polskiej klauzuli *Ścisłe Tajne*.

W pracy [79] zaproponowano wykorzystanie chmury w architekturze PaaS z wykorzystaniem konteneryzacji jako techniki do zapewniania izolacji [56]. Praca [63] poruszono problematykę zapewniania izolacji w centrach danych; izolacja pomiędzy klientami (tenantami) miała być realizowana za pomocą sieci plastrowej, natomiast dostawca usługi ograniczał się jednocześnie tylko do izolacji w kontekście wydajności [56]. Warto również wspomnieć o wdrożeniach w systemach chmurowych rozwiązań zapewniających izolację na poziomie sprzętowym, np. Intel SGX udostępniany na platformie Microsoft Azure [74].

3.6.7. Techniki izolacji w sieci

Izolacja w sieci zasadniczo realizowana jest poprzez zastosowanie odpowiedniego protokołu komunikacyjnego, najczęściej należącego do jednej z warstw modelu ISO/OSI RM. Praca [19]

zawiera analizę i porównanie 15 protokołów pod kątem wydajności, odporności i skalowalności. Techniki izolacji w sieci możemy kategoryzować na różne sposoby, przykładowo:

- Podział według podejścia do tworzenia plastra sieci:
 - tunelowanie - ruch sieciowy jest oddzielony od pozostałej części sieci i współistniejącego ruchu pomiędzy dwoma ustalonymi punktami [56],
 - sieć wirtualna (*Virtual Network*) - fragment ruchu sieciowego zostaje logicznie wydzielony z wiązki ruchu za pomocą wirtualnych podsieci realizowanych z wykorzystaniem fizycznej infrastruktury [56],
 - wielo-klienckość (*multi-tenancy*) - strumienie danych generowane przez różnych klientów są od siebie odseparowane za pomocą zestawu tuneli lub sieci wirtualnych [56].
- Podział według konkretnej implementacji protokołów:
 - protokoły etykietowe, takie jak MPLS (*Multiprotocol Label Switching*) [59],
 - protokoły typu VLAN (*Virtual Local Area Network*) [34, 59] - segmentacja sieci,
 - protokoły typu VPN, tworzące bezpieczne tunele komunikacyjne, przykładowo [59]:
 - IPSec (*Internet Protocol Security*),
 - SSL/TLS (*Secure Socket Layer / Transport Layer Security*),
 - DTLS (*Datagram Transport Layer Security*),
 - MPPE (*Microsoft Point-to-Point Encryption*),
 - SSTP (*Secure Socket Tunneling Protocol*),
 - SSH (*Secure Shell*).

Protokoły użyte do realizacji izolacji mogą zapewniać ją w różnym stopniu. Przykładowo, protokół MPLS pozwala na logiczną separację strumieni danych, ale nie zapewnia poufności przesyłanych danych, narażając strumień danych na podsłuchanie [56]. Protokoły wykorzystujące prymitywy kryptograficzne takie jak szyfrowanie czy podpis cyfrowy pozwalają na silniejszą izolację strumienia danych, zwykle kosztem zmniejszonej wydajności.

Wartym wspomnienia jest również kwestia izolacji sieciowej w kontekście sprzętowym (warstwa 1 modelu ISO/OSI RM). Komunikacja może zostać podsłuchana lub przerwana na poziomie elektrycznym lub optycznym. W praktyce poważnym ryzykiem jest również zdarzenie przzerwania ciągłości wiązki (elektrycznej lub optycznej), co skutkuje brakiem komunikacji. Wypadki tego typu mogą wystąpić na skutek prac budowlanych, przypadkowego uszkodzenia kabli podmorskich lub też zdarzeń o charakterze sabotażu.

Tablica 3.1: Przykładowe techniki izolacji w sieciach RAN, tabela zbudowana na podstawie [56]

Warstwa modelu ISO- /OSI RM	Sposób realizacji izolacji
Fizyczna	multipleksacja P-OFDM (<i>Polar-Orthogonal Frequency Division Multiplexing</i>)[22], f-OFDM (<i>filtered-Orthogonal Frequency Division Multiplexing</i>) [125], komunikacja z wykorzystaniem wielu anten (MIMO - <i>Multiple Input, Multiple Output</i>), SDR (<i>Software Defined Radio</i>)
Łączy danych	protokół MAC , technika CSMA/CA (<i>Carrier Sense Multiple Access with Collision Avoidance</i>), sygnalizacja RTS/CTS
Inne	protokół IPv4, IPv6

3.6.8. Izolacja w sieciach RAN i CN

Izolacja w sieci RAN, zarówno na poziomie węzłów, jak i łączy pomiędzy węzłami sieci była rozpatrywana przez prace [46, 56]. Przykładowe sposoby realizacji izolacji w sieci RAN są podane w tabeli 3.1. W rozdziale 3.3. wprowadzono trzy podstawowe typy izolacji: izolacja pełna, częściowa i brak izolacji. W przypadku wyżej wymienionych technik izolacji w sieciach RAN zjawisko izolacji realizowane jest za pomocą odpowiedniego zarządzania zasobami sieciowymi. Aby uzyskać wysoki poziom izolacji, pożądane jest wyodrębnienie zasobów sieciowych na poczet danego plastra. Przykładowo, takimi zasobami mogą być:

- szczeliny czasowe w multipleksacji w domenie czasu (TDM - *Time-Division Multiplexing*),
- kanały częstotliwościowe w przypadku multipleksacji w domenie częstotliwości (FDM - *Frequency Division Multiplexing*),
- długości fal w przypadku zwielokrotnienia w dziedzinie długości fali (WDM - *Wavelength Division Multiplexing*),
- kanały kodowe w przypadku multipleksacji kodowej (CDM - *Code Division Multiplexing*).

Niektóre techniki jak multipleksacja w dziedzinie częstotliwości, nie zapewniają idealnej separacji kanałów - co wynika z faktu, że zasadniczo każdy sygnał nadawany przez rzeczywisty nadajnik ma nieskończenie szerokie widmo mocy. Niekiedy ma to istotne znaczenie, przykładowo

w sieciach Wi-Fi wzajemny wpływ sąsiednich kanałów jest na tyle znaczący, że sąsiadujące ze sobą sieci nie powinny używać bliskich sobie kanałów częstotliwościowych [59]. Dodatkowo, ze względu na użycie elementów nieliniowych takich jak diody, tranzystory, wzmacniacze operacyjne, nie jest zachowana zasada superpozycji - co oznacza, że na wyjściu takich elementów możemy zaobserwować pewne sygnały uboczne [56]. W dziedzinie optycznej występuje podobne zjawisko, tj. mieszanie czterofalowe, gdzie podczas transmisji trzech sygnałów optycznych o różnych długościach fali, powstaje dodatkowy sygnał, o długości fali (częstotliwości) zależnej od trzech pozostałych [9].

W przypadku warstwy łącza danych (warstwy 2 według modelu ISO/OSI RM), zwykle istnieje protokół typu MAC, który pozwala współdzielić łącze pomiędzy węzłami sieci [56]. Przykładem techniki z tej warstwy (raczej skojarzonej z sieciami Wi-Fi) jest sygnalizacja RTS/CTS, [41, 56]. Rozwiązuje ona problem ukrytego węzła, na jaki podatny jest protokół CSMA (*Carrier Sense Multiple Access*), wykorzystywany do zapewniania wielodostępu w sieciach bezprzewodowych. Ciekawym zjawiskiem jest tutaj wykorzystanie braku izolacji w kanale komunikacyjnym, do synchronizacji węzłów, aby zapewnić izolację (brak kolizji).

Wyższe warstwy modelu ISO/OSI RM (warstwy 3-7) w przypadku części RAN i CN można przeanalizować wspólnie. Na poziomie warstwy sieciowej aktualnie występuje współegzystencja protokołów IPv4 i IPv6 - przykładowo, 41% użytkowników usług Google używa protokołu IPv6 [33]. Obydwa protokoły przewidują pola związane z QoS - TOS w przypadku IPv4 i *Traffic Class* w przypadku IPv6; w obydwu protokołach są to pola 8-bitowe, które pozwalają na zapisanie 256 różnych wartości - co jest wartością stosunkowo małą i nie pozwala na skalowanie rozwiązania [56]. Możliwa jest obsługa niestandardowych rozszerzeń nagłówka IPv6, które mogą przenosić więcej informacji na potrzeby oznaczania pakietów IPv6 i przypisywania ich do poszczególnych plastrów.

Sieć CN, nazywa również siecią szkieletową, co do zasady preferuje łącza przewodowe do realizacji połączeń pomiędzy węzłami, w przeciwieństwie do sieci RAN. Naturalnymi rozwiązaniami w warstwie fizycznej są tutaj łącza miedziane i światłowodowe, co ma również pewien wpływ na protokoły warstwy łącza danych. W przypadku warstwy sieciowej i wyższych, zasadniczo możliwe są takie same techniki izolacji, co w przypadku sieci RAN.

4. Algorytm wyznaczania izolacji

W niniejszym rozdziale przedstawimy algorytm wyznaczania izolacji, który będzie punktem wyjścia do wyznaczania izolacji za pomocą skierowanych liczb rozmytych i technik probabilistycznych.

4.1. HIPERGRAFOWY MODEL SIECI

Celem modelu sieci jest przedstawienie powiązania cech izolacji z węzłami sieci i w dalszej kolejności z plastrami sieci i usługami, jakie są realizowane w ramach tych plastrów. Zagadnienia związane z *izolacją* w innych kontekstach niż kontekst sieciowy (sieci teleinformatycznej) są definiowane i rozumiane zwykle w dość ograniczonym zakresie. Przytoczmy kilka przykładów.

- Izolator termiczny - materiały lub konstrukcje mają różne przewodności cieplne, co pozwala zbudować urządzenia, które dobrze odprowadzają ciepło (radiatory), bądź nie przewodzą ciepła - np. ocieplenie budynków, uchwyty z tworzywa sztucznego, termos próżniowy.
- Izolator elektryczny - powszechnie stosowane są dielektryki, szczególnie w celu zwiększenia pojemności kondensatorów, które jednocześnie bardzo słabo przewodzą prąd elektryczny.
- Rezystancja i impedancja - wielkość określająca stosunek napięcia elektrycznego do prądu elektrycznego. W elektronice mierzone i wyznaczane są różne rodzaje rezystancji i impedancji, szczególnie w przypadku układów nieliniowych i wielowrotników. Wysoka wartość rezystancji oznacza, że po przyłożeniu napięcia do badanego elementu (na konkretne zaciski tego elementu) popłynie mały prąd, zadany przez prawo Ohma.

Zasadniczym problem w określaniu izolacji (w zasadzie dowolnej dziedzinie) jest zdecydowanie, które przepływy izolowanego zjawiska (takiego jak przepływ ciepła, przepływ prądu elektrycznego, przepływ informacji) są najistotniejsze, a wobec czego, jakie elementy skomplikowanego świata rzeczywistego należy ująć w modelu izolacyjnym danego zjawiska. W przypadku izolacji informacyjnej, jaką zajmujemy się w tej pracy, interesują nas sposoby wpływu

na procesy informacyjne realizowane w danym systemie informacyjnym. Jak zostało omówione w podrozdziale 3.4., zagrożenia dla (wpływ na) procesów informacyjnych mogą pochodzić z wnętrza plastra, innych plastrów oraz ze świata zewnętrznego, włączając w tę kategorię również sieci pozaplastrowe (*legacy*). W niniejszej pracy nie będziemy uwzględniać źródła tych zagrożeń pod kątem ilościowym.

W pracy [57] zaproponowano koncepcję wielowarstwowego hipergrafu, gdzie poszczególnym elementom sieciowym przypisano węzły sieci - przedstawiają ją rysunki 4.1 i 4.2. Każdy węzeł zawierał w sobie hipergraf, reprezentujący komponenty składowe wewnątrz węzła nadrzędnego i zależności między nimi. Na najniższym z poziomów (warstwa 5, ang. *Layer 5*) węzeł hipergrafu był opisany zbiorem cech izolacji i wartości właściwych dla danego węzła (patrz rys. 4.2); możliwe również jest dodawanie dodatkowych cech w wyższych warstwach. Wierzchołek w tym modelu może reprezentować urządzenie sieciowe albo łącze pomiędzy urządzeniami sieciowymi. Hiperkrawędź w tym modelu oznaczała połączenie pomiędzy wierzchołkami, być może więcej niż dwoma jednocześnie. Zaletą tego modelu jest duży poziom ogólności i abstrakcji, co łączy się jednocześnie z pewnym skomplikowaniem struktury.

Na potrzeby niniejsze pracy uprościmy model przedstawiony w pracy [57] do następującego modelu sieci.

Niech dany będzie graf $G = (V, E)$, gdzie V to zbiór wierzchołków a E to zbiór krawędzi. Wierzchołek reprezentuje element sieciowy lub łącze sieciowe, na przykład router, przełącznik (*switch*), serwer usług, łącze miedziane, łącze światłowodowe, łącze bezprzewodowe. Krawędź symbolizuje możliwość przepływu informacji pomiędzy wierzchołkami. Każdy z wierzchołków jest opisany indywidualnym zbiorem cech; cechy są szerzej opisane w dalszej części tego rozdziału. Graf G może być podzielony na skończony zbiór podgrafów $G = G_1 \cup G_2 \cup \dots \cup G_S$, niekoniecznie rozłącznych, gdzie dany podgraf symbolizuje pewną logiczną podsieć. Wierzchołek należący do więcej niż jednego podgrafu leży w zakresie zainteresowań interesariuszy związanych z tymi podsieciami, wobec czego jego wpływ na izolację może być silniejszy.

Aby przejść do dalszej części rozważań, należy odpowiedzieć na pytanie, które wierzchołki grafu G należy wziąć pod uwagę przy analizie izolacji. Można zastosować następujące podejścia.

- Rozpatrywać tylko wierzchołki biorące bezpośredni udział w procesie informacyjnym, odpowiadające za przenoszenie (transport) danych w strumieniu danych. Nazwijmy takie wierzchołki *wierzchołkami podstawowymi*.
- Rozpatrywać wierzchołki podstawowe oraz wierzchołki osiągalne z wierzchołków podstawowych za pomocą ścieżek o co najwyżej d krawędziach. Zbiór tych wierzchołków nazwijmy *otoczeniem strumienia danych o promieniu d* .

Zadanie wyznaczenia wierzchołków podstawowych może być zadaniem nietrywialnym, warto tutaj przytoczyć następujące argumenty na poparcie tej tezy.

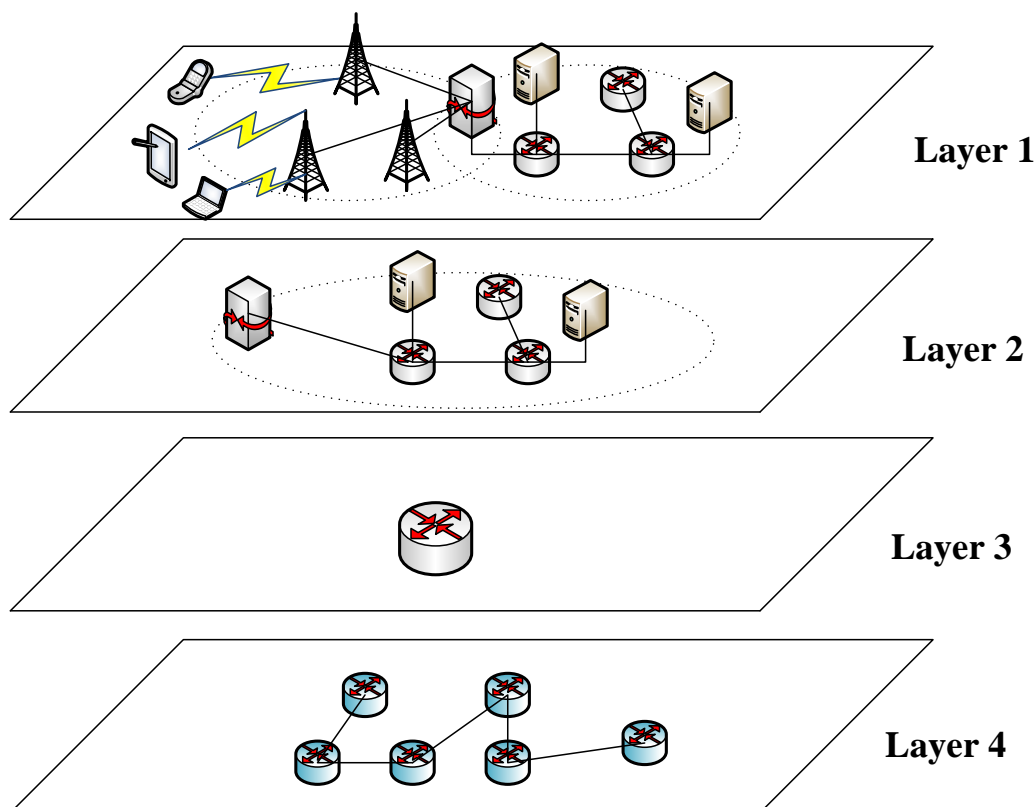
- Trasa jaką podąży strumień informacyjny może się zmieniać w czasie, niekiedy bardzo dynamicznie, ze względu na stosowanie narzędzi do równoważenia obciążenia sieciowego (*load balancers*).
- Dane przetwarzane przez usługę mogą być przesyłane do innych węzłów takich jak serwery bazodanowe, serwery *cache*, innych serwerów usługowych; wyznaczenie tych wierzchołków wymaga dokładnej znajomości architektury systemu, a niekiedy i szerokiej wiedzy na temat architektury korporacyjnej danego przedsiębiorstwa.
- Istotne jest zidentyfikowanie wierzchołków, które mimo braku aktywnego uczestnictwa w założonym procesie informacyjnym, mogą do niego zostać przyłączone na skutek rozbieżności pomiędzy zaplanowaną architekturą systemu informacyjnego, a jego implementacją praktyczną. Przykładowo, administrator systemu bazodanowego może łączyć się do serwera bazodanowego za pośrednictwem publicznej sieci Wi-Fi udostępnianej przez kawiarnię, w której wykonuje pracę zdalną. Ponadto ekran jego komputera nie jest zabezpieczony filtrem prywatyzacyjnym, co pozwala na zarejestrowanie ekranu jego komputera przez kamery monitoringu zainstalowane w lokalu.

W niniejszej pracy będziemy rozpatrywać wierzchołki podstawowe przy analizie izolacji.

4.2. PARAMETRY I WŁAŚCIWOŚCI

Poziom izolacji (zmierzony bądź wyznaczony) zależy od wartości pewnych sieciowych i fizycznych cech (*traits*). W pracach [56, 57, 59] wyróżniono dwie kategorie cech:

- parametry (*parameters*) - cechy ilościowe, najczęściej mierzone, o wartościach ciągłych; pomijamy tutaj ograniczenia wynikające z zasad mechaniki kwantowej czy skończonej reprezentacji liczb w systemach komputerowych.



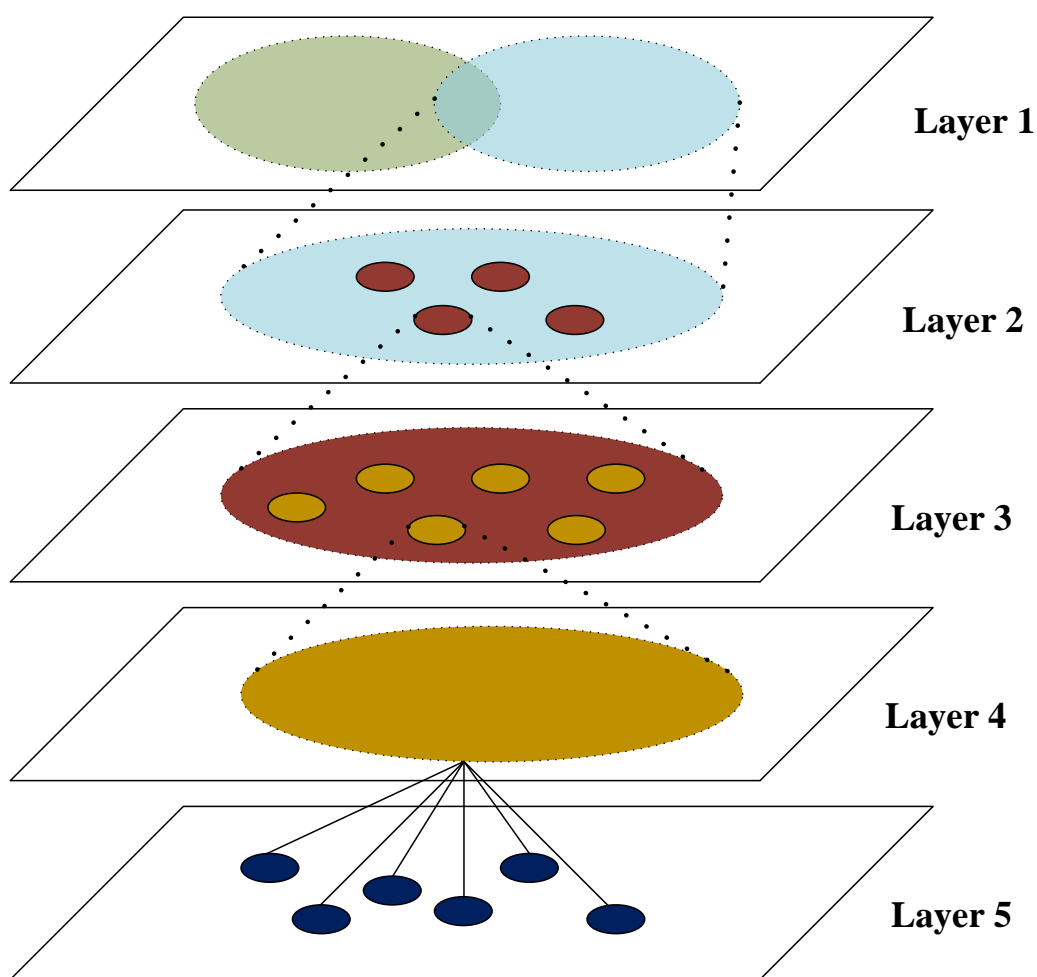
Rysunek 4.1: Wielowarstwowy model sieci - perspektywa sieciowa. Rysunek pochodzi z pracy [57].

— właściwości (*properties*) - cechy jakościowe, które mają charakter enumeracyjny. Mogą być im przypisane liczby ze zbioru Ω , co pozwala je traktować w obliczeniach tak jak parametry.

Istnieje jeszcze trzecia kategoria cech, zwana metrykami (*metrics*) [48, 56, 59]. Metryki są abstrakcyjnymi i syntetycznymi cechami, bez prostej interpretacji sieciowej czy fizycznej, a które jednocześnie związane są z poziomem izolacji.

Cechy opisują elementy sieciowe, co wykorzystujemy do wyznaczania izolacji. Dobór cech jest dla tej operacji kluczowy - powinny one mieć związek z zagadnieniem izolacji, zabezpieczania systemu informacyjnego, odpornością na ataki, czy wzajemne wyczerpanie zasobów przez usługi. Przykładowy zestaw cech został zaproponowany w artykułach [56, 59]; praca [56] proponuje również przykładowe metody wyznaczania wartości niektórych cech.

Wartości cech na potrzeby algorytmu wyznaczania izolacji muszą zostać znormalizowane. W pracach [56, 57] zaproponowano następującą metodę normalizacji, wykorzystującą



Rysunek 4.2: Wielowarstwowy model sieci - perspektywa grafowa. Rysunek pochodzi z pracy [57].

funkcje normalizujące (*normalization functions*) $g : \Lambda \rightarrow \Omega$, gdzie Λ jest dziedziną cechy, a Ω jest ciągłym podzbiorem \mathbb{R} . W niniejszej pracy będziemy przyjmować $\Omega = [0, 1]$. Zbiór Ω powinien być wspólny dla wszystkich cech. Wprowadzamy również oznaczenia:

$$\alpha = \inf_{x \in \Omega} x, \quad (4.1)$$

$$\omega = \sup_{x \in \Omega} x. \quad (4.2)$$

Wartości α odpowiada najgorsza wartość cechy (z perspektywy izolacji), analogicznie ω odpowiada najlepsza wartość cechy (z perspektywy izolacji) [56, 57]. Funkcja normalizująca powinna spełniać zależności [56, 57]:

1. $g(x) \leq g(y)$ wtedy i tylko wtedy, gdy y jest lepszą wartością (z perspektywy izolacji) niż x .
2. $g(\beta) = C$, gdzie C jest stałą zależną od zbioru Ω , natomiast β jest pewną typową wartością dla danej cechy (w kontekście danego elementu sieciowego). Tak jak w pracach [56, 57] proponujemy $C = \frac{\alpha + \omega}{2} = 0.5$. Wartość β może być np. dominantą albo pewną wartością przyjętą na podstawie znajomości kontekstu technicznego i sieciowego.

Przyjmujemy, że każda cecha może być normalizowana za pomocą własnej funkcji normalizującej. W szczególności, ta sama cecha obecna w różnych elementach sieciowych nie musi mieć jednakowej funkcji normalizującej.

W pracach [56, 57] wyszczególniono ze względu na kształt trzy podstawowe rodziny cech:

- cecha rosnąca - im większa wartość cechy, tym lepsza izolacja (np. dostępna przepływność binarna w Mbit/s),
- cecha malejąca - im mniejsza wartość cechy, tym lepsza izolacja (np. BER - *Bit Error Rate* łącza),
- cecha gaussowska - cecha, która jest rosnąca dla wartości mniejszych niż γ , i malejąca dla wartości większych niż γ , gdzie γ to pewna stała (np. jitter między pakietami w strumieniu pakietów).

Tabela 4.1 zawiera przykładowe funkcje normalizacyjne, dla typowych dziedzin. Funkcje te zwykle są parametryzowane za pomocą stałych, co pozwala dopasować kształt ich wykresu do występujących w praktyce wartości.

4.3. PODSTAWY MATEMATYCZNE

Wprowadźmy teraz podstawy matematyczne, zaprezentowane w [56, 57], używany w algorytmie wyznaczania izolacji z wykorzystaniem liczb rzeczywistych. W dalszych częściach niniejszej pracy uogólniamy przedstawione rozumowanie na skierowane liczby rozmyte. Jest ono również podstawą dla algorytmu wykorzystującego metody probabilistyczne.

Załóżmy, że mamy zbiór wierzchołków V_1, V_2, \dots, V_N , opisanych za pomocą wspólnego zbioru cech $\Pi = \{\pi_1, \pi_2, \dots, \pi_M\}$. Dla wierzchołka $V_n, n = 1, 2, \dots, N$ oznaczmy wektor

Tablica 4.1: Przykłady funkcji normalizujących, podanych w pracach [56, 57]

Λ	Rodzina cechy	Najgorsza wartość	Najlepsza wartość	Typowa wartość	Funkcja $g(x)$
\mathbb{R}	rosnąca malejąca	0	$\pm\infty$	β	$g(x) = 1 - 2^{-\frac{x}{\beta}}$
\mathbb{R}	malejąca rosnąca	$\pm\infty$	0	β	$g(x) = 2^{-\frac{x}{\beta}}$
\mathbb{R}	rosnąca	$-\infty$	$+\infty$	$\beta; \beta \neq 0$	$g(x) = \left(1 + e^{\frac{\beta-x}{ \beta }}\right)^{-1}$
\mathbb{R}	rosnąca	$-\infty$	$+\infty$	0	$g(x) = (1 + e^{-x})^{-1}$
\mathbb{R}	malejąca	$+\infty$	$-\infty$	$\beta; \beta \neq 0$	$g(x) = \left(1 + e^{\frac{x-\beta}{ \beta }}\right)^{-1}$
\mathbb{R}	malejąca	$-\infty$	$+\infty$	0	$g(x) = (1 + e^x)^{-1}$
$[q; r] \in \mathbb{R}_{\geq 0}$	rosnąca	q	r	$\frac{q+r}{2}$	$g(x) = \frac{x-q}{r-q}$
$[q; r] \in \mathbb{R}_{\geq 0}$	malejąca	r	q	$\frac{q+r}{2}$	$g(x) = \frac{r-x}{r-q}$
\mathbb{R}	gaussowska	$\pm\infty$	γ	β	$g(x) = 2^{-\left \frac{x-\gamma}{\beta-\gamma}\right }$

wartości cech jako:

$$\mathbf{I}_n \stackrel{\text{def}}{=} (p_{n,1}, p_{n,2}, \dots, p_{n,M})^T, \quad (4.3)$$

natomiast dla całego zbioru wierzchołków wprowadzamy:

$$\mathbf{I}(\{V_1, V_2, \dots, V_N\}) \stackrel{\text{def}}{=} F(\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_N), \quad (4.4)$$

gdzie funkcja $F : (\Omega^M)^N \rightarrow \Omega^M$ jest ogólną funkcją łączącą [56, 57]. Na tym etapie będziemy zakładać, że cechy i ich wartości są wzajemnie niezależne. Wobec tego, możemy zdefiniować

funkcję F jako:

$$F(\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_N) = \begin{pmatrix} f_1(p_{1,1}, p_{2,1}, \dots, p_{N,1}) \\ f_2(p_{1,2}, p_{2,2}, \dots, p_{N,2}) \\ \vdots \\ f_m(p_{1,M}, p_{2,M}, \dots, p_{N,M}) \end{pmatrix}, \quad (4.5)$$

gdzie $f_i : \Omega^N \rightarrow \Omega$, dla $i = 1, 2, \dots, M$ jest funkcją łączącą [56, 57]. Zastanówmy się teraz nad właściwościami funkcji łączącej. Niech $x_{min} = \min\{x_1, x_2, \dots, x_N\}$ i $x_{max} = \max\{x_1, x_2, \dots, x_N\}$. Załóżmy, że funkcja f spełnia zależności [56, 57]:

$$\left(\bigwedge_{x \in \Omega} \right) f(x, x, \dots, x) = x, \quad (4.6)$$

$$\left(\bigwedge_{n=1}^N \right) \left(\bigwedge_{x_n \leq y_n} \right) f(x_1, x_2, \dots, x_n, \dots, x_N) \leq f(x_1, x_2, \dots, y_n, \dots, x_N). \quad (4.7)$$

Powyższe warunki opisują nasze spojrzenie na izolację od strony ilościowej - system zbudowany z komponentów o zadanym poziomie izolacji, efektywnie ma taką samą izolację [57]. Druga właściwość opisuje niemalejący charakter wypadkowej izolacji, gdy nastąpi polepszenie jednego z parametrów izolacji [57]. Z powyższych założeń dostajemy następujące ciągi nierówności [57]:

$$\begin{aligned} f(x_1, x_2, \dots, x_N) &\leq f(x_{max}, x_2, \dots, x_N) \leq \\ &\leq f(x_{max}, x_{max}, \dots, x_N) \leq \dots \leq f(x_{max}, x_{max}, \dots, x_{max}) = x_{max}, \end{aligned} \quad (4.8)$$

$$\begin{aligned} f(x_1, x_2, \dots, x_N) &\geq f(x_{min}, x_2, \dots, x_N) \geq \\ &\geq f(x_{min}, x_{min}, \dots, x_N) \geq \dots \geq f(x_{min}, x_{min}, \dots, x_{min}) = x_{min}. \end{aligned} \quad (4.9)$$

Funkcja, która spełnia nierówności:

$$x_{min} \leq f(x_1, x_2, \dots, x_N) \leq x_{max} \quad (4.10)$$

będzie nazywana *średnią* (ang. *the mean function*) [49, 57]. Przykładową *średnią* jest *średnia potęgowa* (ang. *the generalized mean function*):

$$\omega_q(x_1, x_2, \dots, x_N) = \left(\frac{\sum_{i=1}^N w_i x_i^q}{\sum_{i=1}^N w_i} \right)^{\frac{1}{q}}, \quad (4.11)$$

gdzie $\sum_{i=1}^N w_i > 0$ i $w_i \geq 0, q \in \mathbb{R}$. Zachodzą następujące wzory dla przypadków granicznych [57]:

$$\begin{cases} \omega_{-\infty}(x_1, x_2, \dots, x_N) = \min\{x_1, x_2, \dots, x_N\}, \\ \omega_0(x_1, x_2, \dots, x_N) = \sqrt[N]{x_1 x_2 \dots x_N}, \\ \omega_{+\infty}(x_1, x_2, \dots, x_N) = \max\{x_1, x_2, \dots, x_N\}. \end{cases} \quad (4.12)$$

W pracy [57] rozpatrzono również sytuację, gdy nie ma wspólnego zbioru cech dla wszystkich węzłów, przytoczmy teraz to rozumowanie. Załóżmy, że dla każdego wierzchołka V_n mamy określony zbiór cech $\Pi_n = \{\pi_{n,1}, \pi_{n,2}, \dots, \pi_{n,M_n}\}$, które sumując po wszystkich wierzchołkach, dadzą nam wspólny zbiór:

$$\Pi = \cup_{i=1}^N \Pi_i = \{\pi_1, \pi_2, \dots, \pi_{|\Pi|}\}. \quad (4.13)$$

Zbiór wierzchołków, dla których $\Pi \neq \Pi_n$ dla każdego $n = 1, 2, \dots, N$ nazywamy *wierzchołkami heterogenicznymi* (ang. *non-similar vertices* w pracy [57]). W pracy [57] wprowadzono zbiór $\Omega^* = \Omega \cup \theta, \theta \notin \Omega$, wraz z elementem θ , gdzie dla θ zachodzi:

$$\begin{cases} (\bigwedge_{x \in \Omega}) & x \not\asymp \theta, \\ (\bigwedge_{x \in \Omega}) & x \not\prec \theta, \\ (\bigwedge_{x \in \Omega}) & x \neq \theta. \end{cases} \quad (4.14)$$

Gdy $\Omega = [0; 1]$, to obiektem spełniającym takie formuły jest np. jednostka urojona i [57]. Celem obiektu θ jest wskazanie, że dla danego węzła, wartość konkretnej cechy nie została określona, bądź nie istnieje, co odpowiada wartości *null* w językach programowania [57]. Za pracą [57]

wprowadźmy oznaczenie:

$$r_{n,j} = \begin{cases} \text{wartość } \pi_j \text{ dla } V_n, & \text{gdy } \pi_j \in \Pi_n, \\ \theta, & \text{w p.p.} \end{cases} \quad (4.15)$$

Nawiązując do wzoru (4.5), możemy zdefiniować *ogólną funkcję łączącą dla heterogenicznych wierzchołków* $F^* : (\Omega^*)^{N|\Pi|} \rightarrow (\Omega^*)^{|\Pi|}$, dla której zachodzi [57]:

$$F^*(\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_N) = \begin{pmatrix} f_1^*(r_{1,1}, r_{2,1}, \dots, r_{N,1}) \\ f_2^*(r_{1,2}, r_{2,2}, \dots, r_{N,2}) \\ \vdots \\ f_{|\Pi|}^*(r_{1,|\Pi|}, r_{2,|\Pi|}, \dots, r_{N,|\Pi|}) \end{pmatrix}, \quad (4.16)$$

gdzie funkcja $f^* : (\Omega^*)^N \rightarrow \Omega^*$ jest *funkcją łączącą dla wierzchołków heterogenicznych*, odpowiadająca funkcji f . Analogicznie do wzorów (4.6, 4.7), w pracy [57] podano następujące założenia dla funkcji f^* :

$$\left(\bigwedge_{x \in \Omega^*} \right) f^*(x, x, \dots, x) = x, \quad (4.17)$$

$$\left(\bigwedge_{n=1}^N \right) \left(\bigwedge_{x_n \leq y_n} \right) f^*(x_1, x_2, \dots, x_n, \dots, x_N) \leq f^*(x_1, x_2, \dots, y_n, \dots, x_N), \quad (4.18)$$

$$f^*(x_1, x_2, \dots, x_N) = \theta \iff x_1 = x_2 = \dots = x_N = \theta. \quad (4.19)$$

Nierówność (4.18) wymaga aby wartości x_k i y_k były porównywalne, natomiast na podstawie (4.14), element θ nie spełnia tego warunku z żadnym innym elementem ze zbioru Ω^* [57].

Praca [57] podaje prosty algorytm obliczania wartości funkcji $f^*(x_1, x_2, \dots, x_n)$:

1. Niech $Z = (z_1, z_2, \dots, z_{N'})$ będzie ciągiem elementów ze zbioru $X = (x_1, x_2, \dots, x_N)$ utworzonego poprzez wybranie wszystkich elementów różnych od θ .
2. Użycie np. wzoru (4.11) do obliczeń, korzystając tylko z wartości ze zbioru Z . Wagi powinny zostać zachowane.

Zastanówmy się teraz nad tym, jak wybrać *funkcję łączącą*. Należy wziąć pod uwagę następujące aspekty, zaproponowane w pracy [57].

- Interpretacja cechy pod kątem sieciowym i izolacji - np. dostępna przepływność binarna w szeregowym połączeniu węzłów jest ograniczona z góry przez najwolniejszy element, więc funkcja minimum może być uzasadnionym wyborem.
- Ograniczenia implementacyjne - obliczenia na liczbach całkowitych są zwykle szybsze i dokładniejsze niż na liczbach zmiennoprzecinkowych, szczególnie tych o podwójnej precyzji (IEEE 754[40]). Dokładne typy danych zmiennoprzecinkowych (tzw. typy *money* albo *decimal*) bywają jeszcze wolniejsze, ponieważ nie zawsze mają wsparcie sprzętowe po stronie procesora. Można ten problem ograniczyć, stosując odpowiednie skalowanie, które spowoduje możliwość operowania na liczbach całkowitych.
- Ograniczenia precyzji - obliczenia zmiennoprzecinkowe są obarczone błędami wynikającymi ze skończonej precyzji zapisu danych, które nawarstwiają się przy każdej kolejnej operacji matematycznej, zmniejszając precyzję końcowego wyniku.
- Wartość oczekiwana *funkcji łączącej* - *funkcja normalizująca* ma zwykle zdefiniowany *punkt centralny* C , taki, że $C = g(\beta)$, gdzie $g(x)$ to *funkcja normalizująca*, a β to pewna *typowa wartość* dla danej cechy. Wartość parametru C powinna być określona z uwzględnieniem charakteru zbioru Ω , w szczególności dla zbioru Ω będącego przedziałem liczb rzeczywistych, powinien to być środek tego przedziału. *Funkcja łącząca* może mieć wartość oczekiwaną różną od C , co w przypadku znaczących różnic (w stosunku np. do długości przedziału, który definiuje Ω), może dawać pewne wnioski interpretacyjne, szczególnie bez pełnej znajomości kontekstu obliczeń. Dla rodziny funkcji potęgowych, tylko średnia arytmetyczna ($q = 1$) pozwala na zachowanie zależności:

$$\mathbb{E}f_q(x_1, x_2, \dots, x_N) = g(\beta), \quad (4.20)$$

dla każdego N .

- Kształt *funkcji łączącej* - cechy mogą mieć pewną indywidualną logikę interpretacyjną, w szczególności jak powinna się zachowywać *funkcja łącząca* dwie lub więcej wartości danej cechy. W pracy [57] wyszczególniono następujące podejścia:
 - logika OR - wartość *funkcji łączącej* powinna silniej zależeć od największych podanych wartości parametrów, niż od mniejszych. Takie cechy powinno się łączyć za pomocą

funkcji wypukłych. Przykładem takiej cechy jest *znana liczba podatności CVE* (w kontekście danego węzła sieci).

- logika AND - wartość *funkcji łączącej* powinna silniej zależeć od najmniejszych podanych wartości parametrów, niż od większych. Takie cechy powinny się łączyć za pomocą funkcji wklęsłych. Przykładem takiej cechy jest *ustandaryzowana moc szyfrowania* (w bitach klucza symetrycznego odpowiadającego bezpieczeństwu użytego algorytmu).
- logika neutralna - wartość *funkcji łączącej* powinna w równym stopniu zależeć od najmniejszych podanych wartości parametrów, jak i od największych. Takie cechy można łączyć korzystając z *funkcji łączącej* spełniającej równość:

$$\frac{\partial^2}{\partial x_n^2} f(x_1, x_2, \dots, x_n, \dots, x_N) = 0, \quad (4.21)$$

dla każdego $n = 1, 2, \dots, N$.

4.4. ZAGADNIENIE PORZĄDKU W KONTEKŚCIE IZOLACJI

Z praktycznego punktu widzenia Operator Telekomunikacyjny chciałby mieć możliwość porównania izolacji wyznaczonej dla pewnych różnych fragmentów sieci, bądź też porównywać aktualny stan izolacji w sieci z założonymi poziomami, zdefiniowanymi w kontraktach z klientami. Korzystając z funkcji łączących, można uzyskać *wektor izolacji* [57], który reprezentuje punkt w przestrzeni $(\Omega^*)^{|\Omega|}$. Porównywanie jest tutaj nietrywialne, gdy dwa wektory różnią się na więcej niż jednej współrzędnej (mają więcej niż jedną różną wartość cechy), lub gdy dla danej cechy dokładnie jeden z wektorów przyjmuje wartość θ .

Spróbujmy teraz zdefiniować porządek liniowy dla wektorów izolacji, poprzez zaproponowanie dwukrotnie różniczkowalnej funkcji $\Phi : (\Omega)^{|\Omega|} \rightarrow \Omega$, która przekształci wektor izolacji w skalar. Problem ten został postawiony w [57] i zaproponowano następujące rozwiązanie: z założeń poczynionych podczas definiowania zbioru Ω wiemy, że izolacja rośnie wraz ze wzrostem wartości cech, więc funkcja Φ powinna spełniać nierówność [57]:

$$\frac{\partial}{\partial x_n} \Phi(x_1, x_2, \dots, x_n, \dots, x_N) \geq 0, \quad (4.22)$$

dla każdego $n = 1, 2, \dots, N$. Oprócz tego, chcielibyśmy, aby niskie wartości cech miały większy wpływ na wynik - co można osiągnąć, stosując funkcję wklęsłą, której druga pochodna cząstkowa jest niedodatnia [57]:

$$\frac{\partial^2}{\partial x_n^2} \Phi(x_1, x_2, \dots, x_n, \dots, x_N) \leq 0. \quad (4.23)$$

Przykładową funkcją, która spełnia te dwie nierówności jest funkcja [57]:

$$\Phi_q(x_1, x_2, \dots, x_n, \dots, x_N) = 1 - \left(\frac{1}{N} \sum_{i=1}^N (1 - x_i)^q \right)^{\frac{1}{q}}. \quad (4.24)$$

Jej pierwsza i druga pochodna jest równa odpowiednio [57]:

$$\frac{\partial}{\partial x_n} \Phi_q(x_1, x_2, \dots, x_n, \dots, x_N) = \left(\sum_{i=1}^N (1 - x_i)^q \right)^{\frac{1}{q}-1} \frac{(1 - x_n)^{q-1}}{N} \geq 0, \quad (4.25)$$

$$\frac{\partial^2}{\partial x_n^2} \Phi_q(x_1, x_2, \dots, x_n, \dots, x_N) = \frac{1 - q}{n^2} \left(\sum_{i=1}^N (1 - x_i)^q \right)^{\frac{1}{q}-2} (1 - x_n)^{q-2} \sum_{i=1; i \neq n}^N (1 - x_i)^q. \quad (4.26)$$

Gdy $q \geq 1$, to druga pochodna jest niedodatnia, więc spełnia założenia poczynione we wzorach (4.22) i (4.23) [57].

4.5. KROKI ALGORYTMU

Przejdźmy teraz do opisu algorytmu obliczania izolacji w sieci. Został on zaproponowany w pracy [57], był również rozwinięty w [56].

Zaproponowany algorytm nie podejmuje kwestii wyznaczenia zbioru wierzchołków podstawowych, zakłada, że zostały one zidentyfikowane na podstawie dostępnej dokumentacji systemu informacyjnego, wiedzy eksperckiej dostępnej w ramach organizacji i dokumentacji procesów informacyjnych, które są realizowane w przestrzeni objętej tymi wierzchołkami. Podobna sytuacja jest z cechami przynależnymi do wierzchołków - pewna wiedza kontekstowa dotycząca danej cechy i o wykonywanych procesach informacyjnych jest kluczowa, aby prawidłowo dokonać wstępnej oceny przydatności danej cechy do analizy izolacji.

Algorytm 1 Algorytm wyznaczania izolacji z wykorzystaniem liczb rzeczywistych

- 1: **procedure** CALCULATEISOLATIONREALNUMBERS
 - 2: Wyznacz zbiór cech właściwych dla poszczególnych wierzchołków w badanym hipergrafie sieci.
 - 3: Ustal funkcje normalizujące dla cech przynależnych do poszczególnych wierzchołków.
 - 4: Ustal funkcję łączącą izolację dla poszczególnych cech.
 - 5: Zbierz wartości cech w wierzchołkach.
 - 6: Znormalizuj wartości cech, korzystając z funkcji normalizujących ustalonych w kroku 2.
 - 7: Zagreguj znormalizowane wartości cech z kroku 5, za pomocą funkcji ustalonych w kroku 3.
 - 8: Jeżeli była wydzielona więcej niż jedna podsieć lub poziom w modelu hipergrafowym, agreguj wyniki aż do uzyskania jednego wektora izolacji.
-

Algorytm pozwala na przenikanie się wzajemne podsieci, czyli na współdzielenie węzłów pomiędzy dwoma podsieciami. Przykładem takiego węzła jest brama sieciowa (*gateway*), która łączy dwie podsieci, nierzadko o różnych stosach technologicznych (np. sieci przewodowe i bezprzewodowe). Matematyczną konsekwencją takiego podejścia jest większy udział (wpływ) cech u wspólnionego węzła niż innych węzłów (o pojedynczym użyciu).

5. Skierowane liczby rozmyte

Niniejszy rozdział stanowi krótkie wprowadzenie do tematyki skierowanych liczb rozmytych (SLR). Korzystamy tutaj z aparatu pojęciowego i matematycznego sformułowanego m.in. w [18, 53, 65, 68, 121].

5.1. OGÓLNE INFORMACJE O SKIEROWANYCH LICZBACH ROZMYTYCH

Skierowane Liczby rozmyte zostały zaproponowane w pracy W. Kosińskiego, P. Prokopowicza i D. Ślęzaka [52], jako rozwinięcie idei (L,R)-liczb zaproponowanych przez Dubois i Prade w pracy [21]. Idea skierowanych liczb rozmytych była później przedmiotem innych publikacji naukowych, na przykład [18, 50, 51, 53, 89, 99]. Skierowane liczby rozmyte były również motywem przewodnim rozprawy doktorskiej [68], oraz podstawą do zbudowania nowej struktury matematycznej - świec rozmytych [68, 69], mających zastosowanie w badaniu przebiegów czasowych występujących w modelach ekonometrycznych.

5.2. POJĘCIA I DEFINICJE MATEMATYCZNE

Podstawową koncepcją w wielu pracach o metodach rozmytych jest pojęcie *zbioru rozmytego* (*fuzzy set*). W tej pracy skorzystaliśmy z definicji podanej w [65, 68, 123].

Definicja 5.1 (Zbiór rozmyty)

Zbiór rozmyty A nad przestrzenią X jest zbiorem par:

$$A = \{(x, \mu_A(x)) : x \in X\}, \quad (5.1)$$

gdzie $\mu_A(x) : X \rightarrow [0, 1]$ jest funkcją przynależności (membership/characteristic function) do A .

Definicja 5.2 (Liczba rozmyta)

Jeśli zbiór rozmyty A spełnia warunki:

1. A jest unormowany - $\exists_{x_0 \in X} \mu(x_0) = 1$;
2. A jest wypukły [123] - dla każdego $x, y \in X$ oraz dla każdego $\lambda \in [0, 1]$:

$$\mu_A(\lambda x + (1 - \lambda)y) \geq \min(\mu_A(x), \mu_A(y)); \quad (5.2)$$

3. $\text{Supp}(A) = \{x \in X : \mu_A(x) > 0\}$ (tzw. nośnik A [65, 68]) jest przedziałem;
4. $\mu(x)$ jest przedziałami funkcją ciągłą,

to A jest nazywana liczbą rozmytą (fuzzy number) [21].

Dla liczb rozmytych według powyższej definicji są zdefiniowane operacje dwuargumentowe jak dodawanie czy mnożenie takich liczb, jednakże w praktyce korzystanie z nich obarczone jest pewnymi trudnościami [51]:

- zbiór wypukłych liczb rozmytych nie tworzy pierścienia,
- każda operacja (dodawanie, odejmowanie, itp.) powoduje zwiększenie nośnika liczby rozmytej - czyli zwiększa się rozmycie. Przykładowo, dodawanie liczby i jej liczby przeciwnej da w wyniku zero rozmyte, zamiast zera rzeczywistego [53] - co jest samo w sobie kłopotliwe. Innym problemem jest brak elementu odwrotnego dla mnożenia [68].

Wprowadźmy teraz pojęcie skierowanej liczby rozmytej (*ordered fuzzy number*, SLR), która rozwiązuje kilka istotnych problemów praktycznych [18, 51].

Definicja 5.3 (Skierowana liczba rozmyta)

Skierowaną liczbą rozmytą S nazywamy uporządkowaną parę funkcji [53, 65, 68]:

$$S = (f_S, g_S), f_S, g_S : [0, 1] \rightarrow \mathbb{R}. \quad (5.3)$$

Skierowana liczba rozmyta ma orientację, nazywane również skierowaniem, (*orientation*).

W pracy [68] podano następującą definicję orientacji (skierowania) dla właściwych skierowanych liczb rozmytych.

Definicja 5.4 (Orientacja (skierowanie) właściwej skierowanej liczby rozmytej)

Skierowana liczba rozmyta ma orientację pozytywną (dodatnią) wtedy, gdy kierunek skierowanej liczby rozmytej jest zgodny z kierunkiem osi OX liczb rzeczywistych. Orientacja danej liczby rozmytej A ilustruje tylko położenie części $up f_S$ względem części $down g_S$.

Powyższa definicja (wykorzystująca kierunek osi OX) jest zgodna z opisem orientacji przedstawionym w [50]. Naturalnie, gdy kierunek skierowanej liczby rozmytej jest przeciwny do osi OX, to ma ona orientację negatywną (ujemną) [68]. W pracy [68] zdefiniowano orientację również za pomocą funkcji:

$$Ord(S) = \begin{cases} -1, & \text{gdy } \Theta(S) < f_S(0), \\ 0, & \text{gdy } \Theta(S) = f_S(0) \text{ albo } \Theta(S) \text{ nie istnieje,} \\ 1, & \text{gdy } \Theta(S) > f_S(0). \end{cases} \quad (5.4)$$

Funkcja $\Theta(x)$ jest operatorem wyostrzania skierowanej liczby rozmytej (*defuzzification functional*) [18, 68, 69].

Definicja 5.5 (Operator wyostrzania)

Założmy, że $\Theta(x)$ jest taką funkcją, że $\Theta : \mathcal{F} \rightarrow \mathbb{R}$, gdzie \mathcal{F} to zbiór wszystkich skierowanych liczb rozmytych. Wówczas taką funkcję nazywamy operatorem wyostrzania.

Operator wyostrzania może być wrażliwy na orientację lub nie [18]. Wzór (5.4) w praktyce wymaga znajomości orientacji lub rozróżnia, która część funkcji przynależności odpowiada funkcji f_S , a która funkcji g_S . Jednym ze znanych operatorów wyostrzania jest funkcja środka ciężkości (*the center of gravity*) [18] zadana wzorem:

$$\phi_{COG}(f, g) = \begin{cases} \frac{\int_0^1 (f(s)+g(s))|f(s)-g(s)|ds}{2 \int_0^1 |f(s)-g(s)|ds}, & \text{gdy } \int_0^1 |f(s) - g(s)|ds \neq 0, \\ \int_0^1 f(s)ds, & \text{gdy } \int_0^1 |f(s) - g(s)|ds = 0. \end{cases} \quad (5.5)$$

Gdy obydwie funkcje f_S oraz g_S są monotoniczne, to funkcja przynależności dla skierowanej liczby rozmytej zorientowanej (skierowanej) dodatnio (lub pozytywnie) jest zadana

następującą formułą [50, 68, 69]:

$$\mu_S(x) = \begin{cases} f_S^{-1}(x), & \text{gdy } x \in (L_S, 1_S^-), \\ 1, & \text{gdy } x \in [1_S^-, 1_S^+], \\ g_S^{-1}(x), & \text{gdy } x \in (1_S^+, R_S), \\ 0, & \text{w p.p.} \end{cases} \quad (5.6)$$

Wartość L_S jest lewym końcem przedziału, w którym $f_S(x)$ jest zdefiniowana. Wartość R_S jest prawym końcem przedziału, w którym $g_S(x)$ jest zdefiniowana. Wartości 1_S^- oraz 1_S^+ spełniają następującą zależność:

$$\left(\bigwedge_{x \in [1_S^-, 1_S^+]} \right) \mu_S(x) = 1. \quad (5.7)$$

Dla skierowanych liczb rozmytych zorientowanych (skierowanych) negatywnie (lub ujemnie), mamy następującą formułę opisującą funkcję przynależności:

$$\mu_S(x) = \begin{cases} g_S^{-1}(x), & \text{gdy } x \in (L_S, 1_S^-), \\ 1, & \text{gdy } x \in [1_S^-, 1_S^+], \\ f_S^{-1}(x), & \text{gdy } x \in (1_S^+, R_S), \\ 0, & \text{w p.p.} \end{cases} \quad (5.8)$$

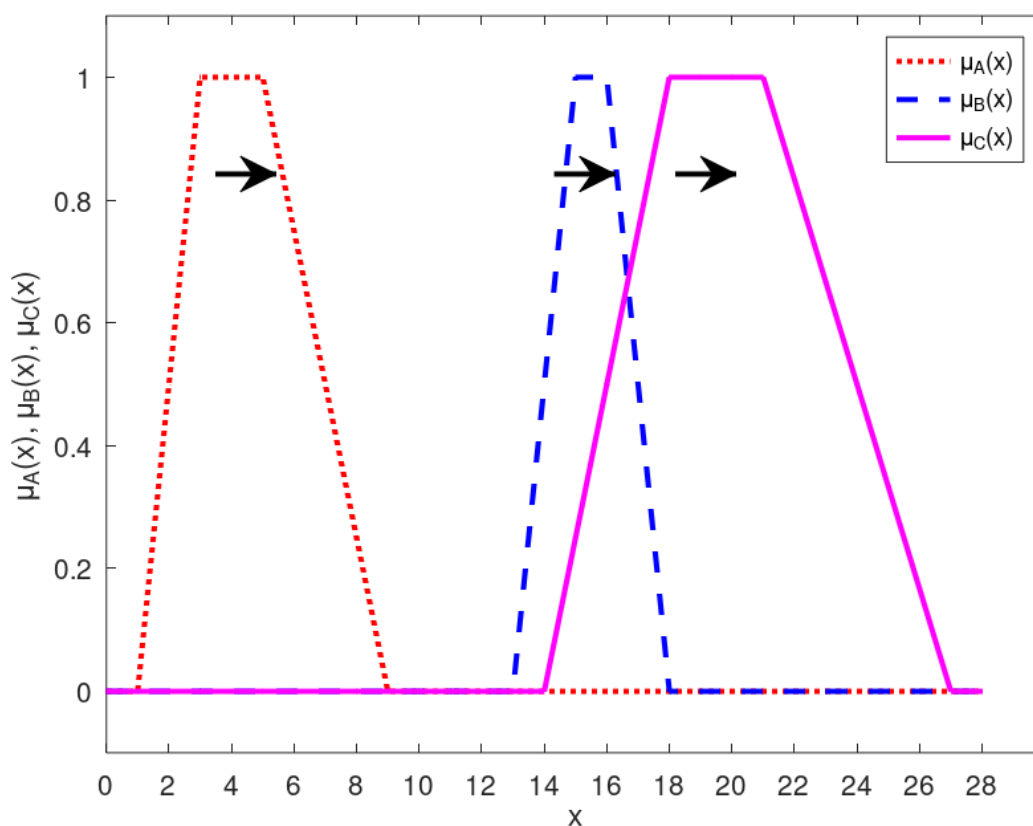
Definicja 5.6 (Operacje dwuargumentowe dla liczb rozmytych)

Dla skierowanych liczb rozmytych, operacje dwuargumentowe są zdefiniowane następująco:

$$C = \diamond(A, B) = A \diamond B = (f_A \diamond f_B, g_A \diamond g_B). \quad (5.9)$$

Definicja (5.6) jest możliwa wprost do zastosowania dla operacji dodawania, odejmowania, mnożenia, dzielenia, [50, 99], podnoszenia do potęgi, logarytmowania [99]. Dzielenie jest dopuszczalne, gdy $|f_B(x)| > 0$ oraz $|g_B(x)| > 0$. Podnoszenie do potęgi i logarytmowanie jest dopuszczalne, gdy te operacje są dobrze zdefiniowane dla liczb rzeczywistych, które występują w odpowiednich funkcjach przynależności. Rysunek 5.1 pokazuje przykład dodawania SLR z

Dodawanie SLR o tym samym skierowaniu: $A + B = C$



Rysunek 5.1: Przykład dodawania dwóch różnych SLR o identycznych, dodatnich skierowaniach.

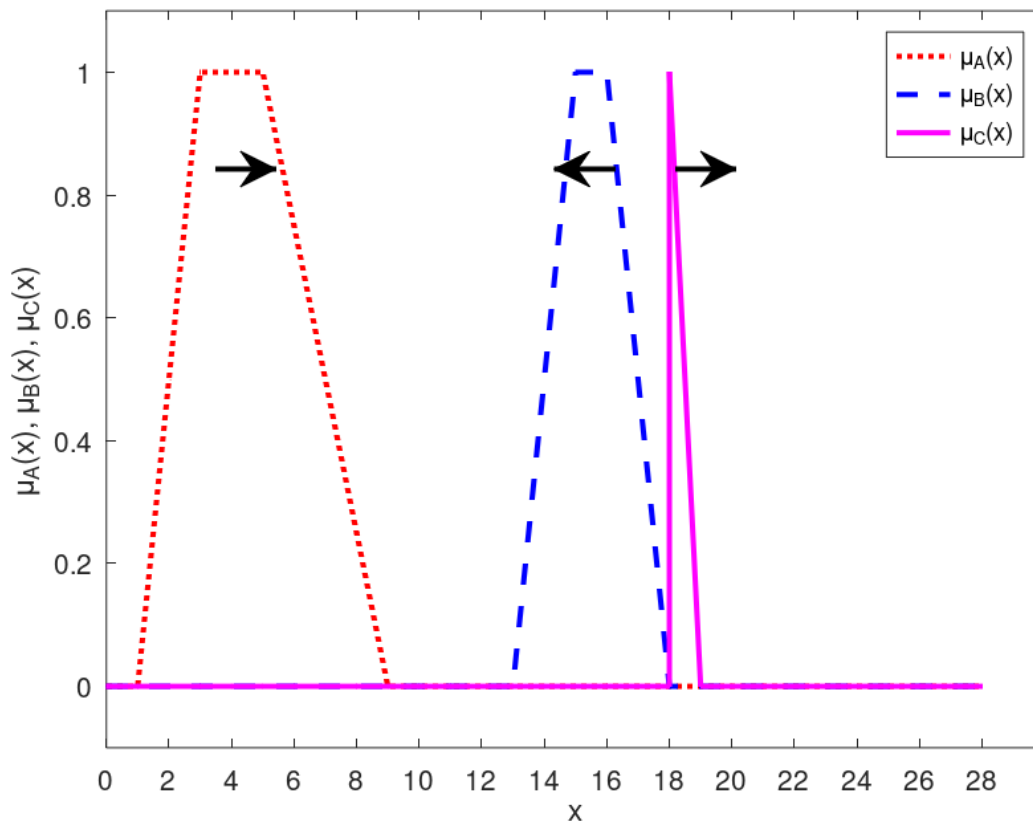
dodatnim skierowaniem. Rysunek 5.2 przedstawia przykład dodawania dwóch SLR o przeciwnym skierowaniu. W rezultacie wykonywanych operacji, SLR może nie mieć zdefiniowanej orientacji, aczkolwiek dalej można rozróżnić funkcję f_S i g_S . Do wyznaczenia orientacji można wykorzystać formułę (5.4) wraz z pewnym operatorem wyostrzania, np. (5.5).

Niech $X = (f, g)$ będzie SLR nad ciałem liczb rzeczywistych. Wprowadźmy *metrykę rozmycia skierowanej liczby rozmytej* $D(X)$, nazywaną dalej w skrócie *metryką rozmycia*, którą zdefiniujemy za pomocą wzoru:

$$Q(X, r) = \int_0^1 (f(y) - r)^2 dy + \int_0^1 (g(y) - r)^2 dy, \quad (5.10)$$

$$D(X) = \min_{r \in \mathbb{R}} Q(X, r). \quad (5.11)$$

Dodawanie SLR o przeciwnych skierowaniach: A + B = C



Rysunek 5.2: Przykład dodawania dwóch różnych SLR o przeciwnych skierowaniach.

Spróbujmy znaleźć ekstrema funkcji $Q(X, r)$ ze względu na r :

$$\frac{\partial Q(X, r)}{\partial r} = -2 \int_0^1 (f(y) + g(y)) dy + 4r. \quad (5.12)$$

Funkcja $Q(X, r)$ osiąga wartość najmniejszą dla danego X , gdy:

$$r = \int_0^1 \frac{f(y) + g(y)}{2} dy = r_0. \quad (5.13)$$

Mamy więc:

$$D(X) = \int_0^1 f^2(y) + g^2(y) dy - \frac{\left(\int_0^1 f(y) + g(y) dy \right)^2}{2}. \quad (5.14)$$

W przypadku gdy $X \in \mathbb{R}$, mamy $f(y) = g(y) = c = \text{const}$ i wówczas:

$$D(X) = \int_0^1 c^2 + c^2 dy - \frac{\left(\int_0^1 c + c dy\right)^2}{2} = 0. \quad (5.15)$$

Rozpatrzmy teraz sytuację liczby rozmytej, która ma oś symetrii $x = c$. Wówczas $g(y) = 2c - f(y)$ oraz $r_0 = c$. Mamy wtedy:

$$D(X) = \int_0^1 f^2(y) + (2c - f(y))^2 dy - 2c^2 = \int_0^1 2f^2(y) - 4cf(y) dy + 2c^2, \quad (5.16)$$

$$D(X) = 2 \int_0^1 (f(y) - c)^2 dy. \quad (5.17)$$

Niech $Y = X + a = (f(y) + a, g(y) + a)$. Wówczas:

$$D(Y) = \int_0^1 (f(y) + a)^2 + (g(y) + a)^2 dy - \frac{\left(\int_0^1 f(y) + g(y) + 2ady\right)^2}{2} = D(X). \quad (5.18)$$

Niech $Y = aX = (af(y), ag(y))$. Wówczas:

$$D(Y) = \int_0^1 (af(y))^2 + (ag(y))^2 dy - \frac{\left(\int_0^1 af(y) + ag(x) dy\right)^2}{2} = a^2 D(X). \quad (5.19)$$

Dwie różne SLR mogą mieć taką samą metrykę rozmycia, mimo zupełnie odmiennego kształtu; w szczególności wielkość nośnika czy kształt funkcji przynależności takich liczb może się bardzo między sobą różnić, przy zachowaniu zadanej metryki rozmycia. W celu porównywania tego typu przypadków, zdefiniujemy *unormowaną metrykę rozmycia* $\|D(X)\|$, która będzie się wyrażać wzorem:

$$\|D(X)\| = \frac{2D(X)}{(m - M)^2}, \quad (5.20)$$

gdzie $m = \min\{\inf f(y), \inf g(y)\}$ oraz $M = \max\{\sup f(y), \sup g(y)\}$. Wielkość $\frac{(m-M)^2}{2}$ jest wartością metryki rozmycia dla liczby rozmytej X_0 , która ma kształt prostokąta opisanego na krzywej funkcji przynależności liczby rozmytej X .

6. Model rozmyty

W niniejszym rozdziale omówimy algorytm wyznaczania izolacji za pomocą skierowanych liczb rozmytych (SLR).

6.1. KONSTRUOWANIE SLR NA PODSTAWIE WYNIKÓW POMIARÓW

W podrozdziale 4.2. wprowadziliśmy podział cech na właściwości i parametry. Właściwości, ze względu na swój enumeratywny charakter, w zasadzie nie cierpią na błędy pomiaru lub zaokrągleń, w przeciwieństwie do parametrów. Dodatkowo wartości cech mogą ulegać pewnym trendom, których również pojedynczy pomiar nie oddaje [89]. Użycie SLR pozwala na uwzględnienie tych zjawisk. W dalszej części tego podrozdziału będziemy przedstawiać algorytm budowania SLR na podstawie wyników pomiarów, który został zaproponowany w pracy [89].

Przejdźmy do zagadnienia stworzenia SLR na podstawie wyników pomiarów. Załóżmy, że mamy serię pomiarów (v_1, v_2, \dots, v_n) przeprowadzonych w chwilach odpowiednio (t_1, t_2, \dots, t_n) . W związku z tym, że dla SLR są zdefiniowane te same operacje algebraiczne co dla liczb rzeczywistych, to aby wykonać algorytm zaproponowany w podrozdziale 4.5., możemy wykonać jedną z dwóch czynności:

1. znormalizować wartości pomiarowe do zbioru Ω i z takich wartości utworzyć SLR,
2. stworzyć SLR na podstawie wartości pomiarowych i taką SLR następnie znormalizować.

Skorzystamy z pierwszego podejścia, ponieważ upraszcza ono zaproponowane rozwiązanie; zostało ono również zaproponowane w pracy [89]. Podejście to zakłada, że najpierw tworzymy dwie funkcje f i g składające się na SLR, a następnie określamy orientację tworzonej liczby. W pracy [89] zaproponowano, aby orientacja była determinowana przez tangens kąta nachylenia linii trendu względem osi czasu. Gdy pomiary są jednakowo ważne (mają równe wagi), to taką

orientację dla serii V można wyznaczyć z poniższych wzorów (na podstawie [29, 89]):

$$\begin{aligned} Ord(V) &= sgn(\text{tg } \varphi_V) = sgn\left(\frac{\sum_{k=1}^n (t_k - \bar{t})(v_k - \bar{v})}{\sum_{k=1}^n (t_k - \bar{t})^2}\right) = \\ &= sgn\left(\frac{n \sum_{k=1}^n t_k v_k - (\sum_{k=1}^n t_k)(\sum_{k=1}^n v_k)}{n \sum_{k=1}^n t_k^2 - (\sum_{k=1}^n t_k)^2}\right), \end{aligned} \quad (6.1)$$

gdzie $\bar{v} = \sum_{k=1}^n v_k$ i $\bar{t} = \sum_{k=1}^n t_k$. Gdy $t_k = (k-1)\tau$, to [89]:

$$\begin{aligned} Ord(V) &= sgn\left(\frac{12 \sum_{k=1}^n \left(k - \frac{n+1}{2}\right) v_k}{\tau n(n-1)^2}\right) = \\ &= sgn\left(\sum_{k=1}^n \left(k - \frac{n+1}{2}\right) v_k\right) = sgn\left(\sum_{k=1}^n \left(k - \frac{n+1}{2}\right) (v_k - \bar{v})\right). \end{aligned} \quad (6.2)$$

Może się zdarzyć, że orientacja obliczona z powyższego wzoru jest nieokreślona ($Ord(V) = 0$). Ze względu na dość dużą precyzję zapisu pomiarów, może to być zjawisko stosunkowo rzadkie w przypadku rzeczywistych wyników pomiarowych. Ten problem można rozwiązać poprzez:

- wybranie losowej orientacji,
- wyznaczenie orientacji na podstawie dłuższej lub krótszej serii pomiarowej,
- wykorzystanie innej funkcji do wyznaczenia trendu. Dla dwóch sygnałów $x[n]$ i $y[n]$ o czasie dyskretnym funkcja korelacji wzajemnej jest zdefiniowana jako [94]:

$$\phi_{xy}[k] = \sum_{m=-\infty}^{\infty} x[m+k]y[m]. \quad (6.3)$$

Zauważmy, że wzór (6.2) oblicza znak korelacji wzajemnej w punkcie $k = 0$ pomiędzy badanym sygnałem (serią pomiarową) a funkcją liniową [89]. W pracy [89] pokazano, że dla niektórych kategorii sygnałów (o znanym kształcie), użycie innej funkcji do wyznaczenia skierowania może dawać lepsze rezultaty. Rozumiemy przez to określenie poprawnego skierowania dla (pseudo)losowych sygnałów o znanym trendzie.

Jeśli znormalizowana funkcja przynależności może być podzielona na dwie ciągłe części: niemalejącą i nierosnącą, to można wyznaczyć funkcje f i g z zależności (5.6). W pracy [89] zaproponowano następujący algorytm:

1. Stwórz znormalizowany histogram z wartości (v_1, v_2, \dots, v_n) na przedziale $[\alpha; \omega]$. Przez znormalizowany histogram rozumiemy histogram, w którym najwyższy słupek ma wartość

1. Niech obwiednia tego histogramu będzie oznaczona jako $\tilde{\mu}(x)$ oraz zdefiniujmy:

$$1_V^- = \min\{x \in \mathbb{R} : \tilde{\mu}(x) = 1\}, \quad (6.4)$$

$$1_V^+ = \max\{x \in \mathbb{R} : \tilde{\mu}(x) = 1\}. \quad (6.5)$$

2. Gdy otrzymana orientacja jest dodatnia, to:

$$f(x) = \begin{cases} \max\{\tilde{\mu}(x), \max_{t < x} f(t)\}, & x \in [\alpha; 1_V^-), \\ 0, & x \in (\alpha - \epsilon; \alpha), \end{cases} \quad (6.6)$$

$$g(x) = \begin{cases} \max\{\tilde{\mu}(x), \max_{t > x} g(t)\}, & x \in (1_V^+; \omega], \\ 0, & x \in (\omega; \omega + \epsilon). \end{cases} \quad (6.7)$$

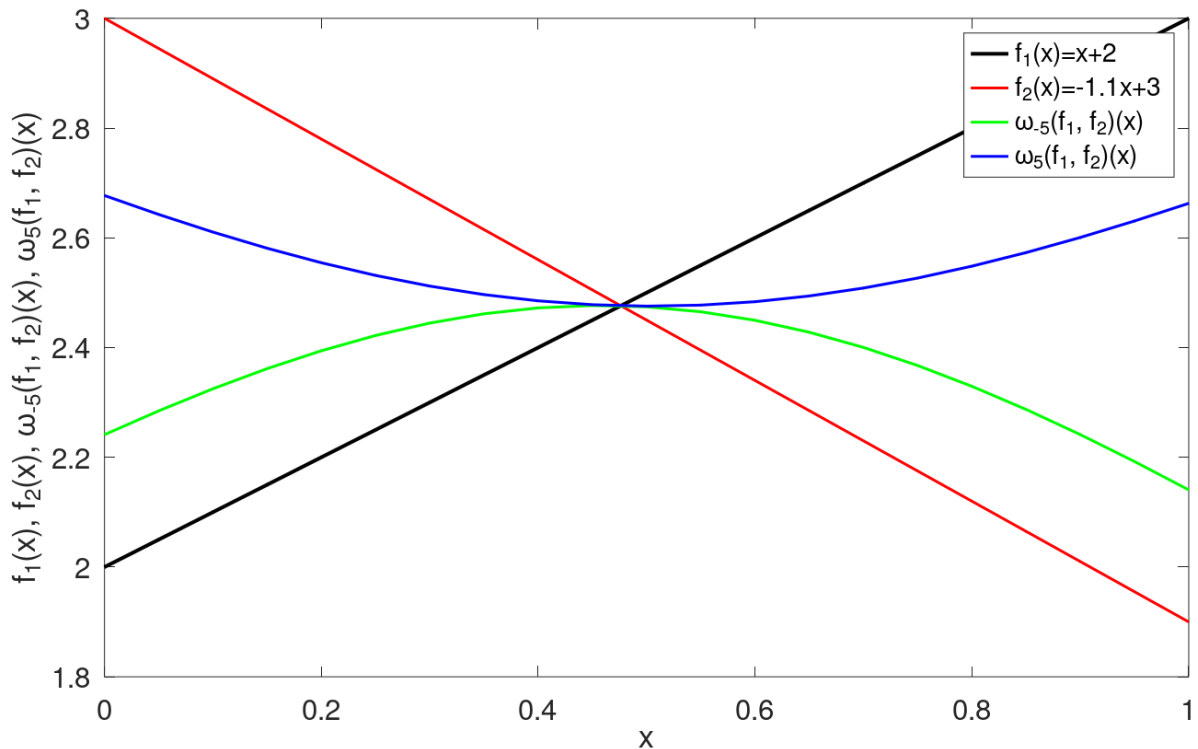
Jeśli orientacja jest ujemna, to:

$$f(x) = \begin{cases} \max\{\tilde{\mu}(x), \max_{t > x} f(t)\}, & x \in (1_V^+; \omega], \\ 0, & x \in (\omega; \omega + \epsilon), \end{cases} \quad (6.8)$$

$$g(x) = \begin{cases} \max\{\tilde{\mu}(x), \max_{t < x} g(t)\}, & x \in [\alpha; 1_V^-), \\ 0, & x \in (\alpha - \epsilon; \alpha). \end{cases} \quad (6.9)$$

W powyższych formułach $\epsilon > 0$ jest pewną stałą, której celem jest wyznaczenie wartości funkcji f i g na krańcach przedziału [89]. Esencją powyższego algorytmu jest skonstruowanie histogramu, którego kolumny najpierw monotonicznie niemaleją, następnie następuje seria kolumn o maksymalnej wysokości, a na końcu monotonicznie nierosną.

Skierowane liczby rozmyte skonstruowane według powyższego algorytmu można następnie złączyć z wykorzystaniem funkcji łączącej, np. średniej potęgowej określonej wzorem (4.11). Co bardzo istotne, utworzone liczby są liczbami *właściwymi*. Zagadnienie liczb *niewłaściwych* jest przedmiotem rozważań następnego podrozdziału.



Rysunek 6.1: Przykład złączenia dwóch funkcji liniowych, który w rezultacie tworzy funkcję nieodwracalną.

6.2. PROBLEM NIEWŁAŚCIWYCH SLR

Korzystając z formuły (4.11) można uzyskać SLR składającą się z funkcji, które nie są bijekcjami (są nieodwracalne); będziemy nazywać takie SLR *niewłaściwymi SLR*. Podany problem w szczególności występuje, gdy łączone są SLR o różnych orientacjach. Rysunek 6.1 przedstawia taką sytuację dla dwóch liniowych funkcji f_1 i f_2 , które zostały następnie złączone. W takiej sytuacji nie ma jednoznacznej metody na wyznaczenie funkcji przynależności. Praca [89] zawiera algorytm heurystyczny (opisany w poprzednim podrozdziale), który pozwala wyodrębnić funkcje f i g , gdy zadana jest funkcja przynależności $\mu(x)$. Skupmy się teraz na rozwiązaniu pokrewnego problemu: dla zadanych f i g , spróbujmy znaleźć funkcję przynależności, która będzie opisywać *właściwą* liczbą rozmytą.

6.2.1. Przybliżenie za pomocą właściwych SLR

Każda SLR składa się z uporządkowanej pary dwóch funkcji, które w wyniku wykonywania działań na SLR mogą być nieodwracalne. Naszym celem jest znalezienie takiej SLR, która

jest *właściwa* i bliska (najbliższa) oryginalnej SLR. Zdefiniujmy wobec tego metrykę, którą będziemy używać do określania odległości między SLR:

$$d(A, B) = \|f_A(x) - f_B(x)\| + \|g_A(x) - g_B(x)\|. \quad (6.10)$$

W praktyce, skierowana liczba rozmyta A jest przedstawiana jako dwa ciągi punktów z płaszczyzny \mathbb{R}^2 , które definiują jej funkcje $f_A(x)$ i $g_A(x)$. Oznaczmy ciąg odpowiadający funkcji $f_A(x)$ jako:

$$F_A = ((a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)), a_1 < a_2 < \dots < a_n, \quad (6.11)$$

Bez straty ogólności, możemy założyć, że $(a_k, b_k) \in \mathbb{R}_{\geq 0}^2$ dla każdego n , ponieważ po zabiegu normalizacji mamy $a_k \in [0; 1]$ oraz $b_k \in \Omega = [0; 1]$. Szukamy ciągu $F_B = ((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)), x_1 < x_2 < \dots < x_n$, który minimalizuje odległość między tymi dwoma ciągami. Możemy założyć, tak jak dla oryginalnego ciągu, że $(x_k, y_k) \in \mathbb{R}_{\geq 0}^2$ dla każdego k . Ciągi F_A i F_B można rozpatrywać jako ciągi liczb zespolonych, lub jako elementy z przestrzeni unitarnej $K = (\mathbb{C}^n, \cdot, +, \langle \cdot, \cdot \rangle)$ ze zdefiniowanym iloczynem skalarnym:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^n u_k \overline{v_k}, \quad (6.12)$$

gdzie $\mathbf{u} = [u_1, u_2, \dots, u_n]^T, u_k = a_k + ib_k$ oraz $\mathbf{v} = [v_1, v_2, \dots, v_n]^T, v_k = x_k + iy_k$. Jako metrykę odległości możemy użyć funkcji Mahalanobisa:

$$D_M(\mathbf{u}, \mathbf{v}) = \sqrt{(\mathbf{u} - \mathbf{v})^T \mathbf{S}^{-1} (\mathbf{u} - \mathbf{v})}. \quad (6.13)$$

Oryginalnie, funkcja ta została zdefiniowana dla wektorów liczb rzeczywistych. W tym przypadku mamy wektor par liczb rzeczywistych, który może być traktowany jak wektor liczb zespolonych, wymagane jest również użycie sprzężenia hermitowskiego zamiast zwykłej transpozycji. Jeśli \mathbf{S} jest macierzą jednostkową, to:

$$D_M(\mathbf{u}, \mathbf{v}) = \sqrt{(\mathbf{u} - \mathbf{v})^T (\mathbf{u} - \mathbf{v})}, \quad (6.14)$$

$$D_M(F_A, F_B) = \sqrt{\sum_{i=k}^n ((a_k - x_k)^2 + (b_k - y_k)^2)}. \quad (6.15)$$

Rozwiązanie minimalizujące odległość musi spełniać równość $a_k = x_k$ dla każdego k , ponieważ ciągi (a_k) i (x_k) są rosnące. Skoro funkcja pierwiastkująca jest ściśle rosnąca, to możemy minimalizować funkcję podpierwiastkową:

$$Z_1(F_A, F_B) = D_M^2(F_A, F_B) = \sum_{k=1}^n (y_k - b_k)^2. \quad (6.16)$$

W każdej SLR jedna z funkcji f i g jest funkcją niemalejącą, a druga nierosnącą. Dla funkcji niemalejącej mamy układ nierówności:

$$\left(\bigwedge_{k=1}^{n-1} \right) y_k \leq y_{k+1}, \quad (6.17)$$

natomiast dla funkcji nierosnącej jest to układ nierówności:

$$\left(\bigwedge_{k=1}^{n-1} \right) y_k \geq y_{k+1}. \quad (6.18)$$

Udowodnijmy dwa istotne lematy, który ograniczają zbiór, w którym należy szukać rozwiązania minimalizującego wartość $Z_1(F_A, F_B)$ podaną wzorem (6.16).

Lemat 6.1 (O położeniu rozwiązania minimalnego dla ciągu niemalejącego)

Niech (y_1, y_2, \dots, y_n) będzie ciągiem spełniającym ciąg nierówności (6.17) i minimalizującym wartość podaną wzorem (6.16). Wówczas:

$$\min(b_1, b_2, \dots, b_n) \leq y_i \leq \max(b_1, b_2, \dots, b_n). \quad (6.19)$$

Dowód. Niech $t = \min(b_1, b_2, \dots, b_n)$. Załóżmy nie wprost, że istnieje ciąg minimalizujący wartość podaną wzorem (6.16), w których dla pewnego m zachodzi $y_m < t$. Wykażemy teraz, że w takim wypadku istnieje ciąg $(z_1, z_2, \dots, z_n) = (t, t, \dots, t, y_{m+1}, y_{m+2}, \dots, y_n)$, dla którego spełniony jest warunek (6.17) i jednocześnie ma mniejszą wartość funkcji (6.16), co daje nam sprzeczność i dowodzi tezy, że $\min(b_1, b_2, \dots, b_n) \leq y_i$. Zaczniemy od nierówności, prawdziwej dla $i = 1, 2, \dots, m$:

$$y_i < t \leq b_i. \quad (6.20)$$

Wobec tego:

$$t + y_i - 2b_i < 0, \quad (6.21)$$

$$y_i - t < 0, \quad (6.22)$$

$$(y_i - b_i)^2 - (t - b_i)^2 = (y_i - t)(t + y_i - 2b_i) > 0, \quad (6.23)$$

$$(t - b_i)^2 < (y_i - b_i)^2. \quad (6.24)$$

Sumując to po wszystkich $i = 1, 2, \dots, m$ mamy:

$$\sum_{i=1}^m (t - b_i)^2 < \sum_{i=1}^m (y_i - b_i)^2, \quad (6.25)$$

Dodajmy stronami składniki odpowiedzialne za wyrazy dla $i = m + 1, m + 2, \dots, n$:

$$\sum_{i=1}^m (t - b_i)^2 + \sum_{i=m+1}^n (y_i - b_i)^2 < \sum_{i=1}^n (y_i - b_i)^2. \quad (6.26)$$

Zauważmy, że lewa strona nierówności (6.26) odpowiada wyrażeniu (6.16) dla ciągu (z_1, z_2, \dots, z_n) , a prawa strona dla ciągu (y_1, y_2, \dots, y_n) , który z założenia nie wprost, miał być ciągiem o najmniejszej wartości wyrażenia (6.16) - co daje nam sprzeczność.

Przejdźmy teraz do dowodu ograniczenia z góry. Niech $T = \max(b_1, b_2, \dots, b_n)$, a ciąg (y_1, y_2, \dots, y_n) jest ciągiem o najmniejszej wartości wyrażenia (6.16) oraz dla pewnego m zachodzi $y_m > T$. Pokażemy, że dla ciągu $(z_1, z_2, \dots, z_n) = (y_1, y_2, \dots, y_{m-1}, T, \dots, T)$ wartość wyrażenia (6.16) jest mniejsza niż dla ciągu (y_1, y_2, \dots, y_n) , co dając sprzeczność, dowodzi, że musi zachodzić $y_i \leq T$ dla każdego $i = 1, 2, \dots, n$.

Zauważmy, że:

$$y_i > T \geq b_i, \quad (6.27)$$

$$T + y_i - 2b_i > 0, \quad (6.28)$$

$$(y_i - b_i)^2 - (T - b_i)^2 = (T + y_i - 2b_i)(y_i - T) > 0, \quad (6.29)$$

$$(T - b_i)^2 < (y_i - b_i)^2. \quad (6.30)$$

Sumując to po wszystkich $i = 1, 2, \dots, m$ mamy:

$$\sum_{i=1}^m (T - b_i)^2 < \sum_{i=1}^m (y_i - b_i)^2. \quad (6.31)$$

Dodajmy stronami składniki odpowiedzialne za wyrazy dla $i = m + 1, m + 2, \dots, n$:

$$\sum_{i=m}^n (y_i - b_i)^2 + \sum_{i=1}^m (T - b_i)^2 < \sum_{i=1}^n (y_i - b_i)^2. \quad (6.32)$$

Zauważmy, że lewa strona nierówności (6.32) odpowiada wyrażeniu (6.16) dla ciągu (z_1, z_2, \dots, z_n) , a prawa strona dla ciągu (y_1, y_2, \dots, y_n) , który z założenia nie wprost, miał być ciągiem o najmniejszej wartości wyrażenia (6.16) - co daje nam sprzeczność. ■

Lemat 6.2 (O położeniu rozwiązania minimalnego dla ciągu nierosnącego)

Niech (y_1, y_2, \dots, y_n) będzie ciągiem spełniającym ciąg nierówności (6.18) i minimalizującym wartość (6.16). Wówczas:

$$\min(b_1, b_2, \dots, b_n) \leq y_i \leq \max(b_1, b_2, \dots, b_n). \quad (6.33)$$

Dowód. Niech $t = \min(b_1, b_2, \dots, b_n)$. Załóżmy nie wprost, że istnieje ciąg minimalizujący wartość (6.16), w których dla pewnego m zachodzi $y_m < t$. Wykażemy teraz, że w takim wypadku istnieje ciąg $(z_1, \dots, z_n) = (y_1, y_2, \dots, y_m, t, t, \dots, t)$, dla którego spełniony jest warunek (6.17) i jednocześnie ma mniejszą wartość funkcji (6.16), co daje nam sprzeczność i dowodzi tezy, że $\min(b_1, b_2, \dots, b_n) \leq y_i$. Zaczniemy od nierówności, prawdziwej dla $i = m, m + 1, \dots, n$:

$$y_i < t \leq b_i. \quad (6.34)$$

Wobec tego:

$$t + y_i - 2b_i < 0, \quad (6.35)$$

$$y_i - t < 0, \quad (6.36)$$

$$(y_i - b_i)^2 - (t - b_i)^2 = (y_i - t)(t + y_i - 2b_i) > 0, \quad (6.37)$$

$$(t - b_i)^2 < (y_i - b_i)^2. \quad (6.38)$$

Sumując to po wszystkich $i = m, \dots, n$ mamy:

$$\sum_{i=m}^n (t - b_i)^2 < \sum_{i=m}^n (y_i - b_i)^2. \quad (6.39)$$

Dodajmy stronami składniki odpowiedzialne za wyrazy dla $i = 1, \dots, m - 1$:

$$\sum_{i=m}^n (t - b_i)^2 + \sum_{i=1}^{m-1} (y_i - b_i)^2 < \sum_{i=1}^n (y_i - b_i)^2. \quad (6.40)$$

Zauważmy, że lewa strona nierówności (6.40) odpowiada wyrażeniu (6.16) dla ciągu (z_1, \dots, z_n) , a prawa strona dla ciągu (y_1, \dots, y_n) , który z założenia nie wprost, miał być ciągiem o najmniejszej wartości wyrażenia (6.16) - co daje nam sprzeczność.

Przejdźmy teraz do dowodu ograniczenia z góry. Niech $T = \max(b_1, b_2, \dots, b_n)$, a ciąg (y_1, y_2, \dots, y_n) jest ciągiem o najmniejszej wartości wyrażenia (6.16) oraz dla pewnego m zachodzi $y_m > T$. Pokażemy, że dla ciągu $(z_1, z_2, \dots, z_n) = (T, \dots, T, y_m, y_{m+1}, \dots, y_n)$ wartość wyrażenia (6.16) jest mniejsza niż dla ciągu (y_1, y_2, \dots, y_n) , co dając sprzeczność, dowodzi, że musi zachodzić $y_i \leq T$ dla każdego $i = 1, 2, \dots, n$.

Zauważmy, że:

$$y_i > T \geq b_i, \quad (6.41)$$

$$T + y_i - 2b_i > 0, \quad (6.42)$$

$$(y_i - b_i)^2 - (T - b_i)^2 = (T + y_i - 2b_i)(y_i - T) > 0, \quad (6.43)$$

$$(T - b_i)^2 < (y_i - b_i)^2. \quad (6.44)$$

Sumując to po wszystkich $i = m, m + 1, \dots, n$ mamy:

$$\sum_{i=m}^n (T - b_i)^2 < \sum_{i=m}^n (y_i - b_i)^2. \quad (6.45)$$

Dodajmy stronami składniki odpowiedzialne za wyrazy dla $i = 1, 2, \dots, m - 1$:

$$\sum_{i=1}^{m-1} (y_i - b_i)^2 + \sum_{i=m}^n (T - b_i)^2 < \sum_{i=1}^n (y_i - b_i)^2, \quad (6.46)$$

Zauważmy, że lewa strona nierówności (6.46) odpowiada wyrażeniu (6.16) dla ciągu (z_1, z_2, \dots, z_n) , a prawa strona dla ciągu (y_1, y_2, \dots, y_n) , który z założenia nie wprost, miał być ciągiem o najmniejszej wartości wyrażenia (6.16) - co daje nam sprzeczność. ■

Problem opisany formułami (6.16) i (6.17) to tzw. zadanie programowania kwadratowego (*Quadratic Programming problem* - QP) [95, 104]; analogicznie jest dla problemu opisanego formułami (6.16) i (6.18). W ogólnej formie, problem QP jest zdefiniowany jako:

$$\begin{aligned} \text{Minimalizuj } & \frac{1}{2} \mathbf{y}^T \mathbf{Q} \mathbf{y} + \mathbf{w}^T \mathbf{y} \\ \text{przy ograniczeniu } & \mathbf{P} \mathbf{y} \preceq \mathbf{r} \\ & \text{oraz } \mathbf{y} \succeq \mathbf{0}, \end{aligned} \tag{6.47}$$

gdzie \mathbf{y} jest wektorem nieznanymi zmiennymi. Notacja $\mathbf{P} \mathbf{y} \preceq \mathbf{r}$ oznacza, że każdy element wektora $\mathbf{P} \mathbf{y}$ jest mniejszy bądź równy odpowiadającemu elementowi wektora \mathbf{r} .

Problem QP może być rozwiązany przez kilka znanych metod, przykładowo [86]:

- metoda zbioru aktywnego (*active set*),
- metody punktu wewnętrznego (*interior-point methods*),
- metoda gradientu (*gradient projection*).

6.2.2. Warunki Karush-Kuhn-Tuckera

Dla problemu QP istnieją warunki konieczne, które musi spełniać rozwiązanie, aby było to rozwiązanie w punkcie ekstremum. Warunki te, nazywane warunkami Karush-Kuhn-Tuckera (KKT) [98] są następujące:

$$\begin{cases} \nabla Z_1(\mathbf{y}) + \sum_{k=1}^{n-1} \mu_k \nabla P_k(\mathbf{y}) = 0, \\ \mathbf{P} \mathbf{y} \leq 0, \\ \mu_k P_k(\mathbf{y}) = 0, \quad k = 1, \dots, n-1. \end{cases} \tag{6.48}$$

Używamy standardowego oznaczenia $\nabla = [\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}, \dots, \frac{\partial}{\partial y_n}]^T$. Stałe $\mu_k \in \mathbb{R}$ to tzw. mnożniki Lagrange'a [98].

W naszej sytuacji mamy następujące stałe, definiujące problem QP dla funkcji niemalejącej:

$$\mathbf{Q} = 2\mathbf{I}, \text{ gdzie } \mathbf{I} \text{ jest macierzą jednostkową } n \times n, \quad (6.49)$$

$$\mathbf{w} = -2\mathbf{b}, \quad (6.50)$$

$$\mathbf{P} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix}, \quad (6.51)$$

$$\mathbf{r} = \mathbf{0}. \quad (6.52)$$

$$\left(\bigwedge_{k=1}^{n-1} \right) P_k(\mathbf{y}) = y_k - y_{k+1}. \quad (6.53)$$

W naszym przypadku z warunków KKT (6.48) otrzymujemy:

$$\frac{\partial Z_1(\mathbf{y})}{\partial y_k} = 2(y_k - b_k), \quad (6.54)$$

$$\sum_{k=1}^{n-1} \mu_k \frac{\partial P_k(\mathbf{x})}{\partial y_j} = \begin{cases} \mu_1, & \text{gd } j = 1, \\ \mu_j - \mu_{j-1}, & \text{gd } j = 2, 3, \dots, n-1, \\ -\mu_{n-1}, & \text{gd } j = n, \end{cases} \quad (6.55)$$

co nam daje następujące równania:

$$2(y_j - b_j) = \begin{cases} -\mu_1, & \text{gd } j = 1, \\ -\mu_j + \mu_{j-1}, & \text{gd } j = 2, 3, \dots, n-1, \\ \mu_{n-1}, & \text{gd } j = n, \end{cases} \quad (6.56)$$

$$\mu_j(y_j - y_{j+1}) = 0. \quad (6.57)$$

Można zauważyć, że dla $j = 2, 3, \dots, n-1$ mamy:

$$2(y_j - b_j)(y_j - y_{j+1})(y_{j-1} - y_j) = (-\mu_j + \mu_{j-1})(y_j - y_{j+1})(y_{j-1} - y_j) = 0. \quad (6.58)$$

Analogiczna reguła do reguły (6.58) może być również otrzymana dla $j = 1$ oraz $j = n$:

$$(y_1 - b_1)(y_1 - y_2) = 0, \quad (6.59)$$

$$(y_n - b_n)(y_{n-1} - y_n) = 0. \quad (6.60)$$

Wniosek 6.1 (Wniosek z warunków KKT dla ciągu niemalejącego)

W rozwiązaniu optymalnym problemu QP opisanego przez formuły (6.16 i (6.17)), punkt y_i ma taką samą wartość co poprzedni punkt y_{i-1} (jeśli istnieje) lub następny punkt y_{i+1} (jeśli istnieje) lub punkt b_i z oryginalnego ciągu.

Rozpatrzmy teraz sytuację funkcji nierosnącej, gdzie zamiast (6.17) zachodzi (6.18), a wobec tego macierz \mathbf{P} ma przeciwny znak. Rozpatrzmy teraz warunki KKT dla takiej funkcji.

$$\frac{\partial Z_1(\mathbf{y})}{\partial y_k} = 2(y_k - b_k), \quad (6.61)$$

$$\sum_{k=1}^{n-1} \mu_k \frac{\partial P_k(\mathbf{x})}{\partial y_j} = \begin{cases} -\mu_1, & \text{gdy } j = 1, \\ -\mu_j + \mu_{j-1}, & \text{gdy } j = 2, 3, \dots, n-1, \\ \mu_{n-1}, & \text{gdy } j = n, \end{cases} \quad (6.62)$$

co nam daje następujące równania:

$$2(y_j - b_j) = \begin{cases} \mu_1, & \text{gdy } j = 1, \\ \mu_j - \mu_{j-1}, & \text{gdy } j = 2, 3, \dots, n-1, \\ -\mu_{n-1}, & \text{gdy } j = n, \end{cases} \quad (6.63)$$

$$\mu_j(-y_j + y_{j+1}) = 0. \quad (6.64)$$

Prowadzi to do analogicznego wniosku, jak w przypadku funkcji niemalejącej, co ilustrują następujące równania. Dla $j = 2, 3, \dots, n-1$ mamy:

$$\begin{aligned} 0 &= \mu_j(-y_j + y_{j+1})(-y_{j-1} + y_j) - \mu_{j-1}(-y_{j-1} + y_j)(-y_j + y_{j+1}) = \\ &= (\mu_j - \mu_{j-1})(-y_j + y_{j+1})(-y_{j-1} + y_j) = 2(y_j - b_j)(-y_j + y_{j+1})(-y_{j-1} + y_j). \end{aligned} \quad (6.65)$$

Analogiczna reguła do reguły (6.65) może być również otrzymana dla $j = 1$ oraz $j = n$:

$$(y_1 - b_1)(-y_1 + y_2) = 0, \quad (6.66)$$

$$(y_n - b_n)(-y_{n-1} + y_n) = 0. \quad (6.67)$$

Wniosek 6.2 (Wniosek z warunków KKT dla ciągu nierosnącego)

W rozwiązaniu optymalnym problemu QP opisanego przez formuły (6.16) i (6.18), punkt y_i ma taką samą wartość co poprzedni punkt y_{i-1} (jeśli istnieje) lub następny punkt y_{i+1} (jeśli istnieje) lub punkt b_i z oryginalnego ciągu.

6.2.3. Redukcja problemu QP do problemu LP

Spróbujmy teraz sprowadzić problem QP do problemu Programowania Liniowego (*Linear Programming problem* - LP). Oznaczmy:

$$L = \sum_{k=1}^n b_k^2 = const, \quad (6.68)$$

co nam daje:

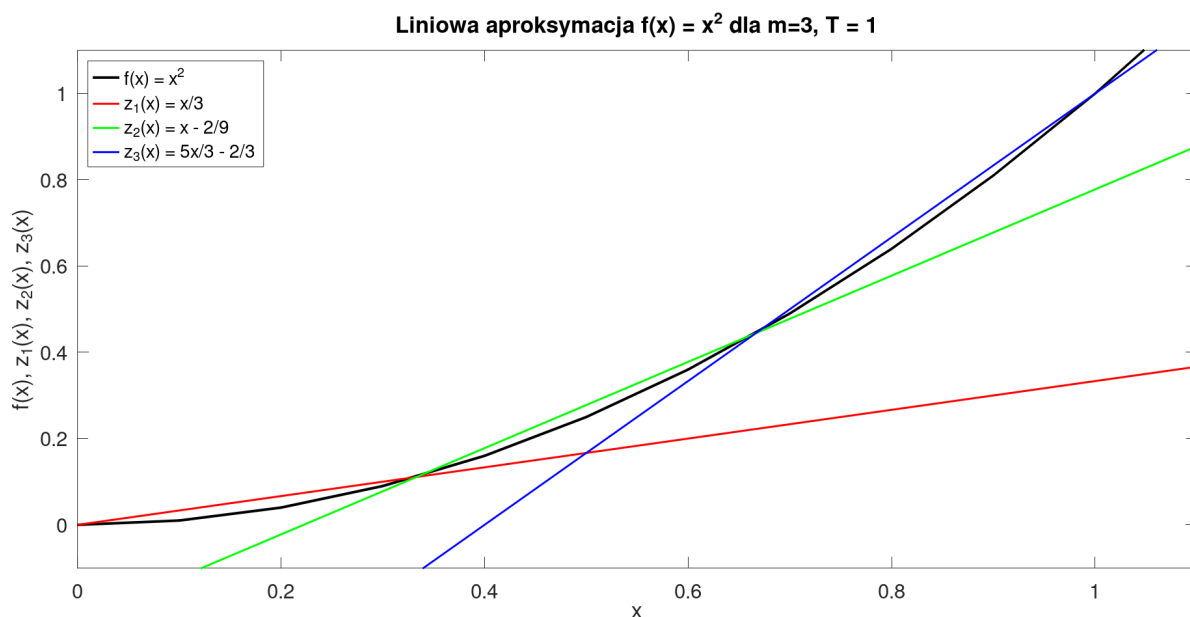
$$Z_1(F_A, F_B) = \sum_{k=1}^n (b_k - y_k)^2 = L - 2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n y_k^2, \quad (6.69)$$

Wystarczy więc minimalizować wyrażenie:

$$Z_2(F_A, F_B) = -2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n y_k^2. \quad (6.70)$$

Rozwiązanie dla $Z_2(F_A, F_B)$ jest również rozwiązaniem dla $Z_1(F_A, F_B)$. Możemy więc przeformułować wyjściowy problem dla ciągów niemalejących do postaci:

$$\begin{aligned} \text{Minimalizuj } Z_2(F_A, F_B) &= -2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n y_k^2, \\ \text{przy ograniczeniu } &\left(\bigwedge_{k=1}^{n-1} \right) y_{k+1} - y_k \geq 0, \\ &\text{oraz } \left(\bigwedge_{k=1}^n \right) y_k \geq 0. \end{aligned} \quad (6.71)$$



Rysunek 6.2: Przykład aproksymacji liniowej.

Aby przekształcić ten problem do problemu LP, trzeba wykorzystać dodatkowe zmienne (*slack variables*) s_k . Daje to następującą postać:

$$\begin{aligned} \text{Minimalizuj } Z_2(F_A, F_B) &= -2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n y_k^2, \\ \text{przy ograniczeniu } \left(\bigwedge_{k=1}^{n-1} \right) & y_{k+1} - y_k - s_k = 0, \\ \text{oraz } \begin{cases} \left(\bigwedge_{k=1}^n \right) & y_k \geq 0, \\ \left(\bigwedge_{k=1}^{n-1} \right) & s_k \geq 0. \end{cases} \end{aligned} \quad (6.72)$$

Analogicznie, dla ciągów nierosnących będzie to problem w postaci:

$$\begin{aligned} \text{Minimalizuj } Z_2(F_A, F_B) &= -2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n y_k^2, \\ \text{przy ograniczeniu } \left(\bigwedge_{k=1}^{n-1} \right) & -y_{k+1} + y_k - s_k = 0, \\ \text{oraz } \begin{cases} \left(\bigwedge_{k=1}^n \right) & y_k \geq 0, \\ \left(\bigwedge_{k=1}^{n-1} \right) & s_k \geq 0. \end{cases} \end{aligned} \quad (6.73)$$

Funkcja $f(x) = x^2$ jest wypukła i dla $x > 0$ może być przybliżana przez zbiór funkcji liniowych [98]. Przykładowa aproksymacja tej funkcji została przedstawiona na rysunku 6.2. Z lematu 6.1 (względnie z lematu 6.2 dla ciągów nierosnących) wiemy, że będziemy poszukiwać rozwiązań y_1, y_2, \dots, y_n na przedziale $[t; T] = [\min\{b_1, b_2, \dots, b_n\}; \max\{b_1, b_2, \dots, b_n\}]$. Podzielmy odcinek równomiernie na m części za pomocą punktów $\tau_0, \tau_1, \dots, \tau_m$. Mam wtedy:

$$\left(\bigwedge_{k=0}^m \right) \tau_k = t + \frac{T-t}{m}k. \quad (6.74)$$

Szukamy równań prostych, które spełniają zależności:

$$\left(\bigwedge_{k=1}^m \right) \begin{cases} \tau_{k-1}^2 = \alpha_k \tau_{k-1} + \beta_k, \\ \tau_k^2 = \alpha_k \tau_k + \beta_k. \end{cases} \quad (6.75)$$

Rozwiązana tych układów równań są następujące:

$$\left(\bigwedge_{k=1}^m \right) \begin{cases} \alpha_k = \tau_{k-1} + \tau_k, \\ \beta_k = -\tau_{k-1}\tau_k. \end{cases} \quad (6.76)$$

Daje to zbiór prostych $z_1(x), z_2(x), \dots, z_m(x)$, opisanych równaniami:

$$\begin{aligned} \left(\bigwedge_{k=1}^m \right) z_k(x) &= \left(2t + \frac{T-t}{m}(2k-1) \right) x - \left(t + \frac{T-t}{m}k \right) \left(t + \frac{T-t}{m}(k-1) \right) = \\ &= \left(2t + \frac{T-t}{m}(2k-1) \right) x - t^2 - \frac{T-t}{m}t(2k-1) - \left(\frac{T-t}{m} \right)^2 k(k-1). \end{aligned} \quad (6.77)$$

Dla takiego zbioru funkcji zachodzą dwa następujące lematy.

Lemat 6.3 (O lokalnym optimum)

Jeśli $\tau_{k-1} \leq x \leq \tau_k$ dla pewnego $1 \leq k \leq m$, to:

$$z_k(x) \geq z_r(x), \quad (6.78)$$

dla każdego $r \in \{1, 2, \dots, m\}$.

Dowód. Obliczmy wartość wyrażenia $z_k(x) - z_r(x)$ na podstawie wzoru (6.77):

$$\begin{aligned}
& z_k(x) - z_r(x) = \\
& = \left(\left(2t + \frac{T-t}{m}(2k-1) \right) x - t^2 - \frac{T-t}{m}t(2k-1) - \left(\frac{T-t}{m} \right)^2 k(k-1) \right) - \\
& \quad - \left(\left(2t + \frac{T-t}{m}(2r-1) \right) x - t^2 - \frac{T-t}{m}t(2r-1) - \left(\frac{T-t}{m} \right)^2 r(r-1) \right) = \\
& = \left(\frac{T-t}{m}(2k-2r) \right) x - \frac{T-t}{m}t(2k-2r) - \left(\frac{T-t}{m} \right)^2 (k(k-1) - r(r-1)) = \\
& \quad = 2(k-r)\frac{T-t}{m}(x-t) - \left(\frac{T-t}{m} \right)^2 ((k-r)(k+r) - (k-r)) = \\
& \quad = (k-r)\frac{T-t}{m} \left(2(x-t) - \left(\frac{T-t}{m} \right) (k+r-1) \right). \tag{6.79}
\end{aligned}$$

Wiemy, że dla pewnego $0 \leq \alpha \leq 1$ zachodzi:

$$x = t + \frac{T-t}{m}(k-1+\alpha), \tag{6.80}$$

a więc:

$$\begin{aligned}
z_k(x) - z_r(x) &= (k-r)\frac{T-t}{m} \left(2\frac{T-t}{m}(k-1+\alpha) - \left(\frac{T-t}{m} \right) (k+r-1) \right) = \\
&= (k-r) \left(\frac{T-t}{m} \right)^2 (2(k-1+\alpha) - (k+r-1)) = \\
&= (k-r) \left(\frac{T-t}{m} \right)^2 (k-r-1+2\alpha). \tag{6.81}
\end{aligned}$$

Niech $q = k - r$. Dla udowodnienia tezy, wystarczy pokazać, że:

$$0 \leq q(q-1+2\alpha). \tag{6.82}$$

Wiemy, że q przyjmuje wartości całkowite z przedziału $[1-m; m-1]$, a powyższa funkcja kwadratowa przyjmuje wartości ujemne na przedziale $(0; 1-2\alpha)$ albo $(1-2\alpha; 0)$, w zależności od tego, czy $\alpha > 0.5$, czy $\alpha < 0.5$. Zauważmy, że żaden z tych dwóch przedziałów nie zawiera liczb całkowitych (przedziały są obustronnie otwarte). Dla wartości granicznych tych przedziałów, funkcja kwadratowa się zeruje. Dla $\alpha = 0.5$ mamy oczywistą nierówność $0 \leq q^2$. ■

Lemat 6.4 (O zależności między relaksacją liniową a funkcją kwadratową)

Jeśli $\tau_{k-1} \leq x \leq \tau_k$, to:

$$z_k(x) \geq f(x) = x^2, \quad (6.83)$$

a równość w wyrażeniu (6.83) zachodzi wtedy i tylko wtedy, gdy $x \in \{\tau_{k-1}, \tau_k\}$.

Dowód. Mamy na podstawie (6.76):

$$\Delta f(x) = z_k(x) - x^2 = (\tau_{k-1} + \tau_k)x - \tau_{k-1}\tau_k - x^2. \quad (6.84)$$

Powyzsza funkcja kwadratowa względem x ma miejsca zerowe w punktach:

$$x_0 = \frac{-(\tau_{k-1} + \tau_k) \pm \sqrt{(\tau_{k-1} + \tau_k)^2 - 4\tau_{k-1}\tau_k}}{-2} = \frac{(\tau_{k-1} + \tau_k) \mp |\tau_k - \tau_{k-1}|}{2}. \quad (6.85)$$

Co oznacza, że funkcja zeruje się w punktach $\{\tau_{k-1}, \tau_k\}$. Współczynnik przy x^2 jest równy -1 , więc pomiędzy tymi punktami funkcja przyjmuje wartości dodatnie, co dowodzi wyjściowej nierówności. ■

Z wykorzystaniem zdefiniowanego zbioru funkcji liniowych, oryginalna funkcja kwadratowa jest przybliżania za pomocą [98]:

$$f(x) \approx \max\{z_k(x) : k = 1, 2, \dots, m\}. \quad (6.86)$$

Aby dokonać relaksacji problemu QP do problemu LP, musimy wprowadzić ograniczenia [98]:

$$\left(\bigwedge_{k=1}^n\right) \left(\bigwedge_{j=1}^m\right) E_k \geq \left(2t + \frac{T-t}{m}(2j-1)\right) y_k - t^2 - \frac{T-t}{m}t(2j-1) - \left(\frac{T-t}{m}\right)^2 j(j-1). \quad (6.87)$$

Wprowadzając nieujemne zmienne $r_{k,j}$, możemy przekształcić nasz problem (w wariacie dla ciągów niemalejących) do poniższego problemu programowania liniowego:

$$\begin{aligned} & \text{Minimalizuj } Z_3(F_A, F_B) = -2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n (E_k), \\ \text{przy ograniczeniu } & \begin{cases} (\bigwedge_{k=1}^{n-1}) & y_{k+1} - y_k - s_k = 0, \\ (\bigwedge_{k=1}^n) (\bigwedge_{j=1}^m) & E_k - (2t + \frac{T-t}{m}(2j-1)) y_k + \\ & + t^2 + \frac{T-t}{m} t(2j-1) + (\frac{T-t}{m})^2 j(j-1) - r_{k,j} = 0, \end{cases} \quad (6.88) \\ & \text{oraz } \begin{cases} (\bigwedge_{k=1}^n) & y_k \geq 0, \\ (\bigwedge_{k=1}^{n-1}) & s_k \geq 0, \\ (\bigwedge_{k=1}^n) & E_k \geq 0, \\ (\bigwedge_{k=1}^n) (\bigwedge_{j=1}^m) & r_{k,j} \geq 0. \end{cases} \end{aligned}$$

Analogicznie, dla ciągów nierosnących jest to problem:

$$\begin{aligned} & \text{Minimalizuj } Z_3(F_A, F_B) = -2 \sum_{k=1}^n b_k y_k + \sum_{k=1}^n (E_k), \\ \text{przy ograniczeniu } & \begin{cases} (\bigwedge_{k=1}^{n-1}) & -y_{k+1} + y_k - s_k = 0, \\ (\bigwedge_{k=1}^n) (\bigwedge_{j=1}^m) & E_k - (2t + \frac{T-t}{m}(2j-1)) y_k + \\ & + t^2 + \frac{T-t}{m} t(2j-1) + (\frac{T-t}{m})^2 j(j-1) - r_{k,j} = 0, \end{cases} \quad (6.89) \\ & \text{oraz } \begin{cases} (\bigwedge_{k=1}^n) & y_k \geq 0, \\ (\bigwedge_{k=1}^{n-1}) & s_k \geq 0, \\ (\bigwedge_{k=1}^n) & E_k \geq 0, \\ (\bigwedge_{k=1}^n) (\bigwedge_{j=1}^m) & r_{k,j} \geq 0. \end{cases} \end{aligned}$$

Zdefiniowany problem składa się z $n(m+3) - 1$ zmiennych i $n(m+1) - 1$ równań, gdzie m jest liczbą funkcji liniowych użytych do aproksymacji funkcji kwadratowej. Problem LP może być rozwiązany np. z użyciem algorytmu Simplex [98] praktycznie w czasie $O(nm)$; w pesymistycznym przypadku wymaga to $O(2^{nm})$ operacji.

Możemy również wyznaczyć błąd aproksymacji wynikły z powodu użycia funkcji liniowych. Na podstawie wzoru (6.84) dla $\tau_{k-1} \leq x \leq \tau_k$ błąd aproksymacji $\Delta f(x)$ ma charakter funkcji kwadratowej, o największej wartości dla wierzchołka paraboli, tzn. dla $x = \frac{\tau_{k-1} + \tau_k}{2}$. Dla takiego punktu mamy:

$$\begin{aligned} \Delta f(x_{max}) &= (\tau_{k-1} + \tau_k) \frac{\tau_{k-1} + \tau_k}{2} - \tau_{k-1} \tau_k - \left(\frac{\tau_{k-1} + \tau_k}{2} \right)^2 = \\ &= \left(\frac{\tau_{k-1} - \tau_k}{2} \right)^2 = \left(\frac{T - t}{2m} \right)^2. \end{aligned} \quad (6.90)$$

Podział problemu znalezienia najbliższej rozmytej liczby wypukłej pociąga za sobą istotne wyzwanie - wyznaczone rozwiązania mogą się wzajemnie *krzyżować (przenikać)*, czyli nie tworzą wypukłej SLR. Można próbować zastosować tutaj następujące podejścia.

- Rozpatrywanie funkcji celu i zbioru ograniczeń wspólnych dla obydwu funkcji tworzących SLR. Niestety, funkcja celu ma wtedy wyjściową postać:

$$\begin{aligned} \min d(A, B) = \min & \sqrt{\sum_{i=k}^n ((a_{f,k} - x_{f,k})^2 + (b_{f,k} - y_{f,k})^2)} + \\ & + \sqrt{\sum_{i=k}^n ((a_{g,k} - x_{g,k})^2 + (b_{g,k} - y_{g,k})^2)}. \end{aligned} \quad (6.91)$$

Nie można dla takiej funkcji celu powtórzyć rozumowania analogicznego do przedstawionego w tym rozdziale.

- Rozpatrywanie innej metryki. Przykładowo, może to być metryka postaci:

$$d_2(A, B) = \|f_A(x) - f_B(x)\|^2 + \|g_A(x) - g_B(x)\|^2. \quad (6.92)$$

Zaletą tego podejścia jest możliwość uzyskania lepszych rozwiązań (mniejszej wartości według tej metryki) niż za pomocą pozostałych metod. Wadą jest zwiększenie wielkości problemu, a algorytm Simplex wykorzystywany do rozwiązywania tego typu problemów ma pesymistyczną złożoność o charakterze wykładniczym.

- Założyć, że niemalejąca część SLR ma wartości mniejsze niż pewna wartość rozgraniczająca, a nierosnąca część odpowiednio większe niż określona wartość.

— Dokonać optymalizacji najpierw jednej funkcji, a następnie korzystając z otrzymanego wyniku, dokonać optymalizacji drugiej funkcji. W tym podejściu do zbioru równań należy dodać nowe równanie (i zmienną), które opisuje zależność pomiędzy otrzymanym wynikiem dla pierwszej funkcji, a poszukiwanym wynikiem dla drugiej funkcji.

W niniejszej pracy skorzystano z ostatniego podejścia, optymalizując najpierw jedną funkcję tworzącą SLR, a następnie drugą funkcję.

6.3. ALGORYTM WYZNACZANIA IZOLACJI Z WYKORZYSTANIEM SLR

Przedstawmy teraz algorytm wyznaczania izolacji z użyciem SLR. Składa się on z algorytmów 2 - 7, które przedstawiają główne zagadnienia związane z obliczaniem izolacji według modelu rozmytego. Główna idea polega na wykonaniu algorytmu 0, ale wykorzystania do obliczeń SLR skonstruowanych na podstawie serii pomiarowych, zamiast pojedynczych pomiarów. Warto skomentować kwestię wag użytych wewnątrz funkcji łączącej - przyjęto tutaj równy wpływ poszczególnych węzłów na wartość izolacji, w praktyce mogą wystąpić przesłanki do użycia innego zestawu wag. Wagi użyte do łączenia izolacji pomiędzy podsieciami zależą od wkładu danej podsieci w ramach danej cechy. Wkład jest określany poprzez liczbę węzłów opisanych danymi parametrem w danej podsieci.

6.4. PORZĄDEK W MODELU ROZMYTYM

Zajmijmy się teraz nad zagadnieniem porównywania poziomów izolacji w ramach modelu rozmytego. W podrozdziale 4.4. wprowadziliśmy sposób tworzenia porządku liniowego z wektora izolacji za pomocą funkcji:

$$\Phi_q(x_1, x_2, \dots, x_n, \dots, x_N) = 1 - \left(\frac{1}{N} \sum_{i=1}^N (1 - x_i)^q \right)^{\frac{1}{q}}. \quad (6.93)$$

W przypadku liczb rzeczywistych, powyższe przekształcenie rozwiązuje problem możliwości porównywania poziomów izolacji w sposób kompletny - dowolne dwie liczby rzeczywiste

można ze sobą porównać. Inna sytuacja występuje w przypadku liczb rozmytych oraz skierowanych liczb rozmytych. Opiszmy teraz prostą sytuację określania porządku dla dwóch *prostokątnych* liczb rozmytych; wykorzystamy w tym celu podejście przedstawione w [128], gdzie rozpatrywano porządek w dziedzinie przedziałów.

Założmy, że mamy dwie liczby rozmyte (niekoniecznie skierowane) A i B , określone odpowiednio na przedziałach $[a_1, a_2]$ i $[b_1, b_2]$; $a_1 \leq a_2, b_1 \leq b_2$ oraz ich funkcje przynależności są równe 1 na tych przedziałach. Rozpatrzmy następujące przypadki wzajemnego rozmieszczenia tych liczb.

1. $a_1 \leq a_2 \leq b_1 \leq b_2$. Gdy zachodzi $a_1 = b_2$, to $A = B$, w przeciwnym wypadku uznajemy, że $A < B$.
2. $a_1 \leq b_1 \leq a_2 \leq b_2$. Gdy zachodzi $a_1 = b_2 \vee (a_1 = b_1 \wedge a_2 = b_2)$, to $A = B$, w przeciwnym wypadku uznajemy, że $A < B$.
3. $b_1 \leq a_1 \leq a_2 \leq b_2$. Gdy zachodzi $b_1 = b_2$, to $A = B$. W pozostałych przypadkach liczba B jest jednocześnie *większa* i *mniejsza* od A (w sensie potocznym, nie porządku matematycznego). W takiej sytuacji, gdy:

$$a_1 - b_1 < b_2 - a_2 \implies A < B, \quad (6.94)$$

oraz:

$$a_1 - b_1 > b_2 - a_2 \implies A > B. \quad (6.95)$$

Powyższy warunek można również geometrycznie zinterpretować w następujący sposób: *wzajemne położenie osi symetrii liczb rozmytych A i B wyznacza porządek tych dwóch liczb.*

Ze względu na charakter danych wejściowych i przekształceń matematycznych, najczęściej porównaniu nie będą podlegać *prostokątne* liczby rozmyte, wobec czego potrzebne jest nieco inne podejście do tego zagadnienia. Z perspektywy badania izolacji, mamy następujące fundamentalne wymagania:

1. Dla znanego z góry poziomu izolacji T chcemy określić, czy wyznaczony poziom izolacji A spełnia warunek $T \leq A$.
2. Dla dwóch poziomów izolacji A i B chcemy określić, który z nich jest wyższy, tzn. czy zachodzi $A \leq B$ czy $B \leq A$.

Do realizacji pierwszego wymagania wystarcza posługiwanie się częściowym porządkiem - w takim wypadku, gdy T i A nie są w relacji $T \leq A$ (nie zachodzi relacja $T \leq A$), to nie zachodzi potrzeba weryfikowania, czy T i A są w relacji $A \leq T$ (czy zachodzi relacja $A \leq T$).

Drugie wymaganie oznacza konieczność określenia porządku liniowego, który pozwala porównywać dwa dowolne elementy z zadanego zbioru. Możliwe jest określenie takiego porządku liniowego na zbiorze liczb rozmytych, np. w sposób zdefiniowany w pracy [120]. Pomysł przedstawiony przez autorów polega na zdefiniowaniu podzbioru gęstego zbioru $[0; 1]$ i sprawdzaniu cech funkcji przynależności dla kolejnych wartości ze zdefiniowanego zbioru gęstego. Zbiór taki jest przeliczalny, ale nieskończony, co oznacza, że proces porównywania dwóch liczb rozmytych może trwać w zasadzie dowolnie długo.

Zastanówmy się teraz nad wymaganiami, jakie mamy wobec porządku liniowego, z perspektywy domeny izolacji. Podobnie jak w przypadku tworzenia funkcji łączących izolacje, istotny jest *charakter* użytego porządku liniowego. Kluczowe jest tutaj wybranie odpowiedniego rozstrzygnięcia dla przypadków, w których:

$$Supp(A) \cap Supp(B) \neq \emptyset, \quad (6.96)$$

czyli nośniki porównywanych liczb rozmytych pokrywają się na pewnym niepustym przedziale - a więc nie ma mocnego rozgraniczenia pomiędzy tymi liczbami rozmytymi. Interesują nas tutaj przedziały niezdegenerowane do pojedynczego punktu; dla takiego przypadku możliwe jest dość proste zdefiniowanie porządku liniowego dla dwóch liczb rozmytych.

Zdefiniujmy relację $\rho \subset \mathcal{F} \times \mathcal{F}$, gdzie \mathcal{F} to zbiór wszystkich SLR. Niech $\rho = \{(A, B) \in \mathcal{F} \times \mathcal{F} : Compare(A, B) = 1\}$. Dla dowolnych skierowanych liczb rozmytych $A, B \in \mathcal{F}$, algorytm 8 opisujący funkcję $Compare(A, B)$ zwraca jedną z trzech możliwości:

$$Compare(A, B) = \begin{cases} 1 & \iff A \prec B, \\ -1 & \iff B \prec A, \\ 0 & \iff A = B. \end{cases} \quad (6.97)$$

Oznacza to, że według tego algorytmu, zawsze zachodzi $\rho(A, B) \vee \rho(B, A) \vee A = B$. Wprowadźmy relację pomocniczą $\preceq = \rho \cup \{(E, E) : E \in \mathcal{F}\}$ dla której zachodzi:

$$\preceq (A, B) \vee \preceq (B, A), \quad (6.98)$$

wobec czego para (\mathcal{F}, \preceq) tworzy porządek liniowy, czyli dowolne dwa elementy ze zbioru \mathcal{F} można porównać korzystając z relacji \preceq .

Algorytm 8 realizuje ideę porządku leksykograficznego (dokładnie typu *The Lexical 1* z [120]) - liczba, która zaczyna się wcześniej (jej funkcja przynależności dla mniejszych argumentów zaczyna przyjmować niezerowe wartości) jest mniejsza; dodatkowo uzupełniono algorytm o uwzględnienie skierowania właściwego dla SLR. W przypadku weryfikacji, czy $(T, A) \in \preceq$ (równoważnie, czy $T \preceq A$) jest to pożądane zachowanie, ponieważ sens wyznaczania ostrych poziomów granicznych polega na tym, aby liczby poniżej wartości granicznej nie przechodziły pozytywnie tego testu. Dostajemy dzięki temu jasną informację, że poziom minimalny został przez daną liczbę rozmytą przekroczony. Dodatkowo, możliwe jest definiowanie maksymalnego poziomu unormowanej metryki rozmycia zadanej wzorem (5.20), aby zapewnić, że wyznaczona SLR ma niski poziom rozmycia.

Rozpatrzmy teraz sytuację, że chcielibyśmy dokonywać sprawdzenia, czy $(A, T) \in \preceq$ (równoważnie, czy $A \preceq T$). Okazuje się, że taki test wykonany za pomocą algorytmu 8 nie będzie zwracać wyników w sposób analogiczny do porównania $T \preceq A$ - przykładowo SLR z nośnikiem $[0, 1; 0, 9]$ będzie należała do relacji \preceq wraz z dowolnym $T \in [0, 1; 0, 9]$.

Algorytm 2 Algorytm wyznaczania izolacji z wykorzystaniem SLR.

```
1: procedure CALCULATEISOLATIONSLR
2:   Wyznacz zbiór wierzchołków podstawowych  $V = \{V_1, V_2, \dots, V_N\}$  dla danego strumienia danych.
3:   for  $V_i \in V$  do
4:     Wyznacz zbiór cech  $\Pi_i = \{\pi_{i,1}, \pi_{i,2}, \dots, \pi_{i,M_i}\}$  właściwych dla  $V_i$ .
5:     Ustal funkcje normalizujące  $(g_{i,1}, g_{i,2}, \dots, g_{i,M_i})$  dla cech  $(\pi_{i,1}, \pi_{i,2}, \dots, \pi_{i,M_i})$  przynależnych do  $V_i$ .
6:     Ustal funkcje  $F_1, F_2, \dots, F_{|\Pi|}$  łączące izolację dla poszczególnych cech ze zbioru  $\Pi = \cup_{i=1}^N \Pi_i = \{\pi_1, \pi_2, \dots, \pi_{|\Pi|}\}$ .
7:     for  $V_i \in V$  do
8:       for  $j = 1, 2, \dots, M_i$  do
9:         Zbierz wartości  $(x_1, x_2, \dots, x_k)$  cechy  $\pi_{i,j}$  dla wierzchołka  $V_i$ .
10:         $(x'_1, x'_2, \dots, x'_k) = (g_{i,j}(x_1), g_{i,j}(x_2), \dots, g_{i,j}(x_k))$  - normalizacja wartości cech, korzystając z funkcji normalizujących ustalonych w kroku 4.
11:        hits,  $x_{min}$ ,  $x_{max}$ , m, left, right = CreateHistogram( $x'_1, x'_2, \dots, x'_k$ )
12:        order = CalculateOrder( $x'_1, x'_2, \dots, x'_k$ )
13:         $slr[V_i, \pi_{i,j}] = \text{CreateSLR}(\text{hits}, \text{left}, \text{right}, \text{order})$  // wyznacz SLR dla danej cechy danego węzła
14:     for  $G_s \in \{G_1, G_2, \dots, G_S\}$  do
15:       for  $\pi_i \in \Pi$  do
16:          $W = \{V_u \in V : \pi_i \in \Pi_u\}$  //wierzchołki z  $G_s$  z daną cechą
17:         tmp = W.Select(  $v \Rightarrow slr[v, \pi_i]$  )
18:         weights = repeat(1.0 / len(tmp), len(tmp)) //wektor takich samych wag sumujących się do 1
19:          $r_{i,s} = \text{MergeFunction}(\text{tmp}, \text{weights})$ 
20:         mergedResults[ $\pi_i$ ].Add( $r_{i,s}$ )
21:         globalWeights[ $\pi_i$ ].Add(len(tmp)) //dodaje skalar do listy wag
22:     for  $\pi_i \in \Pi$  do
23:       weights = globalWeights[ $\pi_i$ ].Select(  $v \Rightarrow v/\text{globalWeights}[\pi_i].\text{Sum}()$  ) // wagi muszą się sumować do 1
24:       isolationVector[i] = MergeFunction(mergedResults[ $\pi_i$ ], weights)
25:   finalIsolation = CalculateFinalIsolation(isolationVector)
26:   finalResult = FinalIsolationWithConvexForm(finalIsolation) //Wyznacz SLR będącą właściwą SLR, która jest najbliższa SLR otrzymanej w punkcie poprzednim.
```

Algorytm 3 Algorytm tworzenia unormowanego i skorygowanego histogramu

```
1: procedure CREATEHISTOGRAM( $x'_1, x'_2, \dots, x'_k$ )
2:    $x_{min} = \min(x'_1, x'_2, \dots, x'_k)$ 
3:    $x_{max} = \max(x'_1, x'_2, \dots, x'_k)$ 
4:    $m = \text{floor}(\sqrt{k})$  //liczba kolumn histogramu
5:    $size = (x_{max} - x_{min})/m$ 
6:   for  $x_i \in X$  do
7:      $p = \text{floor}((x_i - x_{min})/size)$ 
8:     if  $p == m$  then
9:        $p = m - 1$  //prawy koniec przedziału należy do ostatniej kolumny histogramu
10:     $hits[p] += 1$ 
11:    $h_{max} = \max(hits)$ 
12:   for  $i = 0, 1, \dots, m - 1$  do
13:      $hits[i] /= h_{max}$  //normalizacja histogramu
14:    $left = 0$ 
15:    $right = m - 1$ 
16:    $notFound = \text{true}$ 
17:   for  $i = 0; i < m$  AND  $notFound$ ;  $++i$  do
18:     if  $hits[i] == 1$  then
19:        $left = i$ 
20:        $notFound = \text{false}$ 
21:    $notFound = \text{true}$ 
22:   for  $i = m - 1; i \geq 0$  AND  $notFound$ ;  $--i$  do
23:     if  $hits[i] == 1$  then
24:        $right = i$ 
25:        $notFound = \text{false}$ 
26:   for  $i = 1, 2, \dots, left - 1$  do // lewa strona niemalejąca
27:     if  $hits[i] < hits[i - 1]$  then
28:        $hits[i] = hits[i - 1]$ 
29:   for  $i = left + 1, left + 2, \dots, right - 1$  do
30:      $hits[i] = 1$  // odcinek o  $\mu(x) = 1$ 
31:   for  $i = m - 2 \dots right + 1$  do // prawa strona nierosnąca
32:     if  $hits[i] < hits[i + 1]$  then
33:        $hits[i] = hits[i + 1]$ 
34:   return  $hits, x_{min}, x_{max}, m, left, right$ 
```

Algorytm 4 Algorytm wyznaczania skierowania na podstawie wyników pomiarów

```
1: procedure CALCULATEORDER( $x'_1, x'_2, \dots, x'_k$ )
2:    $res = 0$ 
3:    $mx = \text{mean}(x'_1, x'_2, \dots, x'_k)$ 
4:   for  $i = 1, 2, \dots, k$  do
5:      $res += (i - (k + 1)/2) * (x'_i - mx)$ 
6:   return  $\text{sgn}(res)$ 
```

Algorytm 5 Algorytm tworzenia SLR na podstawie histogramu

```
1: procedure CREATESLR(hits, left, right, order)
2:   resLeft = hits.Take(left+1).Select(h => (h.left, h.value)) // left jest liczony od 0, bierzemy też wartość dla  $\mu(x) = 1$ 
3:   resLeft += hits.Take(left).Select(h => (h.right, h.value))
4:   resLeft += (0, hits[0].left) // dodajemy punkt z  $\mu(x) = 0$ , aby mieć pokryty przedział [0;1]
5:   resLeft.Distinct().OrderBy(r => r.Value1).ThenBy(r => r.Value2) // najpierw sortowanie po pierwszej kolumnie, a potem po drugiej
6:   resRight = hits.Skip(right+1).Select(h => (h.left, h.value))
7:   resRight += hits.Skip(right).Select(h => (h.right, h.value))
8:   resRight += (0, hits[0].right) // dodajemy punkt z  $\mu(x) = 0$ , aby mieć pokryty przedział [0;1]
9:   resRight.Distinct().OrderBy(r => r.Value1).ThenBy(r => r.Value2) // najpierw sortowanie po pierwszej kolumnie, a potem po drugiej
10:  if order > 0 then
11:    return f = resLeft, g = resRight, order // f, g
12:  else
13:    return f = resRight, g = resLeft, order // f, g
```

Algorytm 6 Przykładowa funkcja łącząca

```
1: procedure MERGEFUNCTION(W, weights)
2:   q=1 //parametr zależny od cechy, tutaj przykładowa wartość. Dla q = 0 wymagana jest osobna implementacja - średnia geometryczna
3:   res = ((0, 1), (0, 1)) //zero w SLR
4:   for i = 0; i < len(W); ++ i do
5:     res = res + pow(W[i], q) * weights[i]
6:   return pow(res, 1/q)
```

Algorytm 7 Przykładowa funkcja wyliczająca izolację końcową - 4.24

```
1: procedure CALCULATEFINALISOLATION(W)
2:   q=2 //parametr zależny od cechy, tutaj przykładowa wartość. Dla q = 0 wymagana jest osobna implementacja - średnia geometryczna
3:   N = len(W)
4:   res = ((0, 1), (0, 1)) //0 jako SLR
5:   oneSLR = ((1, 1), (1, 1)) //1 jako SLR
6:   for i = 0; i < N; ++ i do
7:     res = res + pow(oneSLR - W[i], q)
8:   res = res / N
9:   return oneSLR - pow(res, 1/q)
```

Algorytm 8 Algorytm określania relacji porządku pomiędzy skierowanymi liczbami rozmytymi.

```

1: procedure COMPAREOFN( $A = ((a_1, v_1), (a_2, v_2), \dots, (a_n, v_n)), B =$ 
   ( $(b_1, w_1), (b_2, w_2), \dots, (b_m, w_m)$ ))
2:    $A = ((a_1, v_1), (a_2, v_2), \dots, (a_n, v_n))$  //  $a_1 \leq a_2 \leq \dots \leq a_n$ 
3:    $B = ((b_1, w_1), (b_2, w_2), \dots, (b_m, w_m))$  //  $b_1 \leq b_2 \leq \dots \leq b_m$ 
4:    $C = \{a_1, a_2, \dots, a_n\} \cup \{b_1, b_2, \dots, b_m\} = \{c_1, c_2, \dots, c_k\}$ 
5:    $A_R = A.First( (a, b) \Rightarrow b == 1 ).index$  // indeks krotki w ciagu, liczony od 1
6:    $B_R = B.First( (a, b) \Rightarrow b == 1 ).index$ 
7:    $R = \min(A_R, B_R)$ 
8:   for  $i = 1, 2, \dots, R$  do
9:     if  $A(c_i) < B(c_i)$  then //wartość  $A$  w punkcie  $c_i$ 
10:      return -1 //  $B \prec A$ 
11:     if  $A(c_i) > B(c_i)$  then
12:      return 1 //  $A \prec B$ 
13:   for  $i = R + 1, R + 2, \dots, m$  do
14:     if  $A(c_i) < B(c_i)$  then
15:      return 1 //  $A \prec B$ 
16:     if  $A(c_i) > B(c_i)$  then
17:      return -1 //  $B \prec A$ 
18:   if  $Ord(A) < Ord(B)$  then //operator skierowania, bądź też informacja o wyznaczonym
   skierowaniu
19:     return 1 //  $A \prec B$ 
20:   if  $Ord(A) > Ord(B)$  then
21:     return -1 //  $B \prec A$ 
22:   return 0 //  $A = B$ 

```

7. Przykład obliczeniowy

Przedstawmy teraz przykładowe obliczenia prowadzące do wyznaczenia poziomu izolacji dla pewnej prostej sieci plastrowej. Przykład podzielimy na dwa kroki: stworzenie SLR i wykonanie algorytmu wyznaczania izolacji z obliczonymi skierowanymi liczbami rozmytymi.

Załóżmy, że mamy sieć, w której zdefiniowano następujące cechy, przypisane wierzchołkom hipergrafu przedstawionego na rysunku 7.1. Sieć ta została podzielona na dwie podsieci, przedstawione na rysunku 7.2, zostały również do poszczególnych węzłów sieci przypisane zidentyfikowane cechy związane z izolacją. Wierzchołek gNB należy do obydwu podsieci.

Cechy przypisane do poszczególnych wierzchołków są zdefiniowane w tabeli 7.1 wraz z funkcjami normalizującymi i zestawem danych liczbowych określających kontekst danej cechy w danym węźle. Przykładowe wyniki pomiarowe zostały zebrane w tabeli 7.2, a ich znormalizowane wartości zostały zebrane w tabeli 7.3.

Rysunki 7.3 - 7.10 przedstawiają stworzone histogramy i odpowiadające im skierowane liczby rozmyte. Przyjęto, że liczba kolumn histogramu jest równa pierwiastkowi kwadratowemu z liczby próbek ($4 = \sqrt{16}$), co pozwala zachować balans rozdzielczości w obydwu osiach histogramu. Histogram budowany jest każdorazowo na przedziale pomiędzy najmniejszą i największą wartością próbki, aby rozdzielczość w osi OX była zmaksymalizowana. Przedziały w ramach histogramu są równe. Funkcja f danej liczby rozmytej jest wykreślona kolorem czerwonym, funkcja g kolorem niebieskim.

Warto w tym miejscu odnotować jedną istotną uwagę, dotyczącą zagadnienia, czy f i g są *funkcjami* w klasycznym matematycznym rozumieniu. *Funkcja* jest przyporządkowaniem dla każdego elementu ze zbioru Q elementu ze zbioru R , gdzie Q i R to dowolne zbiory. W przypadku standardowych funkcji, dla których $Q = R = \mathbb{R}$, jest problem z interpretacją poziomych i pionowych odcinków obwiedni histogramu - przyporządkowanie nie jest tutaj jednoznaczne (poziomych - dla f i g , pionowych dla hipotetycznych f^{-1} i g^{-1}). Można interpretować za to f i g jako funkcje typu $\mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}$, gdzie dla liczby rzeczywistej przypisujemy odcinek, być może zdegenerowany do punktu. Powstaje wtedy problem określenia f^{-1} i g^{-1} , które tworzą funkcję

przynależności, która jest funkcją $\mu(x) : \mathbb{R} \rightarrow \mathbb{R}$. Tworzy to dwuwymiarowe rozmycie, gdzie nie jest określona w sposób *ostry* wartość liczby (dlatego jest rozmyta), ale i przynależność danej wartości do liczby (rozmytej). Pomocna tutaj jest interpretacja kontekstowa - wiemy, że SLR została zbudowana na bazie histogramu, a więc wartość funkcji przynależności w danym punkcie powinna odpowiadać poziomowi wynikającemu z histogramu. W przypadku punktów należących do obydwu kolumn histogramu (narożników obwiedni), należy przyjąć konwencję, do której kolumny on należy. Przyjmujemy konwencję, że do kolumny histogramu należy jego wartość minimalna. Punkt dla którego $x = 1$ należy do ostatniej kolumny histogramu.

Rysunki 7.14 i 7.15 przedstawiają ciekawe przypadki, jakie mogą wystąpić po wykonaniu operacji łączenia SLR. Dla przejrzystości przykładu do łączenia liczb zastosowano średnią arytmetyczną. Widzimy, że nie dla każdej wartości x mamy jednoznaczną wartość funkcji przynależności μ - są to niewłaściwe skierowane liczby rozmyte. Używamy konwencji nazewniczej $T3||T6||T8$, która oznacza, że dana liczba rozmyta przedstawia cechę scaloną z SLR wyznaczonych dla cech o indeksach T3, T6 i T8. Tabela 7.4 zawiera informacje w jaki sposób dokonano złączenia wartości poszczególnych cech w ramach danych podsieci (RAN i CN). Do operacji złączenia użyto funkcji łączącej (4.11) z parametrem $q = 1$ (średnia arytmetyczna) i równymi wagami. Wybór parametru był podyktowany przejrzystością przykładu obliczeniowego. Tabela 7.5 zawiera sposób uzyskania wektora izolacji z izolacji wyznaczonych dla poszczególnych podsieci. Wagi dobrane są w sposób taki, aby zrównoważyć wpływ pojedynczych cech wśród różnych węzłów; w takim przypadku węzeł uwspólniony (cecha T3) będzie miał większy udział niż pozostałe cechy z tej samej rodziny cech. Poszczególne SLR tworzące wektor izolacji zostały przedstawione na rysunkach 7.17 - 7.20.

Rysunek 7.21 przedstawia skierowaną liczbę rozmytą będącą wynikiem przekształcenia (4.24), dla $q = 2$ i $N = 4$. Jest to niewłaściwa skierowana liczba rozmyta, a kolejnym krokiem jest odnalezienie właściwej skierowanej liczby rozmytej, która jest możliwie jak najbliższa (w sensie zadanej metryki) zadanej niewłaściwej SLR. W tym celu należy określić skierowanie niewłaściwej SLR, aby móc stworzyć problem optymalizacyjny przedstawiony w podrozdziale 6.2.. Można do problemu wyznaczania skierowania niewłaściwej SLR podejść na kilka sposobów.

- Skorzystać z operatora wyostrzania SLR i wyznaczyć skierowanie na podstawie wzoru (5.4). Niestety, dobór operatora wyostrzania SLR jest w gruncie rzeczy arbitralny, więc nie jest to zupełne rozwiązanie problemu.
- Wyznaczyć trend na podstawie znanych punktów niewłaściwej SLR, podobnie jak w przypadku wyznaczania orientacji stworzonej SLR z szeregu czasowego (pomiarowego). Wadą tego rozwiązania jest nieuwzględnianie faktycznej kolejności funkcji f i g względem siebie.
- Skorzystać z faktu, że wiemy, która część wykresu SLR należy do funkcji f , a która do funkcji g i próbować zastosować definicję 5.4.

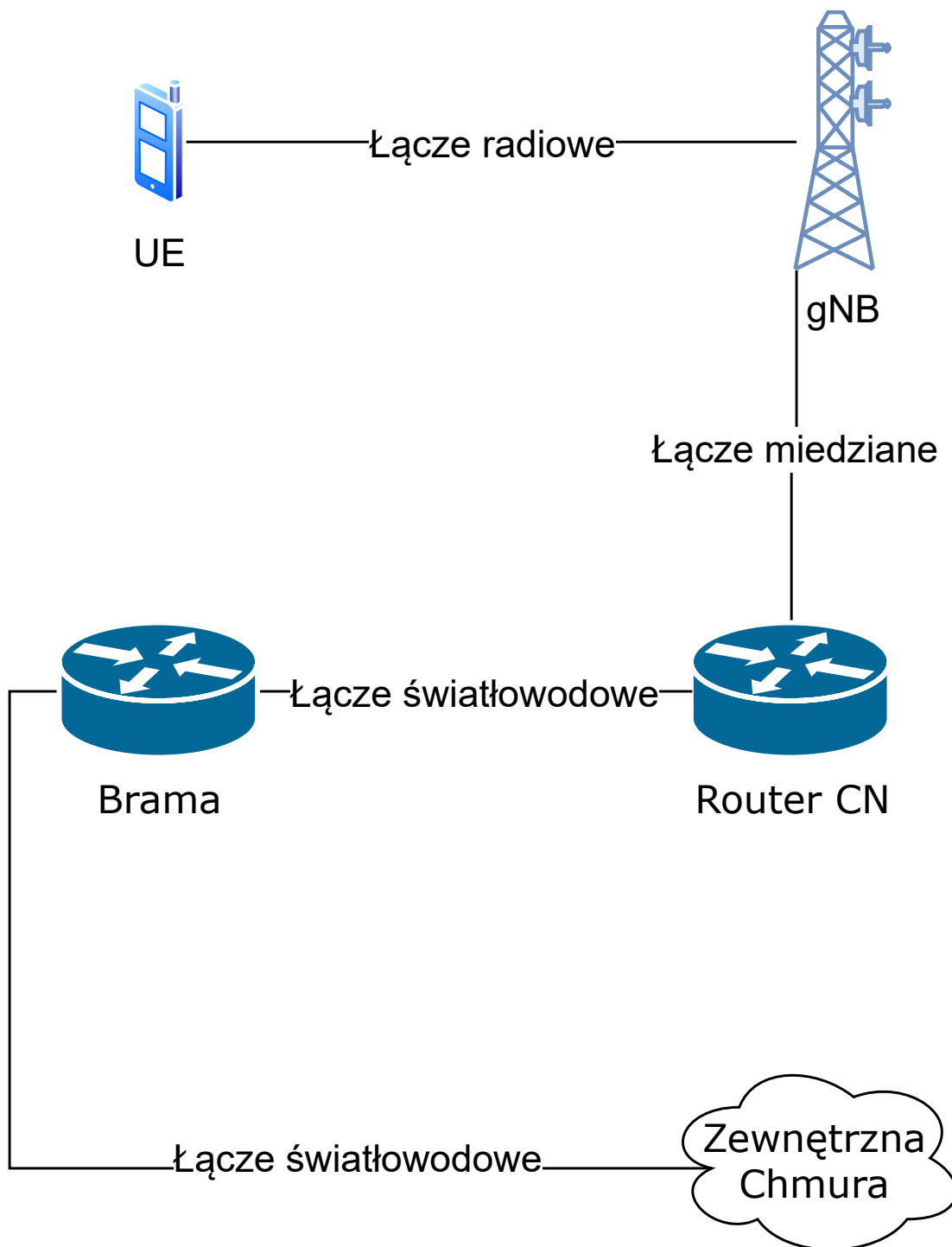
W niniejszej pracy będziemy wykorzystywać ostatnie podejście, które przedstawiono poprzez algorytm 9. W zaprezentowanym podejściu określanie skierowania jest wymagane na początkowym etapie - przy przechodzeniu z dziedziny serii pomiarowych do dziedziny skierowanych liczb rozmytych, oraz na etapie poszukiwania najbliższej SLR wypukłej. Podczas wykonywania dwuargumentowych działań na SLR nie ma potrzeby znajomości skierowania liczb i określania skierowania wyniku operacji, wystarczy znajomość funkcji f i g .

Algorytm 9 wyznacza skierowanie liczby rozmytej korzystając z informacji, jak zachowuje się funkcja f i g w sąsiedztwie jądra liczby rozmytej (przedziału, w którym funkcja przynależności jest równa 1). Celem było użycie tego fragmentu nośnika SLR, do którego dana SLR przynależy w największym stopniu, a więc powinien on wyznaczać kluczowe właściwości tej SLR.

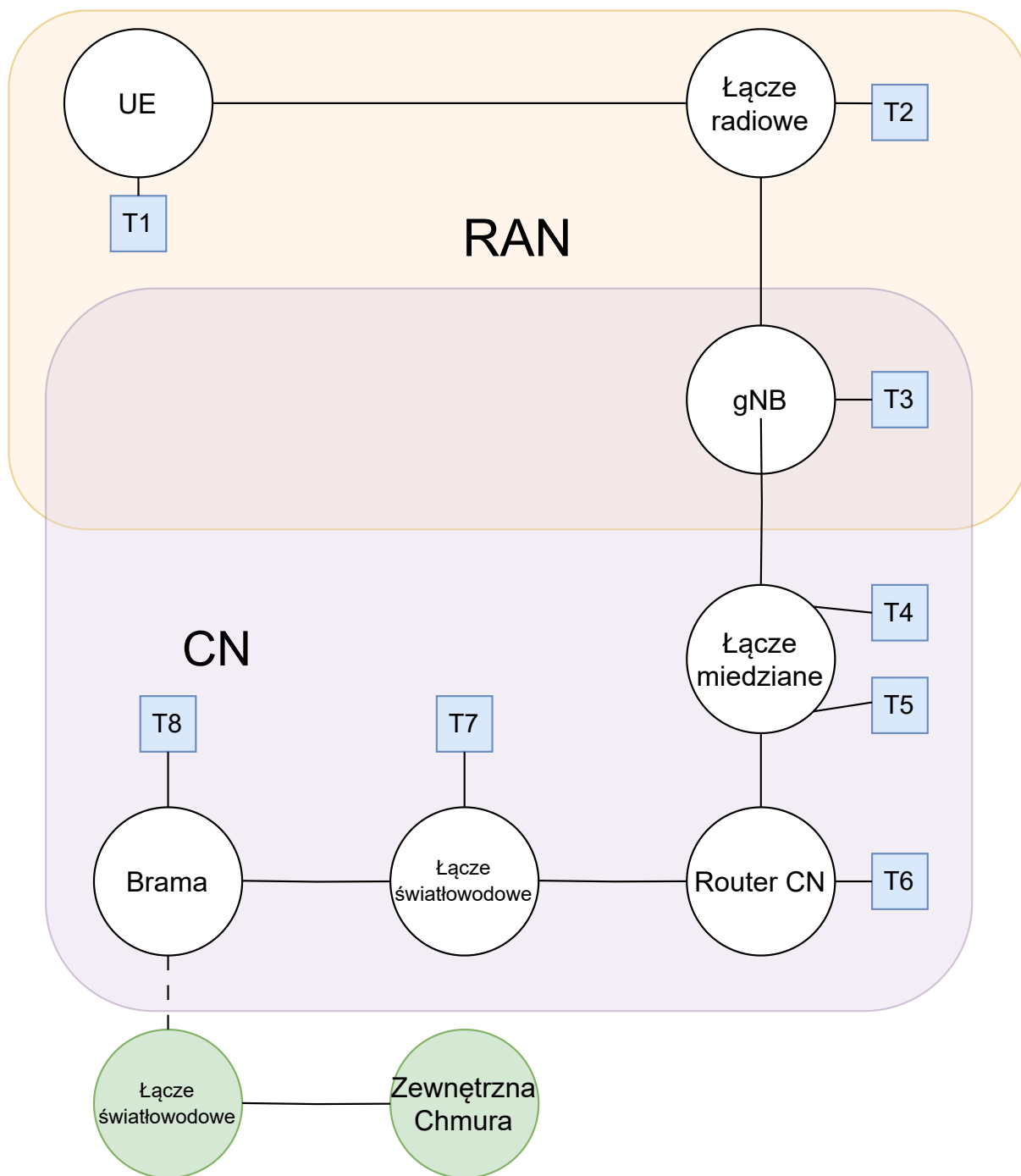
Rysunek 7.22 przedstawia rezultat rozwiązania problemu optymalizacji zadanego wzorami (6.88) i (6.89). Wykorzystano w tym celu algorytm Simplex wykonywany przez program Octave, który dostarcza implementacji tego algorytmu. W niniejszym przykładzie dokonano najpierw optymalizacji funkcji f (rezultat zaznaczony kolorem czarnym), a następnie funkcji g (linia koloru fuksji), korzystając z otrzymanego rozwiązania. Właściwa liczba rozmyta zachowuje skierowanie określone dla niewłaściwej liczby rozmytej.

Algorytm 9 Algorytm określania skierowania niewłaściwej SLR

```
1: procedure CALCULATEORDER( $A = (f_A, g_A)$ )
2:    $f_A = ((a_1, v_1), (a_2, v_2), \dots, (a_n, v_n))$  //  $0 \leq a_1 \leq a_2 \leq \dots < a_n = 1$ 
3:    $g_A = ((b_1, w_1), (b_2, w_2), \dots, (b_m, w_m))$  //  $0 \leq b_1 \leq b_2 \leq \dots < b_m = 1$ 
4:   // SLR były unormowane, więc funkcje  $f_A$  i  $g_A$  miały zdefiniowane wartości dla  $x = 1$ 
5:   if  $v_n < w_m$  then
6:     return  $Ord(A) = 1$ 
7:   if  $v_n > w_m$  then
8:     return  $Ord(A) = -1$ 
9:   //SLR ma  $\mu_A(x) = 1$  dla dokładnie jednego punktu
10:   $(a, v) = f_A.Last((a, v) \Rightarrow v \neq v_n)$  // ostatni element z  $f_A$ , który na drugiej pozycji krotki
    spełnia  $v \neq v_n$ 
11:   $(b, w) = f_B.Last((b, w) \Rightarrow w \neq w_m)$ 
12:  if  $v == \text{null}$  AND  $w == \text{null}$  then // nie znaleziono takich elementów, jest to liczba ostra
13:    return  $Ord(A) = 0$ 
14:  if  $v == \text{null}$  then // nie znaleziono elementu dla funkcji  $f_A$ 
15:     $v = v_n$ 
16:  if  $w == \text{null}$  then // nie znaleziono elementu dla funkcji  $g_A$ 
17:     $w = w_m$ 
18:  if  $v < w$  then
19:    return  $Ord(A) = 1$ 
20:  if  $v > w$  then
21:    return  $Ord(A) = -1$ 
```



Rysunek 7.1: Przykładowa sieć do obliczeń, schemat zgodny z przykładem podanym w pracy [57].



Rysunek 7.2: Przykładowa sieć do obliczeń - zestaw węzłów podstawowych w sieciach RAN i CN, schemat zgodny z przykładem podanym w pracy [57].

Tablica 7.1: Zbiór funkcji normalizujących i znormalizowanych wartości, zgodnie z przykładem podanym w pracy [57]

Część	Element	Cecha	Indeks	Funkcja normalizująca	Wartość typowa / skrajna
RAN	UE w RAN	Zastosowany język programowania	T1	n/d	n/d
	Łącze radiowe w RAN	Siła użytego szyfrowania symetrycznego	T2	$g(x) = \frac{x-a}{b-a}$	$\begin{cases} a = 0 \\ b = 256bit \end{cases}$
RAN + CN	gNB	Średni czas pomiędzy analizami podatności	T3	$g(x) = 2^{-\frac{x}{\beta}}$	$\beta = 4h$
CN	Łącze miedziane w CN	Siła użytego szyfrowania symetrycznego	T4	$g(x) = \frac{x-a}{b-a}$	$\begin{cases} a = 0 \\ b = 256bit \end{cases}$
		Emisja elektromagnetyczna	T5	$g(x) = \left(1 + e^{\frac{x-\beta}{ \beta }}\right)^{-1}$	$\beta = 25dB\mu V/m$
	Router CN w CN	Średni czas pomiędzy analizami podatności	T6	$g(x) = 2^{-\frac{x}{\beta}}$	$\beta = 4h$
	Łącze światłowodowe w CN	Siła użytego szyfrowania symetrycznego	T7	$g(x) = \frac{x-a}{b-a}$	$\begin{cases} a = 0 \\ b = 256bit \end{cases}$
	Brama pomiędzy CN i zewnętrzną chmurą	Średni czas pomiędzy analizami podatności	T8	$g(x) = 2^{-\frac{x}{\beta}}$	$\beta = 4h$

Tablica 7.2: Zbiór przykładowych danych pomiarowych

Próbka	T1	T2	T3	T4	T5	T6	T7	T8
1	0,20	224	18,5	192	51	4,5	208	18,0
2	0,10	232	10,0	232	67	0,5	136	13,0
3	0,30	216	9,5	184	34	6,5	8	2,5
4	0,90	0	14,0	72	24	18,0	32	9,5
5	0,80	112	10,5	72	55	2,0	216	10,0
6	0,95	216	10,5	216	13	9,0	240	8,5
7	0,95	240	6,0	224	48	6,5	216	8,5
8	0,15	24	2,0	88	33	1,0	152	0,5
9	0,35	184	18,0	40	62	12,0	112	6,5
10	0,25	112	3,0	48	76	2,5	96	5,5
11	0,30	112	1,5	48	9	10,5	168	12,5
12	0,25	24	11,0	96	29	6,5	136	8,0
13	0,75	208	16,0	128	78	6,5	192	11,5
14	0,05	216	4,0	216	37	18,5	152	3,0
15	0,80	192	4,0	88	28	11,5	136	19,5
16	0,00	160	8,0	136	52	1,0	200	14,0

Tablica 7.3: Zbiór znormalizowanych danych pomiarowych

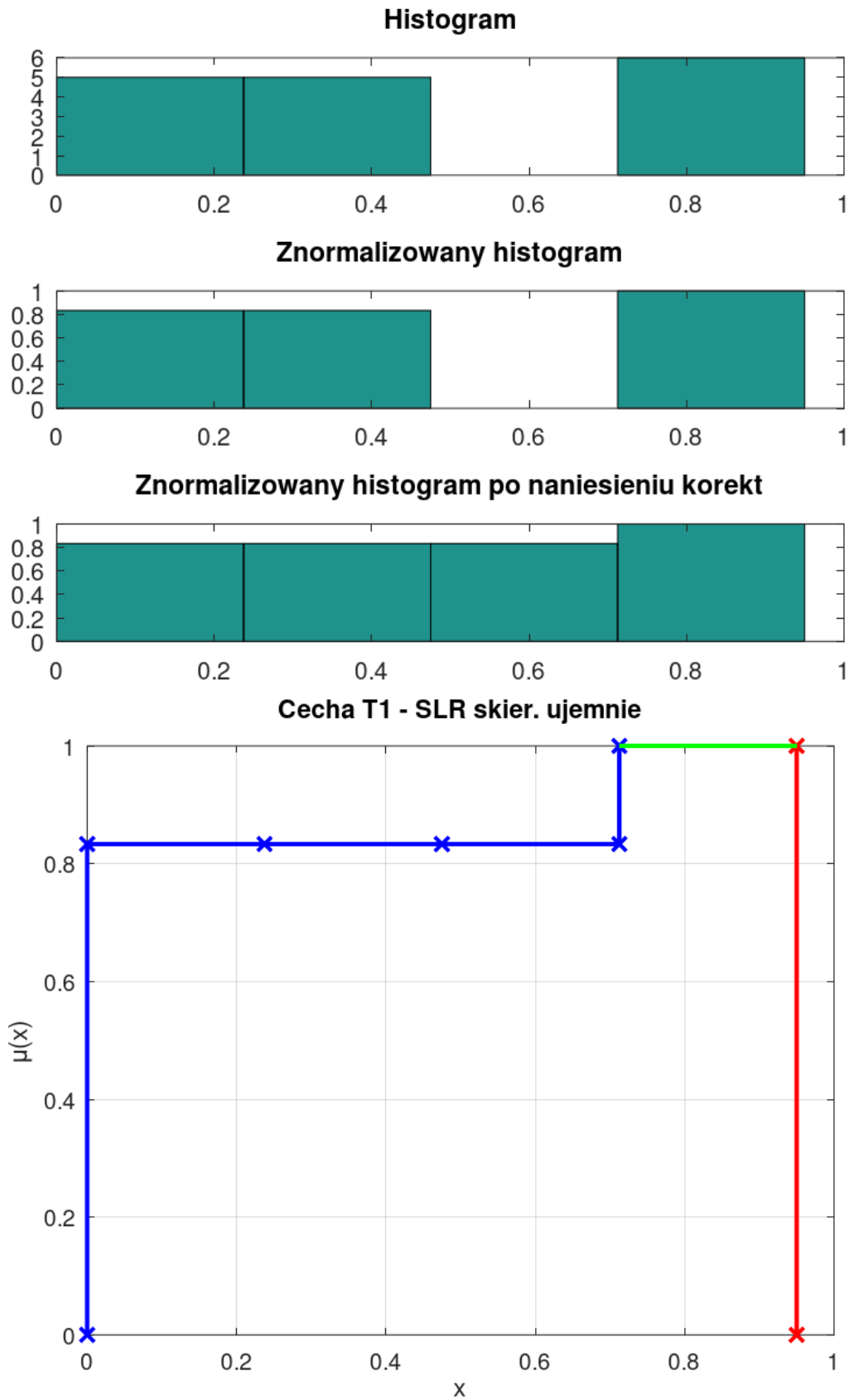
Próbka	T1	T2	T3	T4	T5	T6	T7	T8
1	0,2000	0,8750	0,0405	0,7500	0,2611	0,4585	0,8125	0,0442
2	0,1000	0,9063	0,1768	0,9063	0,1571	0,9170	0,5313	0,1051
3	0,3000	0,8438	0,1928	0,7188	0,4110	0,3242	0,0313	0,6484
4	0,9000	0,0000	0,0884	0,2813	0,5100	0,0442	0,1250	0,1928
5	0,8000	0,4375	0,1621	0,2813	0,2315	0,7071	0,8438	0,1768
6	0,9500	0,8438	0,1621	0,8438	0,6177	0,2102	0,9375	0,2293
7	0,9500	0,9375	0,3536	0,8750	0,2850	0,3242	0,8438	0,2293
8	0,1500	0,0938	0,7071	0,3438	0,4207	0,8409	0,5938	0,9170
9	0,3500	0,7188	0,0442	0,1563	0,1854	0,1250	0,4375	0,3242
10	0,2500	0,4375	0,5946	0,1875	0,1151	0,6484	0,3750	0,3856
11	0,3000	0,4375	0,7711	0,1875	0,6548	0,1621	0,6563	0,1146
12	0,2500	0,0938	0,1487	0,3750	0,4601	0,3242	0,5313	0,2500
13	0,7500	0,8125	0,0625	0,5000	0,1072	0,3242	0,7500	0,1363
14	0,0500	0,8438	0,5000	0,8438	0,3823	0,0405	0,5938	0,5946
15	0,8000	0,7500	0,5000	0,3438	0,4700	0,1363	0,5313	0,0341
16	0,0000	0,6250	0,2500	0,5313	0,2535	0,8409	0,7813	0,0884

Tablica 7.4: Izolacja wewnątrz sieci RAN i CN

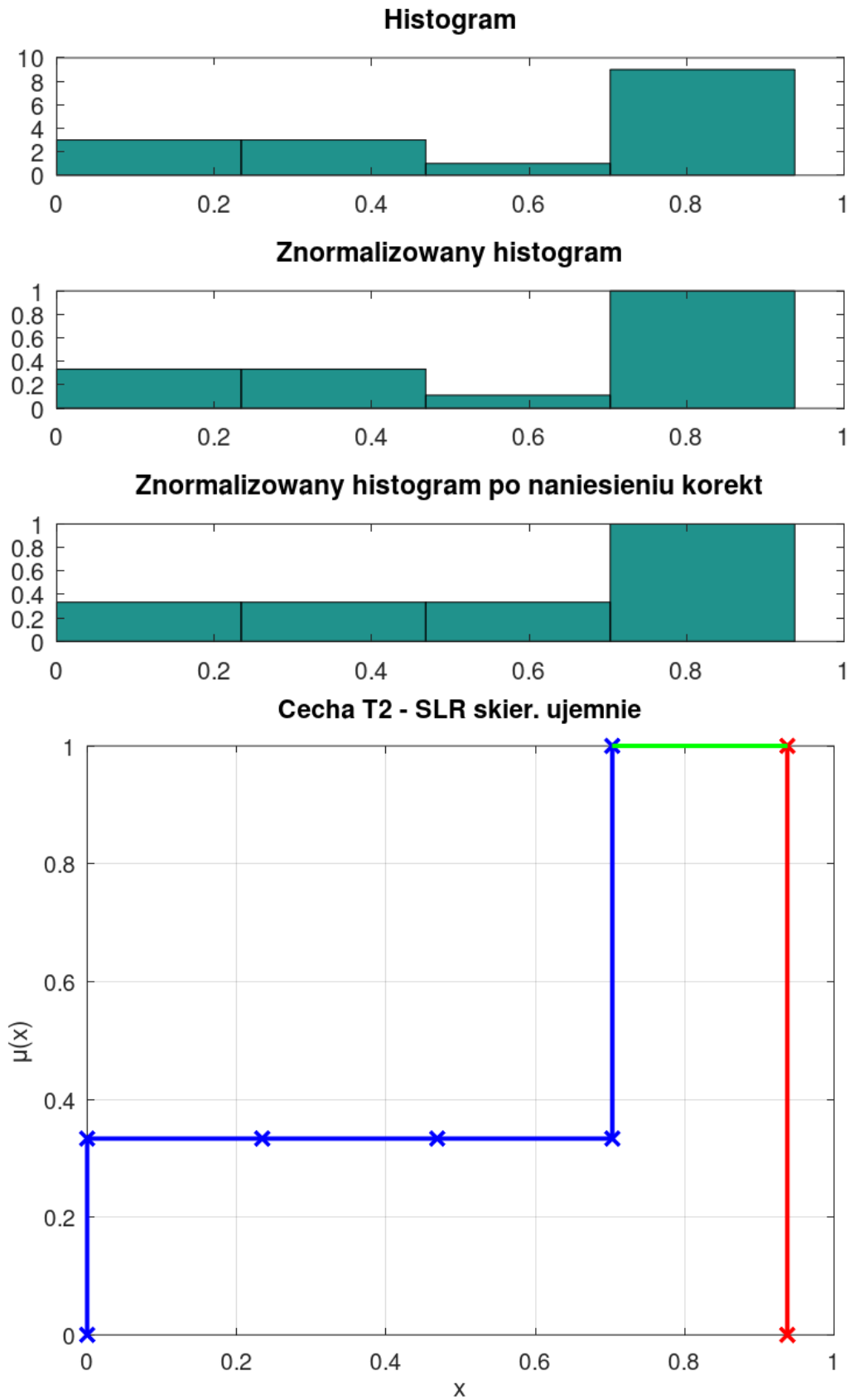
Część	Parametr	Wartości z cech	Złączone wartości
RAN	Zastosowany język programowania	T1	T1
	Siła użytego szyfrowania symetrycznego	T2	T2
	Średni czas pomiędzy analizami podatności	T3	T3
CN	Siła użytego szyfrowania symetrycznego	T4, T7	$\frac{T4+T7}{2}$
	Emisja elektromagnetyczna	T5	T5
	Średni czas pomiędzy analizami podatności	T3, T6, T8	$\frac{T3+T6+T8}{3}$

Tablica 7.5: Wektor izolacji

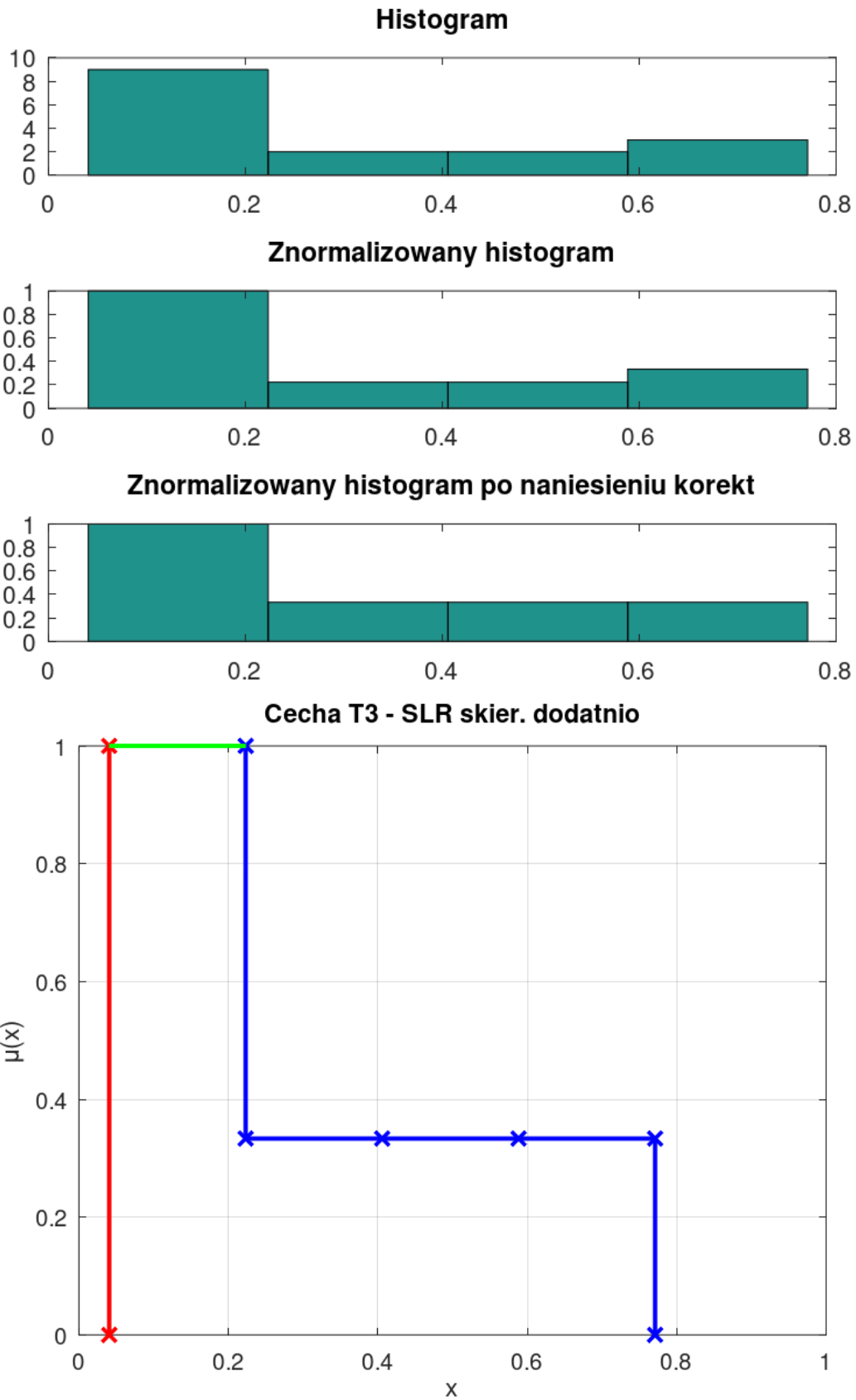
Cecha	Wartości z cech	Wagi	Wartości wektora izolacji
Zastosowany język programowania	T1	1	T1
Siła użytego szyfrowania symetrycznego	T2, T4 T7	1/3, 2/3	$\frac{1}{3}T2 + \frac{2}{3}T4 T7 = \frac{1}{3}T2 + \frac{2}{3} \cdot \frac{T4+T7}{2}$
Średni czas pomiędzy analizami podatności	T3, T3 T6 T8	1/4, 3/4	$\frac{1}{4}T3 + \frac{3}{4} \cdot \frac{T3+T6+T8}{3}$
Emisja elektromagnetyczna	T5	1	T5



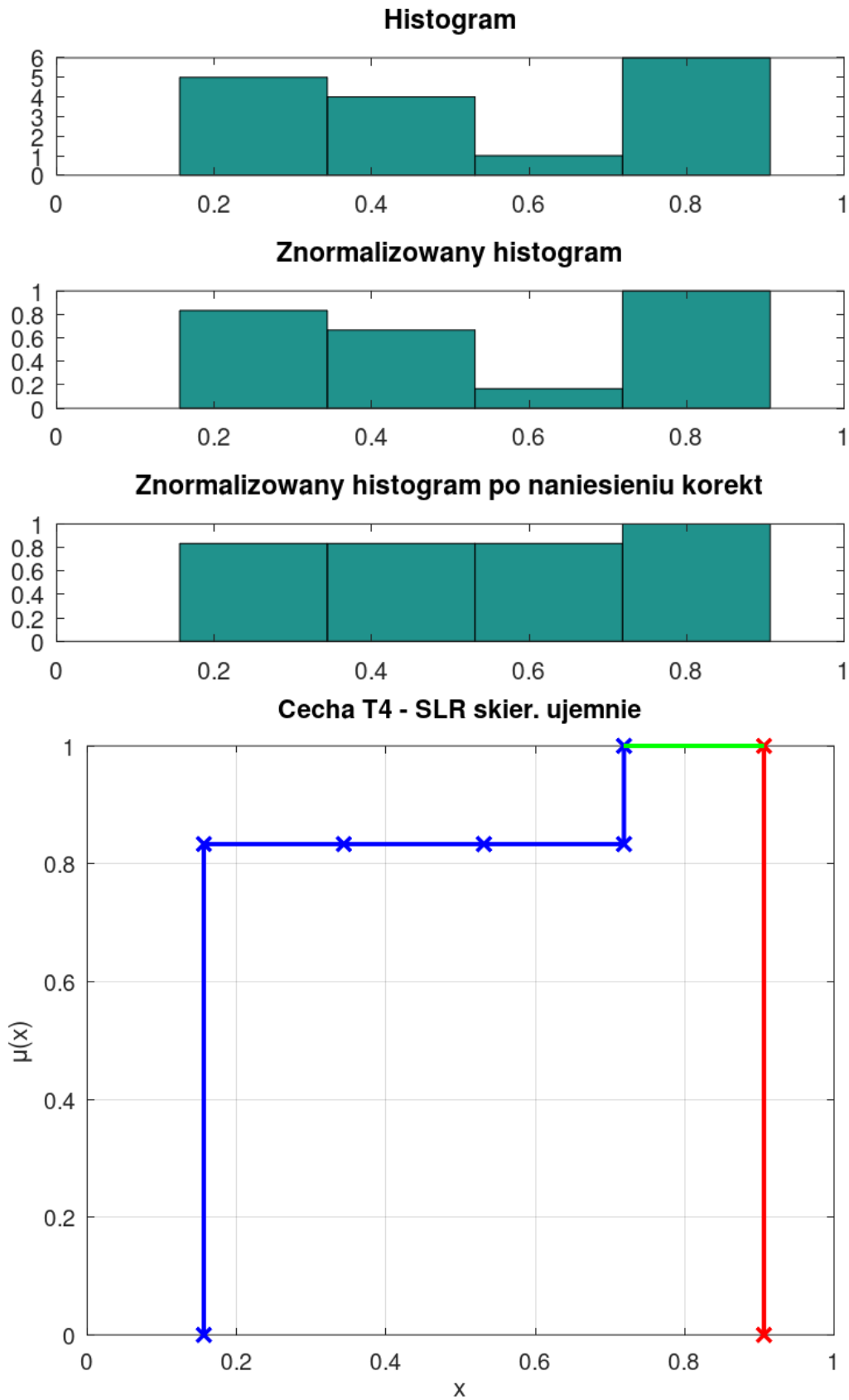
Rysunek 7.3: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T1.



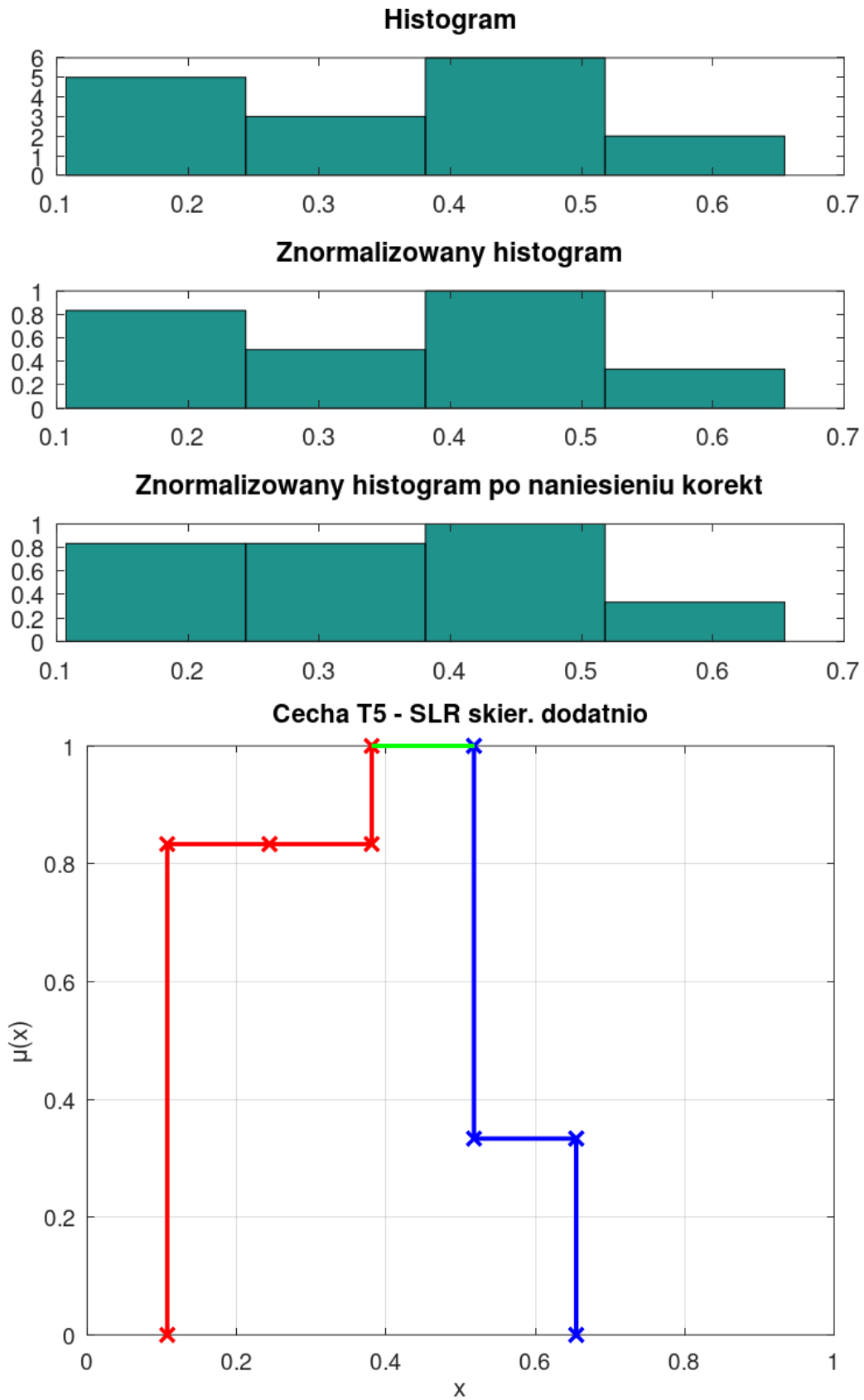
Rysunek 7.4: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T2.



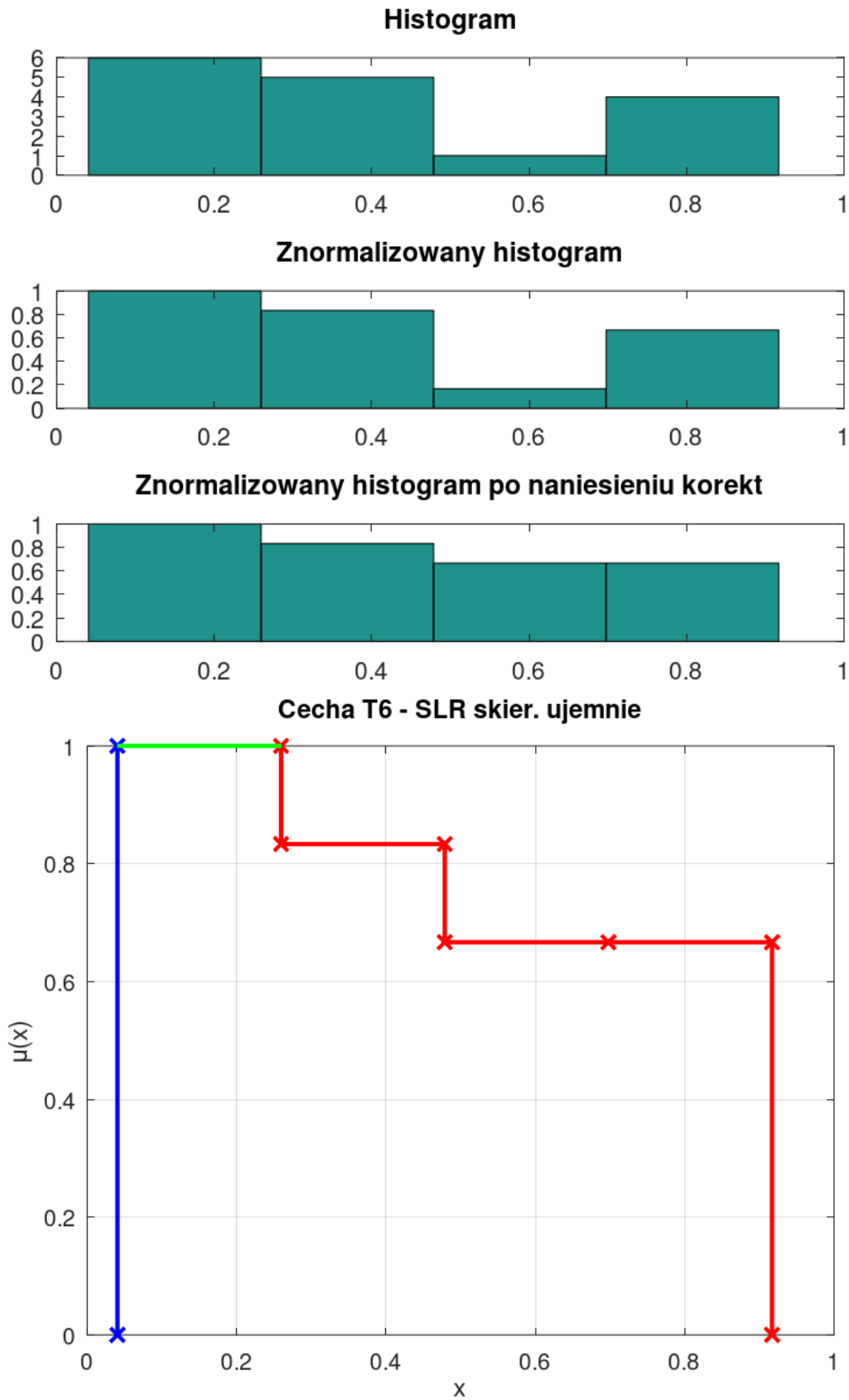
Rysunek 7.5: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T3.



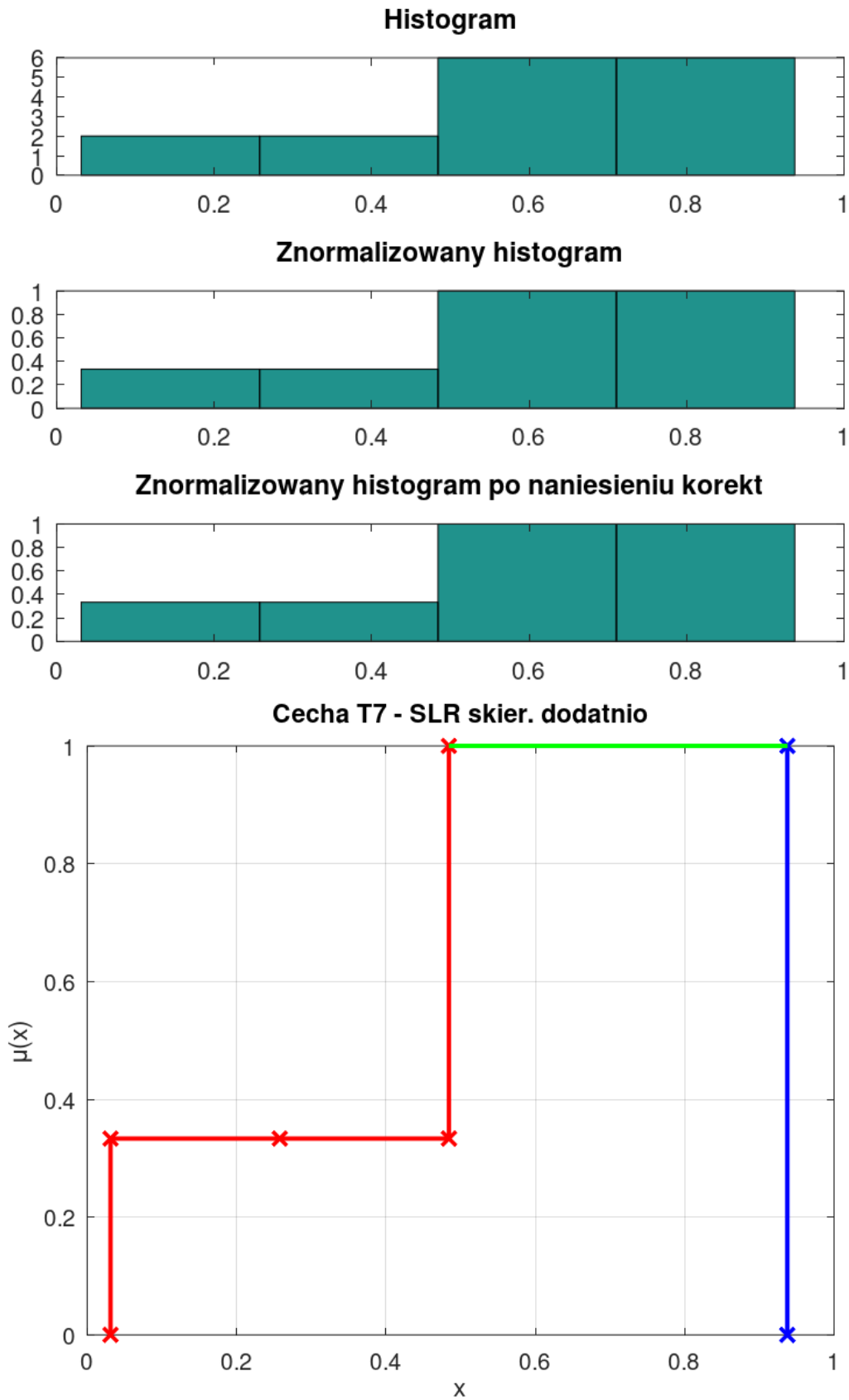
Rysunek 7.6: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T4.



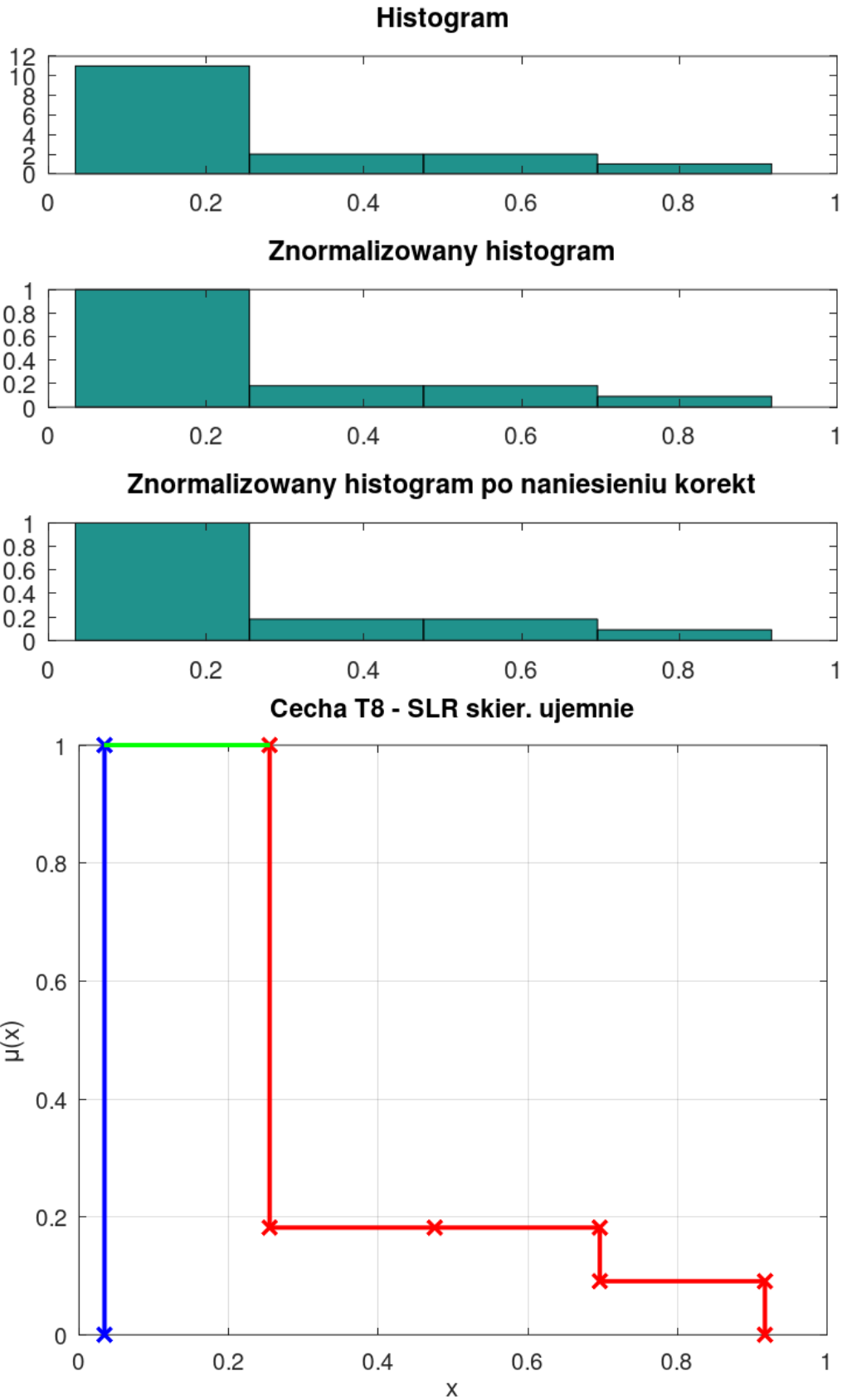
Rysunek 7.7: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T5.



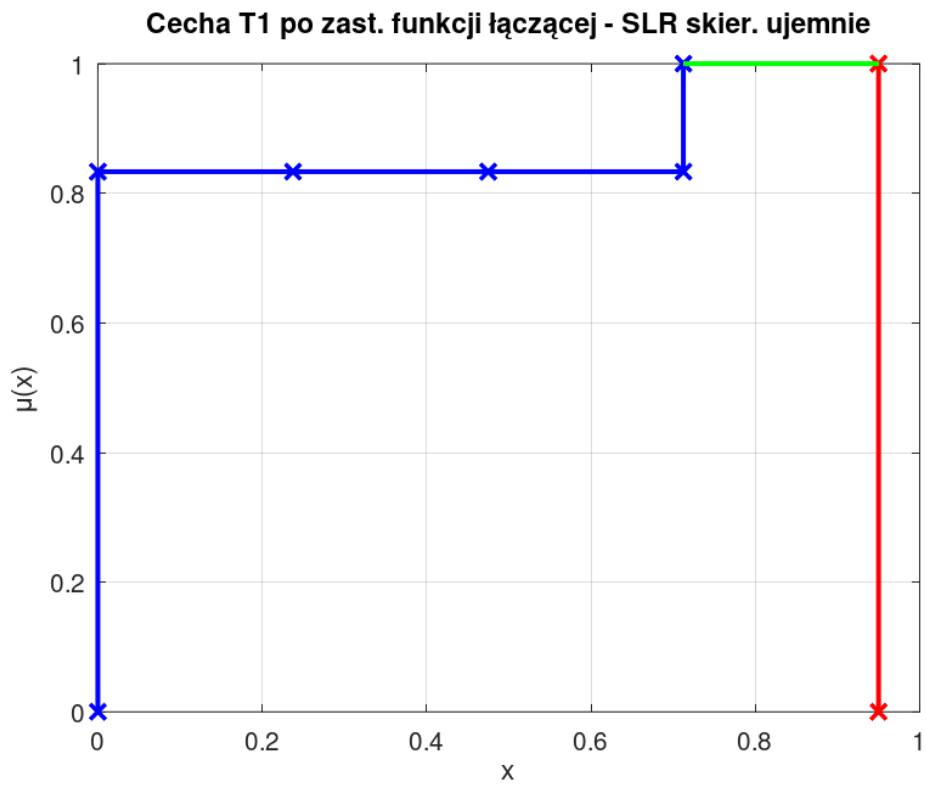
Rysunek 7.8: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T6.



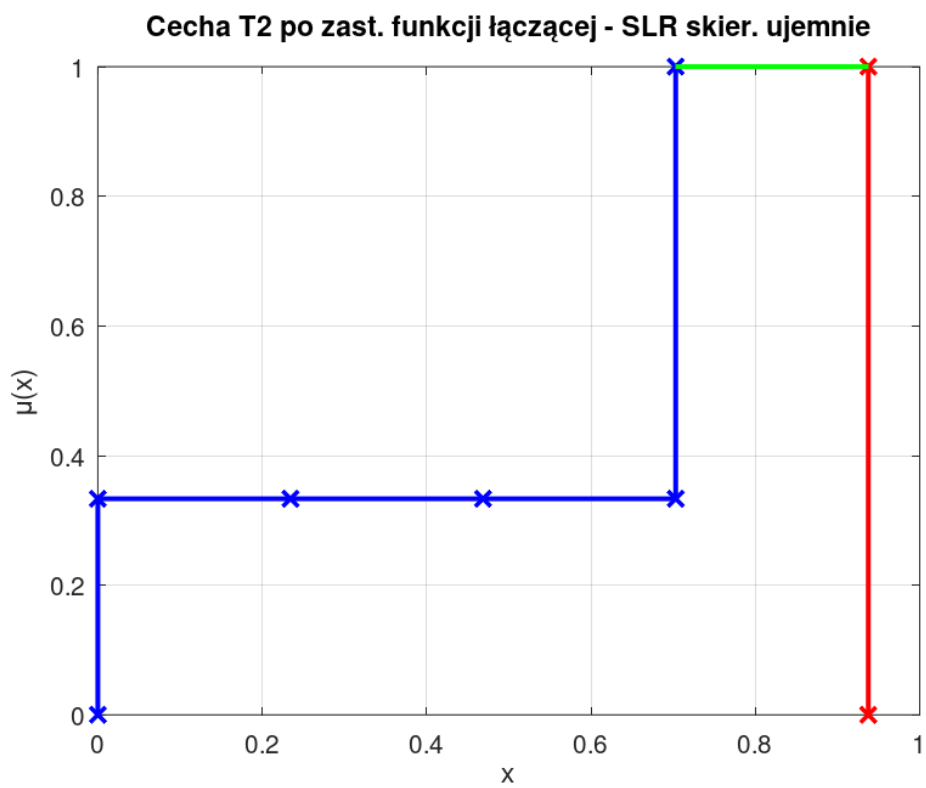
Rysunek 7.9: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T7.



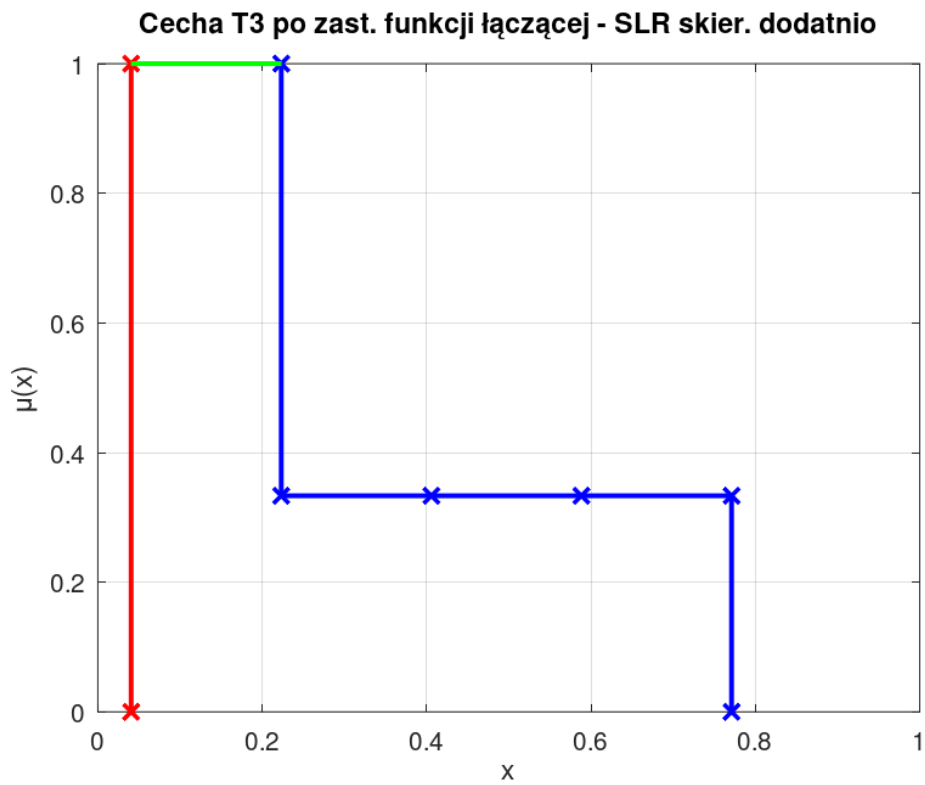
Rysunek 7.10: Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T8.



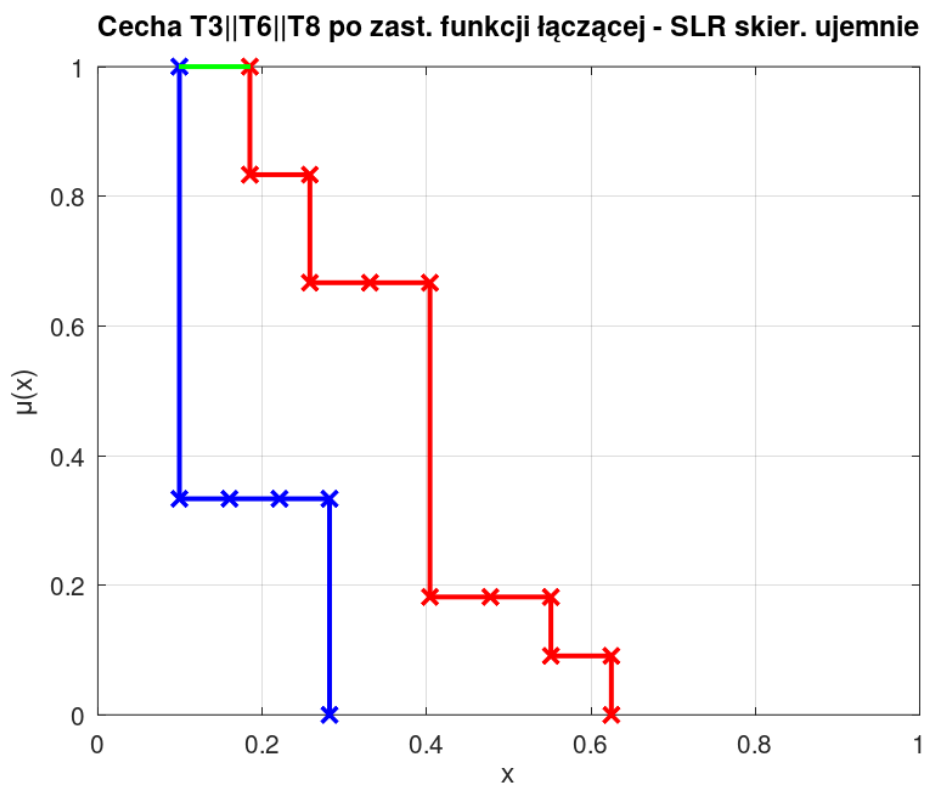
Rysunek 7.11: Wyznaczona SLR (po operacji złączenia) dla cechy T1.



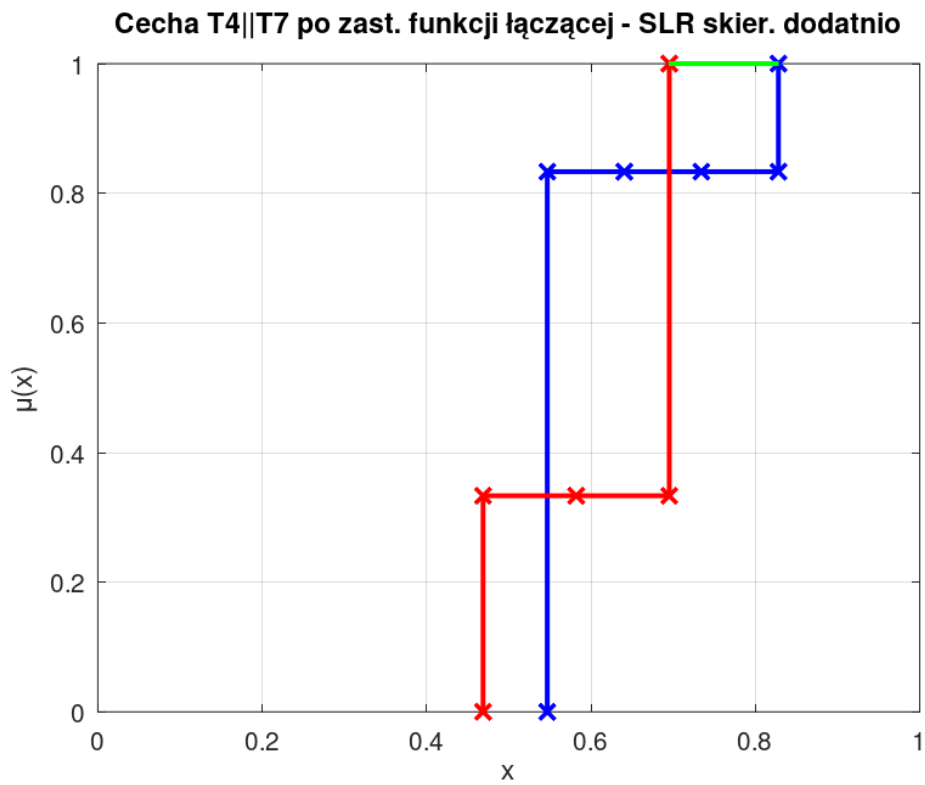
Rysunek 7.12: Wyznaczona SLR (po operacji złączenia) dla cechy T2.



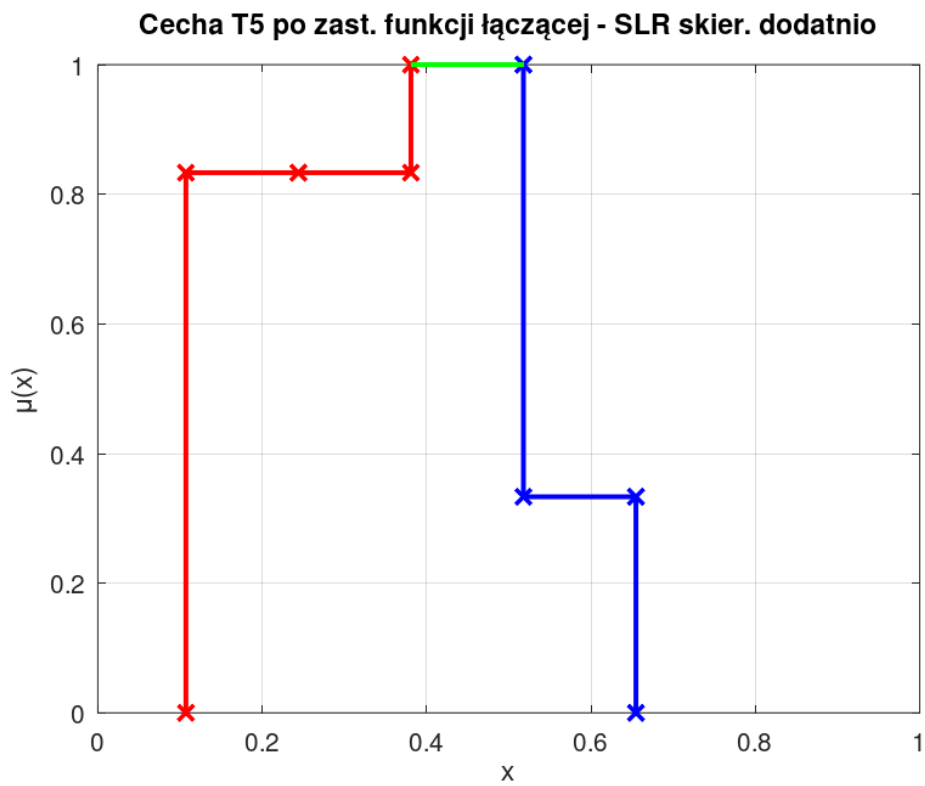
Rysunek 7.13: Wyznaczona SLR (po operacji złączenia) dla cechy T3.



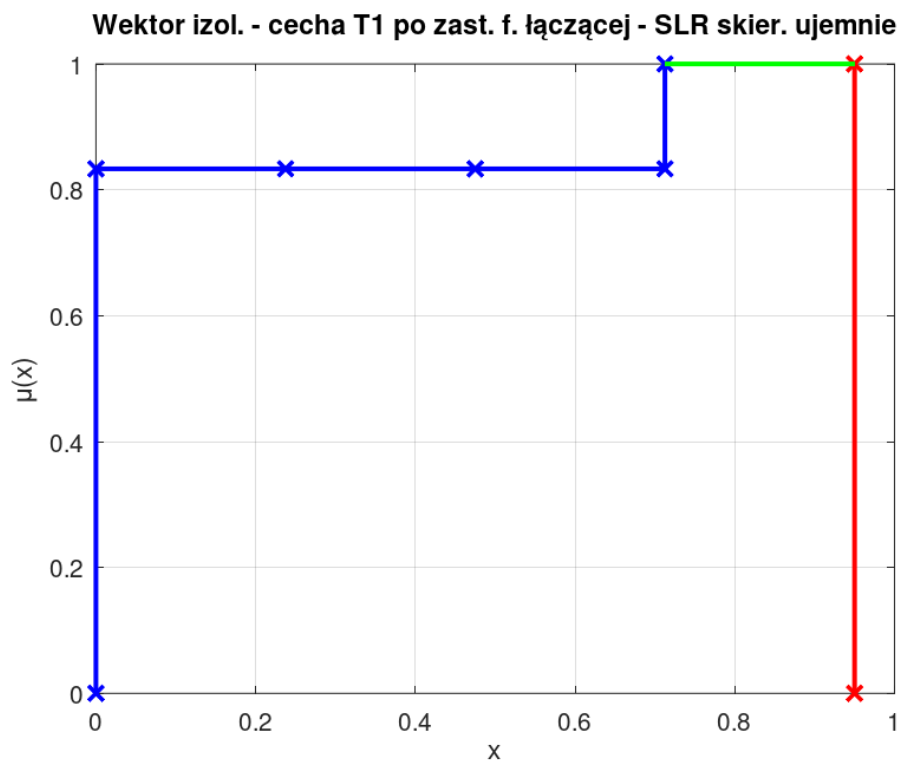
Rysunek 7.14: Wyznaczona SLR (po operacji złączenia) dla cechy T3||T6||T8.



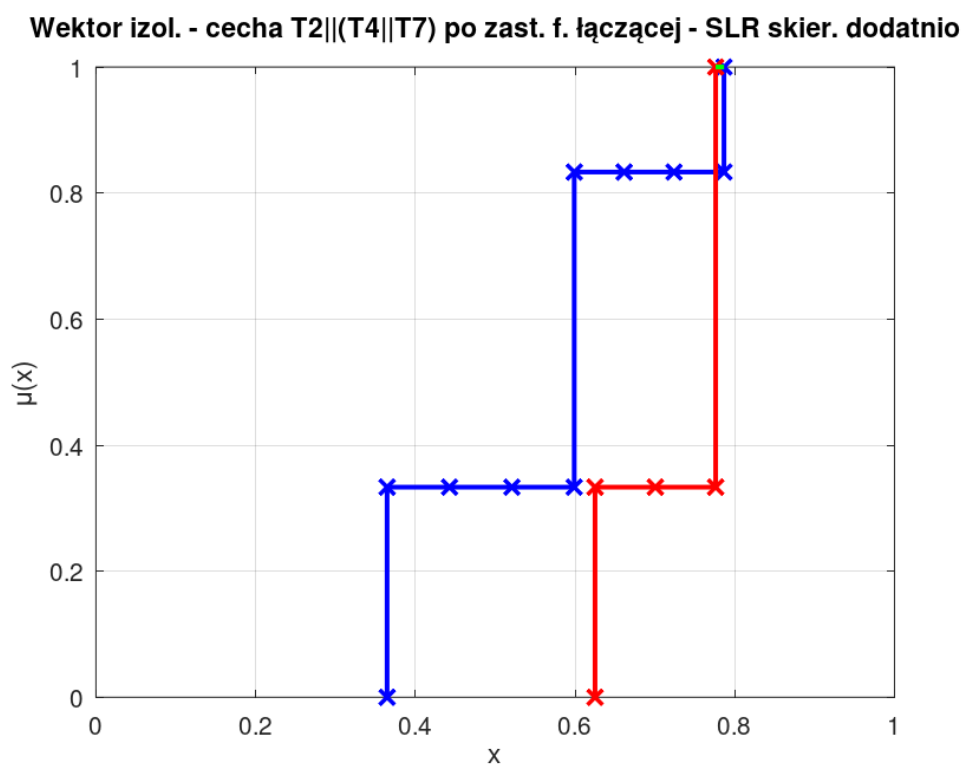
Rysunek 7.15: Wyznaczona SLR (po operacji złączenia) dla cechy T4||T7.



Rysunek 7.16: Wyznaczona SLR (po operacji złączenia) dla cechy T5.

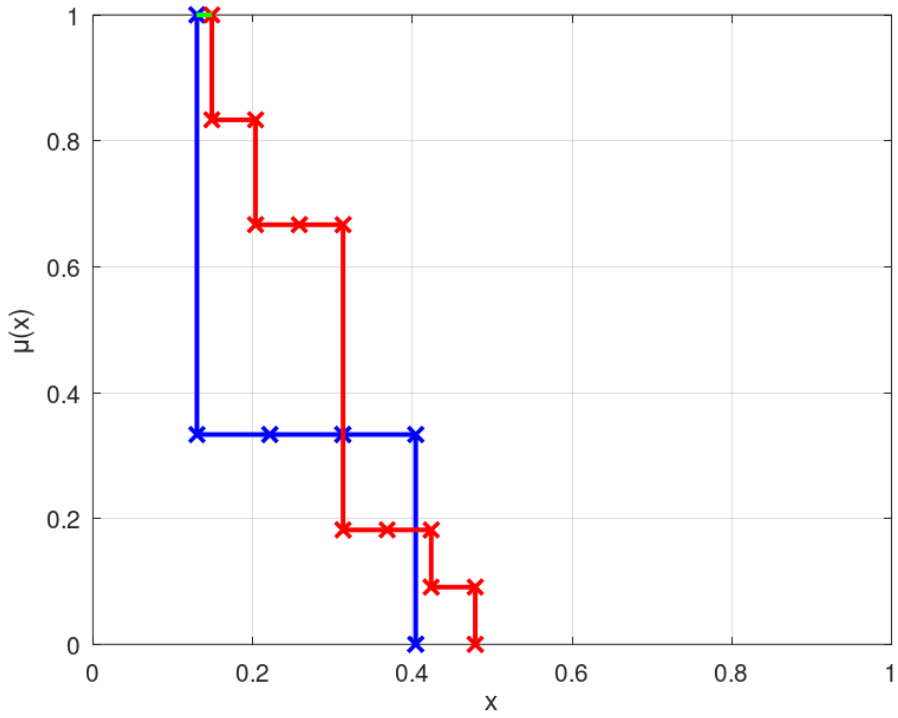


Rysunek 7.17: Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy T1.



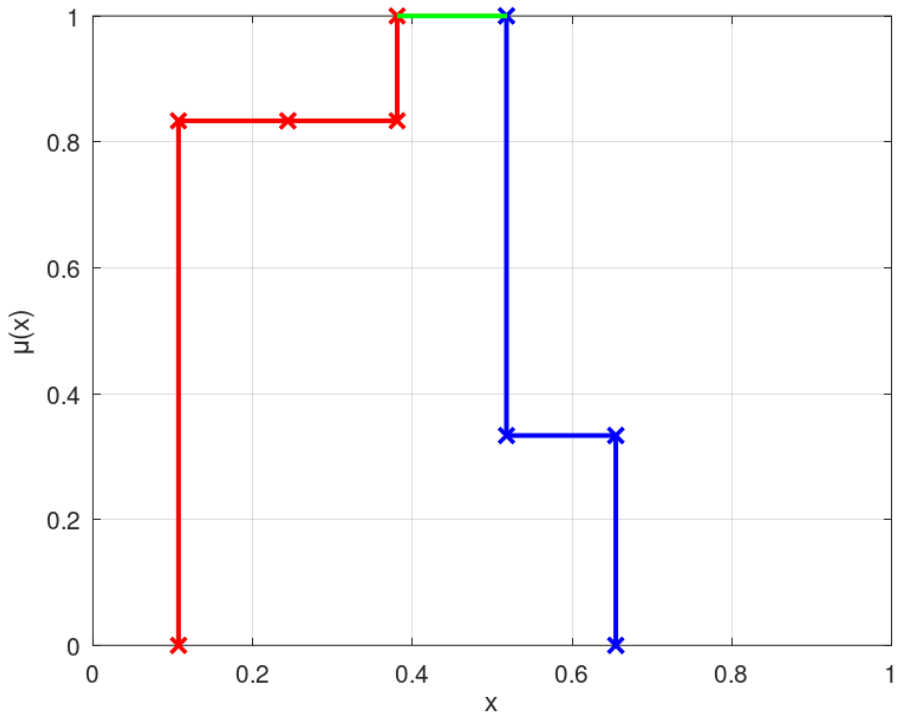
Rysunek 7.18: Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy T2||(T4||T7).

Wektor izol. - cecha T3||(T3||T6||T8) po zast. f. łączącej - SLR skier. ujemnie

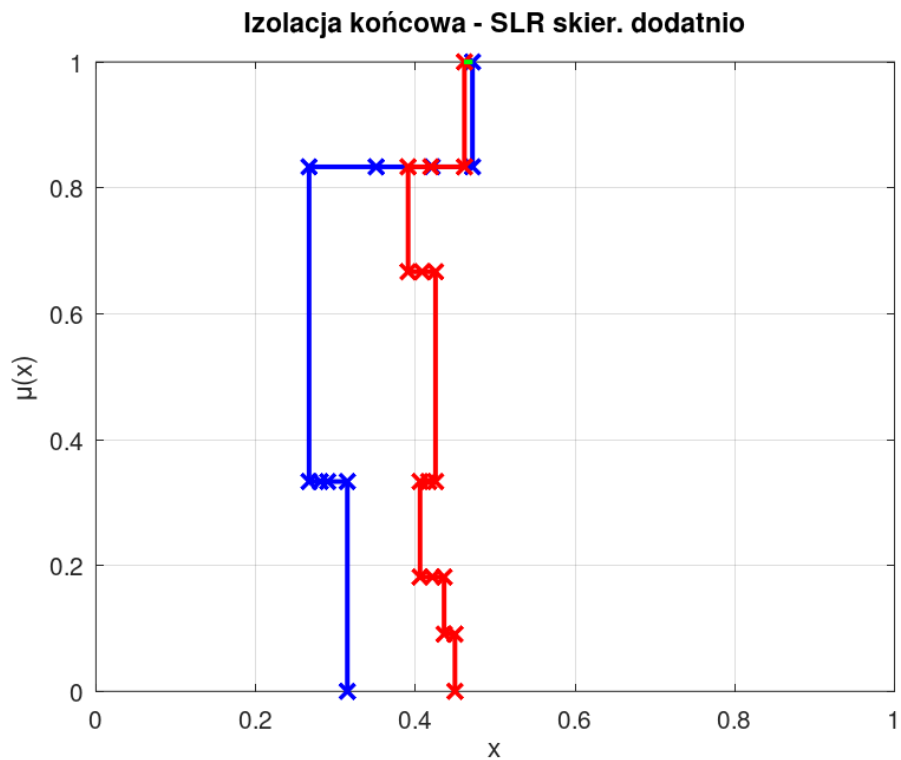


Rysunek 7.19: Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy T3||(T3||T6||T8).

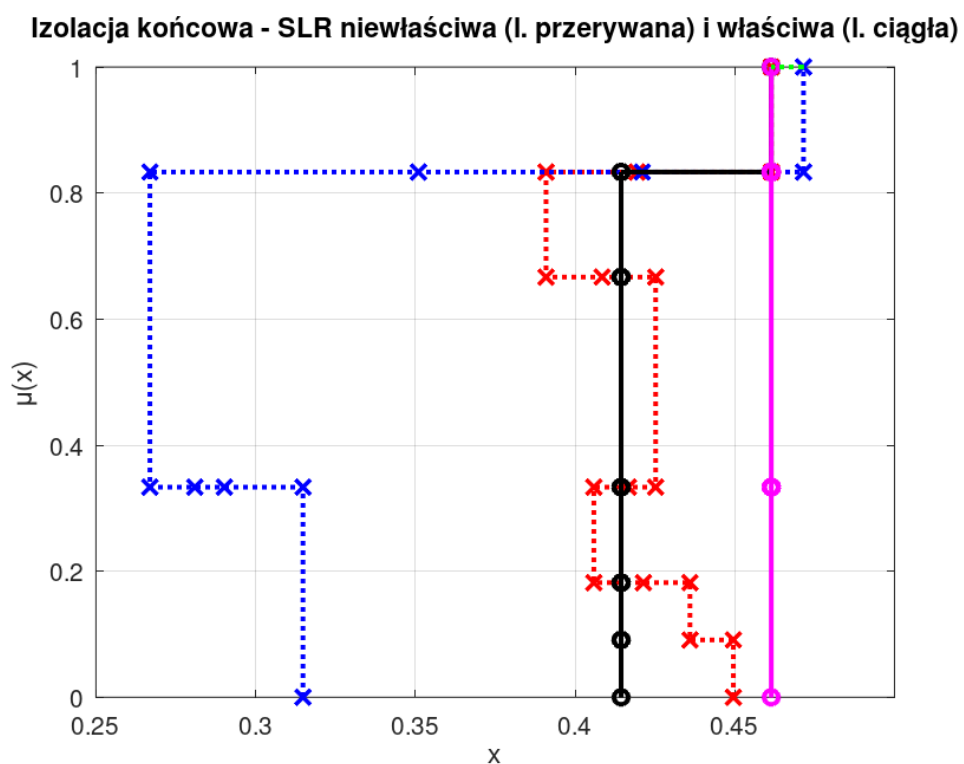
Wektor izol. - cecha T5 po zast. f. łączącej - SLR skier. dodatnio



Rysunek 7.20: Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy T5.



Rysunek 7.21: Wyznaczona SLR opisująca końcową izolację w sieci.



Rysunek 7.22: Wyznaczona SLR opisująca końcową izolację w sieci - wariant właściwy wyznaczony za pomocą programowania liniowego.

8. Model probabilistyczny

W niniejszym rozdziale przedstawimy założenia i model probabilistyczny izolacji w sieciach plastrowych.

8.1. ZAŁOŻENIA MODELU

Oczekujemy od modelu, aby spełniał następujące założenia.

- **Model probabilistyczny wykorzystuje ten sam zbiór danych co model rozmyty, bazujący na skierowanych liczbach rozmytych.** Pozwala to na porównanie wyników otrzymywanych z obydwu podejść, poprzez zredukowanie wpływu różnicy w ilości informacji (w rozumieniu Shannona [111]) niesionej przez same dane.
- **Model probabilistyczny wykorzystuje ten sam zbiór wierzchołków podstawowych.** Jest to założenie uzupełniające powyższe, aby uniknąć obciążenia (pozytywnego bądź negatywnego) wynikającego z dysponowania dodatkową informacją.
- Wyniki pomiarowe jakie rozpatrujemy w niniejszym rozdziale zostały uprzednio znormalizowane za pomocą funkcji normalizujących, opisanych szerzej w podrozdziale 4.2..

8.2. MODEL

Założmy, że dla węzła V_i i cechy r_j mamy ciąg wyników pomiarowych (lub określenia wartości cechy) $\mathbf{V}_{i,j} = (v_{i,j,1}, v_{i,j,2}, \dots, v_{i,j,n})^T$. Przyjmijmy:

$$\mathbf{H}(\mathbf{X}) = \begin{bmatrix} h_1(\mathbf{X}) \\ h_2(\mathbf{X}) \\ \vdots \\ h_K(\mathbf{X}) \end{bmatrix}. \quad (8.1)$$

Funkcje h_1, h_2, \dots, h_K są pewnymi funkcjami o znaczeniu statystycznym. Załóżmy, że dla węzła V_i i cechy r_j zdefiniowany jest ciąg zbiorów $(P_{i,j,1}, P_{i,j,2}, \dots, P_{i,j,K})$, które określają *pożądane* wartości odpowiednich funkcji (h_1, h_2, \dots, h_K) - dla tych wartości nie zachodzi złamanie izolacji. Oznaczmy:

$$s_{i,j,k} = \begin{cases} 1 \Leftrightarrow h_{i,j,k}(\mathbf{V}_{i,j}) \in P_{i,j,k}, \\ 0 \text{ w p.p.} \end{cases} \quad (8.2)$$

Stwórzmy zmienną losową $\mathbf{I}_{i,j}$:

$$\mathbf{I}_{i,j} = \sum_{k=1}^K s_{i,j,k} / K. \quad (8.3)$$

Zmienna losowa $\mathbf{I}_{i,j}$ odpowiada izolacji węzła. Rozpatrzmy teraz kilka przykładowych funkcji o znaczeniu statystycznym, jakie mogą być użyte do wyznaczania izolacji.

8.2.1. Funkcja minimum i maksimum

Niech:

$$\mathbf{Y} = \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N), \quad (8.4)$$

$$\mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N). \quad (8.5)$$

Zastanówmy się, jaki rozkład ma zmienna losowa \mathbf{Y} . Mamy dystrybuantę:

$$F_{\mathbf{Y}}(t) = \mathbb{P}(\mathbf{Y} \leq t) = \mathbb{P}(\min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N) \leq t) = 1 - \mathbb{P}(\min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N) > t). \quad (8.6)$$

Zauważmy, że:

$$\min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N) > t \Leftrightarrow \mathbf{X}_1 > t \wedge \mathbf{X}_2 > t \wedge \dots \wedge \mathbf{X}_N > t. \quad (8.7)$$

Jeśli zmienne losowe $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ są niezależne, to wówczas mamy:

$$F_{\mathbf{Y}}(t) = 1 - \prod_{i=1}^N \mathbb{P}(\mathbf{X}_i > t) = 1 - \prod_{i=1}^N (1 - \mathbb{P}(\mathbf{X}_i \leq t)). \quad (8.8)$$

a gdy te zmienne mają ten sam rozkład, oznaczony dla uproszczenia \mathbf{X} , to:

$$F_{\mathbf{Y}}(t) = 1 - (1 - \mathbb{P}(\mathbf{X} \leq t))^N = 1 - (1 - F_{\mathbf{X}}(t))^N. \quad (8.9)$$

Rozpatrzmy zmienną losową \mathbf{Z} . Zauważmy, że:

$$\mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N) \leq t \Leftrightarrow \mathbf{X}_1 \leq t \wedge \mathbf{X}_2 \leq t \wedge \dots \wedge \mathbf{X}_N \leq t. \quad (8.10)$$

Przy założeniu niezależności zmiennych losowych $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$, mamy:

$$F_{\mathbf{Z}}(t) = \mathbb{P}(\mathbf{Z} \leq t) = \mathbb{P}(\max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N) \leq t) = \prod_{i=1}^N (\mathbb{P}(\mathbf{X}_i \leq t)) = (F_{\mathbf{X}}(t))^N. \quad (8.11)$$

W przypadku ciągłych i różniczkowalnych rozkładów prawdopodobieństwa możemy wyznaczyć gęstość rozkładu prawdopodobieństwa zmiennej losowej \mathbf{Y} i \mathbf{Z} :

$$\begin{aligned} f_{\mathbf{Y}}(t) &= \frac{d}{dt} (1 - (1 - F_{\mathbf{X}}(t))^N) = \\ &= N(1 - F_{\mathbf{X}}(t))^{N-1} \frac{d}{dt} F_{\mathbf{X}}(t) = N(1 - F_{\mathbf{X}}(t))^{N-1} f_{\mathbf{X}}(t). \end{aligned} \quad (8.12)$$

$$f_{\mathbf{Z}}(t) = \frac{d}{dt} (F_{\mathbf{X}}(t))^N = N (F_{\mathbf{X}}(t))^{N-1} f_{\mathbf{X}}(t). \quad (8.13)$$

Momenty zwykłe rzędu m są dane wzorami:

$$\mathbb{E}\mathbf{Y}^m = N \int_{-\infty}^{\infty} t^m (1 - F_{\mathbf{X}}(t))^{N-1} f_{\mathbf{X}}(t) dt. \quad (8.14)$$

$$\mathbb{E}\mathbf{Z}^m = N \int_{-\infty}^{\infty} t^m (F_{\mathbf{X}}(t))^{N-1} f_{\mathbf{X}}(t) dt. \quad (8.15)$$

Definicja 8.1 (Symetryczny rozkład prawdopodobieństwa ciągłej zmiennej losowej)

Rozkład ciągłej zmiennej losowej \mathbf{X} nazywamy symetrycznym rozkładem prawdopodobieństwa (w skrócie rozkładem symetrycznym) wtedy i tylko wtedy, gdy jego gęstość prawdopodobieństwa spełnia zależność:

$$f_{\mathbf{X}}(t) = f_{\mathbf{X}}(-t). \quad (8.16)$$

Zmienną losową o takim rozkładzie oznaczamy jako $\mathbf{X} \sim \text{Sym}$.

Definicja 8.2 (Symetryczny rozkład prawdopodobieństwa dyskretnej zmiennej losowej)

Rozkład dyskretnej zmiennej losowej \mathbf{X} nazywamy symetrycznym rozkładem prawdopodobieństwa (w skrócie rozkładem symetrycznym) wtedy i tylko wtedy, gdy jego funkcja masy prawdopodobieństwa [115] spełnia zależność:

$$pmf_{\mathbf{X}}(t) = pmf_{\mathbf{X}}(-t). \quad (8.17)$$

Zmienną losową o takim rozkładzie oznaczamy jako $\mathbf{X} \sim \text{Sym}$.

Zachodzą poniższe lematy, których dowody są załączone w dodatku A.

Lemat 8.1 (o momentach funkcji minimum dla symetrycznych ciągłych zmiennych losowych)

Niech $\mathbf{Y} = \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ i $\mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ będą zmiennymi losowymi zależnymi od ciągu niezależnych zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$, które mają taki sam ciągły rozkład prawdopodobieństwa $\mathbf{X} \sim \text{Sym}$. Wówczas dla $m \in \mathbb{N}$ zachodzi równość:

$$\mathbb{E}(\mathbf{Y}^m) = (-1)^m \mathbb{E}(\mathbf{Z}^m). \quad (8.18)$$

Lemat 8.2 (o momentach funkcji minimum dla symetrycznych dyskretnych zmiennych losowych)

Niech $\mathbf{Y} = \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ i $\mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ będą zmiennymi losowymi zależnymi od ciągu niezależnych zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$, które mają taki sam dyskretny rozkład prawdopodobieństwa $\mathbf{X} \sim \text{Sym}$. Wówczas dla $m \in \mathbb{N}$ zachodzi równość:

$$\mathbb{E}(\mathbf{Y}^m) = (-1)^m \mathbb{E}(\mathbf{Z}^m). \quad (8.19)$$

Wniosek 8.1 (Wariancja funkcji minimum i maksimum)

Zachodzi równość wariancji dla funkcji minimum i maksimum:

$$\text{Var} \mathbf{Y} = \mathbb{E} \mathbf{Y}^2 - \mathbb{E}^2 \mathbf{Y} = (-1)^2 \mathbb{E} \mathbf{Z}^2 - (-\mathbb{E} \mathbf{Z})^2 = \text{Var} \mathbf{Z}. \quad (8.20)$$

Rozpatrzmy teraz kilka podstawowych rozkładów prawdopodobieństwa.

Rozkład jednostajny ciągły

Niech $X \sim U[a; b]$ oraz $a < b$. Dla $t \leq a$ mamy $F_Y(t) = 0$ oraz dla $b \leq t$ mamy $F_Y(t) = 1$. W przypadku gdy $a \leq t \leq b$ mamy $F_X(t) = \frac{t-a}{b-a}$ oraz:

$$F_Y(t) = 1 - (1 - F_X(t))^n = 1 - \left(1 - \frac{t-a}{b-a}\right)^n = 1 - \left(\frac{b-t}{b-a}\right)^n. \quad (8.21)$$

Gęstość prawdopodobieństwa w przedziale $[a; b]$ jest wyrażana wzorem:

$$f_Y(t) = n \left(\frac{b-t}{b-a}\right)^{n-1} \frac{1}{b-a}, \quad (8.22)$$

a wartość oczekiwana jest równa:

$$\begin{aligned} \mathbb{E}Y &= n \int_a^b t \left(\frac{b-t}{b-a}\right)^{n-1} \frac{1}{b-a} dt = \frac{-n}{(b-a)^n} \int_a^b -t (b-t)^{n-1} dt = \\ &= \frac{-n}{(b-a)^n} \int_a^b (b-t)^n - b(b-t)^{n-1} dt = \\ &= \frac{-n}{(b-a)^n} \left(\frac{-(b-a)^{n+1}}{-(n+1)} - \frac{-b(b-a)^n}{-n} \right) = \\ &= -n \left(\frac{(b-a)}{(n+1)} - \frac{b}{n} \right) = b - \frac{n(b-a)}{(n+1)} = a + \frac{b-a}{n+1}. \end{aligned} \quad (8.23)$$

Wyznamy teraz drugi moment, aby móc obliczyć wariancję dla takiego rozkładu podstawowego.

$$\mathbb{E}Y^2 = n \int_a^b t^2 \left(\frac{b-t}{b-a}\right)^{n-1} \frac{1}{b-a} dt = \frac{n}{(b-a)^n} \int_a^b t^2 (b-t)^{n-1} dt. \quad (8.24)$$

Zauważmy, że:

$$\int t^2 (b-t)^{n-1} dt = \frac{-(b-t)^{n+2}}{n+2} + \frac{2b(b-t)^{n+1}}{n+1} - \frac{b^2(b-t)^n}{n} + C, \quad (8.25)$$

gdzie $C \in \mathbb{R}$ jest stałą całkowania. Wobec tego mamy:

$$\begin{aligned}
\mathbb{E}\mathbf{Y}^2 &= n \int_a^b t^2 \left(\frac{b-t}{b-a} \right)^{n-1} \frac{1}{b-a} dt = \\
&= \frac{n}{(b-a)^n} \left(\frac{-(b-t)^{n+2}}{n+2} + \frac{2b(b-t)^{n+1}}{n+1} - \frac{b^2(b-t)^n}{n} \right) \Big|_a^b = \\
&= \frac{n}{(b-a)^n} \left(\frac{(b-a)^{n+2}}{n+2} - \frac{2b(b-a)^{n+1}}{n+1} + \frac{b^2(b-a)^n}{n} \right) = \\
&= n \left(\frac{(b-a)^2}{n+2} - \frac{2b(b-a)}{n+1} + \frac{b^2}{n} \right).
\end{aligned} \tag{8.26}$$

Wobec tego wariancja jest równa:

$$\begin{aligned}
Var\mathbf{Y} &= \mathbb{E}\mathbf{Y}^2 - \mathbb{E}^2\mathbf{Y} = n \left(\frac{(b-a)^2}{n+2} - \frac{2b(b-a)}{n+1} + \frac{b^2}{n} \right) - \left(-n \left(\frac{(b-a)}{(n+1)} - \frac{b}{n} \right) \right)^2 = \\
&= \frac{n(b-a)^2}{(n+2)(n+1)^2}.
\end{aligned} \tag{8.27}$$

Na podstawie wniosku 8.1 wariancja funkcji maksimum wyraża się takim samym wzorem co wyprowadzona powyżej wariancja dla funkcji minimum.

Rozkład jednostajny dyskretny

Załóżmy, że \mathbf{X} ma K -punktowy rozkład jednostajny (dyskretny). Rozkład taki ma dystrybuantę w punkcie c_k równą:

$$F_{\mathbf{X}}(c_k) = \frac{k}{K}, \tag{8.28}$$

co oznacza, że dystrybuanta zmiennej losowej \mathbf{Y} w punkcie c_k wyraża się wzorem:

$$F_{\mathbf{Y}}(c_k) = 1 - \left(1 - \frac{k}{K} \right)^n. \tag{8.29}$$

Wartość oczekiwana jest równa:

$$\mathbb{E}\mathbf{Y} = \sum_{k=1}^K c_k \left(\left(1 - \frac{k-1}{K} \right)^n - \left(1 - \frac{k}{K} \right)^n \right). \tag{8.30}$$

Rozkład wykładniczy

Niech $\mathbf{X} \sim Exp(\lambda)$ oraz $\lambda > 0$. Dystrybuanta tego rozkładu ma wzór $F_{\mathbf{X}}(t) = (1 - e^{-\lambda t}) \cdot \mathbf{1}_{[0, \infty)}(t)$, a jego gęstość prawdopodobieństwa określona jest wzorem $\lambda e^{-\lambda t} \cdot \mathbf{1}_{[0, \infty)}(t)$. Wobec tego mamy:

$$F_{\mathbf{Y}}(t) = 1 - (1 - F_{\mathbf{X}}(t))^n = (1 - e^{-\lambda nt}) \cdot \mathbf{1}_{[0, \infty)}(t). \quad (8.31)$$

Gęstość prawdopodobieństwa wyrażana jest wzorem:

$$f_{\mathbf{Y}}(t) = \lambda n e^{-\lambda nt} \cdot \mathbf{1}_{[0, \infty)}(t), \quad (8.32)$$

a wartość oczekiwana jest równa:

$$\mathbb{E}\mathbf{Y} = \int_0^{\infty} t \lambda n e^{-\lambda nt} dt. \quad (8.33)$$

Niech $q = \lambda nt$, wówczas mamy:

$$\mathbb{E}\mathbf{Y} = \frac{1}{\lambda n} \int_0^{\infty} q e^{-q} dq = \frac{1}{\lambda n} (-e^{-q}(q+1)) \Big|_0^{\infty} = \frac{1}{\lambda n}. \quad (8.34)$$

8.2.2. Funkcja maksimum

Niech:

$$\mathbf{Z} = h(\mathbf{X}) = \max(x_1, x_2, \dots, x_n). \quad (8.35)$$

Dla rozkładów dyskretnych, które przyjmują wartości c_1, c_2, \dots, c_K , mamy:

$$\mathbb{E}\mathbf{Z} = \sum_{k=1}^K c_k (F_{\mathbf{Z}}(c_k) - F_{\mathbf{Z}}(c_{k-1})), \quad (8.36)$$

gdzie $F_{\mathbf{Z}}(c_0) = 0$. Rozpatrzmy teraz kilka podstawowych rozkładów prawdopodobieństwa.

Rozkład jednostajny ciągły

Niech $x_1 \sim U[a; b]$ oraz $a < b$. Dla $t \leq a$ mamy $F_{\mathbf{Z}}(t) = 0$ oraz dla $b \leq t$ mamy $F_{\mathbf{Z}}(t) = 1$. W przypadku gdy $a \leq t \leq b$ mamy $F_{\mathbf{X}}(t) = \frac{t-a}{b-a}$ oraz:

$$F_{\mathbf{Z}}(t) = (F_{\mathbf{X}}(t))^n = \left(\frac{t-a}{b-a}\right)^n. \quad (8.37)$$

Gęstość prawdopodobieństwa w przedziale $[a; b]$ jest wyrażana wzorem:

$$f_{\mathbf{Z}}(t) = n \left(\frac{t-a}{b-a}\right)^{n-1} \frac{1}{b-a}, \quad (8.38)$$

a wartość oczekiwana jest równa:

$$\begin{aligned} \mathbb{E}\mathbf{Z} &= n \int_a^b t \left(\frac{t-a}{b-a}\right)^{n-1} \frac{1}{b-a} dt = \frac{n}{(b-a)^n} \int_a^b t (t-a)^{n-1} dt = \\ &= \frac{n}{(b-a)^n} \int_a^b ((t-a)^n + a(t-a)^{n-1}) dt = \\ &= \frac{n}{(b-a)^n} \left(\frac{(b-a)^{n+1}}{n+1} + \frac{a}{n} (b-a)^n \right) = \frac{n(b-a)}{n+1} + a. \end{aligned} \quad (8.39)$$

Rozkład jednostajny dyskretny

Załóżmy, że \mathbf{X} ma K -punktowy rozkład jednostajny (dyskretny). Rozkład taki ma dystrybuantę w punkcie c_k równą:

$$F_{\mathbf{X}}(c_k) = \frac{k}{K}, \quad (8.40)$$

co oznacza, że dystrybuanta zmiennej losowej \mathbf{Z} w punkcie c_k wyraża się wzorem:

$$F_{\mathbf{Z}}(c_k) = \left(\frac{k}{K}\right)^n. \quad (8.41)$$

Wartość oczekiwana jest równa:

$$\mathbb{E}\mathbf{Z} = \sum_{k=1}^K c_k \left(\left(\frac{k}{K}\right)^n - \left(\frac{k-1}{K}\right)^n \right). \quad (8.42)$$

8.2.3. Estymator wartości oczekiwanej

Estymator wartości oczekiwanej z próby (x_1, x_2, \dots, x_n) ma wartość:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (8.43)$$

Wartość oczekiwana tego estymatora jest średnią arytmetyczną wartości oczekiwanych poszczególnych próbek:

$$\mathbb{E}\bar{X} = \mathbb{E} \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n \mathbb{E}x_i. \quad (8.44)$$

Gdy poszczególne próbki mają ten sam rozkład prawdopodobieństwa, to wówczas:

$$\mathbb{E}\bar{X} = \frac{1}{n} \sum_{i=1}^n \mathbb{E}x_i = \mu, \quad (8.45)$$

gdzie $\mu = \mathbb{E}x_i$ dla $i = 1, 2, \dots, n$. Oznacza to, że jest to estymator nieobciążony. Dla typowych rozkładów prawdopodobieństwa znane są wartości oczekiwane w zależności od parametrów tych rozkładów.

8.2.4. Wyznaczanie zbiorów

Rozpatrzmy teraz sposoby wyznaczania zbiorów P . Istnieje kilka metod definiowania takiego zbioru.

- zbiór wynikający z wymagań biznesowych - w ramach danego parametru mogą być zadane z góry dopuszczalne wartości izolacji, szczególnie w przypadku parametrów o charakterze enumeratywnym.
- metoda zbioru arbitralnego - zbiór P wyznaczany jest arbitralnie, a stopień swobody zapewniany jest poprzez dobór odpowiedniej funkcji normalizującej. Zauważmy, że pierwotne wyniki pomiarowe można normalizować na nieskończenie wiele sposobów i dlatego znajomość kontekstu danego parametru pozwala podejmować decyzję, czy jego wartość jest wystarczająco dobra (zapewnia izolację), czy też nie.

Skupmy się na drugim podejściu. Zmienne jakie bierzemy pod uwagę, zostały wcześniej znormalizowane za pomocą funkcji normalizujących, co oznacza, że wartość 0.5 jest uznawana za *typową*. Wprowadźmy teraz *zbiór wartości progowych* $0 < \phi_{i,j} \leq 1$, gdzie $\phi_{i,j}$ jest wartością

taką, że uznajemy iż cecha r_j w węźle V_i zachowuje izolację (nie zachodzi złamanie izolacji ze względu na wartość tej cechy w tym wierzchołku).

Zastanówmy się teraz, jakie są warunki konieczne do tego, aby izolacja była zachowana, przy założeniu, że dla danej cechy r_j i wierzchołka V_i wystarczy, aby:

$$x_1, x_2, \dots, x_n \geq \phi_{i,j}. \quad (8.46)$$

Skoro każdy element z ciągu x_1, x_2, \dots, x_n jest większy niż $\phi_{i,j}$, to również zachodzi:

$$\min\{x_1, x_2, \dots, x_n\} \geq \phi_{i,j}, \quad (8.47)$$

$$\bar{\mathbf{X}} = \frac{1}{n} \sum_{k=1}^n x_k \geq \phi_{i,j}. \quad (8.48)$$

Nieobciążony estymator wariancji jest równy:

$$\begin{aligned} \overline{Var \mathbf{X}} &= \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{\mathbf{X}})^2 = \frac{1}{n-1} \left(n\bar{\mathbf{X}}^2 - 2\bar{\mathbf{X}} \sum_{k=1}^n x_k + \sum_{k=1}^n x_k^2 \right) = \\ &= \frac{1}{n-1} \left(\sum_{k=1}^n x_k^2 - n\bar{\mathbf{X}}^2 \right). \end{aligned} \quad (8.49)$$

Zauważmy, że w naszym przypadku zachodzi:

$$\left(\bigwedge_{k=1}^n \right) \phi_{i,j} \leq x_k \leq 1, \quad (8.50)$$

oraz mamy:

$$|x_k - \bar{\mathbf{X}}| \leq 1 - \phi_{i,j}, \quad (8.51)$$

co nam daje:

$$\overline{Var \mathbf{X}} = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{\mathbf{X}})^2 \leq \frac{n}{n-1} (1 - \phi_{i,j})^2. \quad (8.52)$$

Dostajemy wobec tego funkcje i zbiory jak w tabeli 8.1.

Tablica 8.1: Funkcje statystyczne i odpowiadające im zbiory

Funkcja	Formuła	Zbiór $P_{i,j,k}$
$h_{i,j,1}$	$\min\{x_1, x_2, \dots, x_n\}$	$[\phi_{i,j}; 1]$
$h_{i,j,2}$	$\bar{X} = \frac{1}{n} \sum_{k=1}^n x_k$	$[\phi_{i,j}; 1]$
$h_{i,j,3}$	$\overline{Var X} = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{X})^2$	$[0; \frac{n}{n-1} (1 - \phi_{i,j})^2]$

8.3. ALGORYTM WYZNACZANIA IZOLACJI W MODELU PROBABILISTYCZNYM

Przedstawmy teraz algorytm 10, służący do wyznaczania izolacji w modelu probabilistycznym. Algorytm jest podobny do algorytmu wykorzystującego skierowane liczby rozmyte - używa zbioru danych pomiarowych przypisanych do cech opisujących poszczególne węzły sieci. Różnicą w stosunku do algorytmu dla liczb rozmytych jest inny sposób przekształceń danych znormalizowanych - w niniejszym algorytmie użyto przekształceń z wykorzystaniem zbioru funkcji o znaczeniu statystycznym zamiast przekształceń prowadzących do zbudowania skierowanych liczb rozmytych. Algorytm 11 przedstawia funkcję łączącą w wariancie dla modelu probabilistycznego - używane są w niej liczby rzeczywiste.

8.4. PRZYKŁAD OBLICZENIOWY W MODELU PROBABILISTYCZNYM

Przeprowadźmy teraz obliczenia izolacji dla przykładowej sieci przedstawionej wcześniej w niniejszej pracy na rysunkach 7.1 i 7.2. Dane dotyczące poszczególnych cech pozostają takie same jak w przykładzie przedstawionym dla modelu rozmytego - zob. tablice 7.1, 7.2 i 7.3. Tablica 8.2 zawiera zbiór wartości progowych wyznaczonych dla tego przykładu. Wartość progowa $\phi_{i,j}$ jest wyznaczana poprzez zastosowanie funkcji normalizującej do pewnej wartości nieznormalizowanej, którą uznano za zapewniającą izolację - z perspektywy Operatora Telekomunikacyjnego, usługodawcy lub użytkownika. Może to być pochodna innych założeń związanych z bezpieczeństwem systemu w danym zakresie, np. długość klucza użytego do szyfrowania jest związana z przewidywanym okresem, w którym tekst zaszyfrowany powinien

Tablica 8.2: Wartości progowe $\phi_{i,j}$

Cecha V_i / r_j	Wartość przed normalizacją	$\phi_{i,j}$
T1 V_1/r_1	0.6	0,6000
T2 V_2/r_1	192 bity	0,7500
T3 V_3/r_1	3h	0,5946
T4 V_4/r_1	192 bity	0,7500
T5 V_4/r_2	19dB μ V/m	0,5597
T6 V_5/r_1	3h	0,5946
T7 V_6/r_1	192 bity	0,7500
T8 V_7/r_1	3h	0,5946

pozostać tajny, co przy aktualnym stanie wiedzy z zakresu kryptoanalizy, dostępnej mocy obliczeniowej i rozwiązań technologicznych, pozwala na przewidywanie, że w danym okresie czasu praktycznie niemożliwe będzie odtworzenie tekstu jawnego z odpowiednio dużym prawdopodobieństwem. Może się oczywiście zdarzyć, że atakujący szybko odgadnie klucz szyfrujący, ale przy odpowiednio dużej przestrzeni potencjalnych kluczy jest to zdarzenie o znikomym prawdopodobieństwie. Wartości funkcji $h_{i,j,1}$, $h_{i,j,2}$, $h_{i,j,3}$ dla tych cech zostały przedstawione w tablicy 8.3.

Tablica 8.4 zawiera wartości izolacji po operacji złączenia w ramach poszczególnych podsieci (części). Tak jak w przypadku modelu rozmytego, skorzystano ze średniej potęgowej o równych wagach i parametrze $q = 1$ (średnia arytmetyczna). Tablica 8.5 przedstawia sposób wyliczania wartości wektora izolacji na podstawie wyników uzyskanych dla podsieci; wagi odpowiadają udziałom poszczególnych podsieci w ramach danej cechy. Izolacja końcowa jest

równa:

$$\begin{aligned} \mathbf{I} &= 1 - \sqrt{\frac{1}{4} \sum_{i=1}^4 (1 - isolationVector[i])^2} = \\ &= 1 - \sqrt{\frac{1}{4} ((1 - 0,3333)^2 + (1 - 0,1111)^2 + (1 - 0,3333)^2 + (1 - 0,3333)^2)} = \\ &= 0,2714. \end{aligned} \quad (8.53)$$

Przejdźmy teraz do dyskusji otrzymanych wyników. Otrzymane wartości liczbowe w naturalny sposób zależą od użytych funkcji normalizujących i od przyjętych zbiorów $P_{i,j,k}$. Parametr $\phi_{i,j}$ powinien być dobrany tak, aby w przypadku danej funkcji normalizacyjnej było możliwe rozróżnienie wartości pośrednich izolacji. Zbyt duża wartość $\phi_{i,j}$ spowoduje częste pojawianie się wyników zerowych, zbyt mała nie będzie gwarantować, że faktycznie izolacja jest zachowana w kontekście danej cechy i danego węzła. Poprawę modelu można również uzyskać dzięki dodaniu dodatkowych metryk, na przykład mierząc wzrost (spadek) wartości parametru pomiędzy próbkami - dla cech wolnozmiennych jest to sygnał do weryfikacji stanu danej cechy, a w przypadku cech szybkozmiennych (np. dostępna pamięć RAM, liczba użytkowników w systemie, liczba otwartych portów, stany kolejek, itd.) zwykle można określić przedział, w jakim taka zmienność jest uznawana za naturalną.

Algorytm 10 Algorytm wyznaczania izolacji w modelu probabilistycznym.

```
1: procedure CALCULATEISOLATIONPROB
2:   Wyznacz zbiór wierzchołków podstawowych  $V = \{V_1, V_2, \dots, V_N\}$  dla danego strumienia danych.
3:   for  $V_i \in V$  do
4:     Wyznacz zbiór cech  $\Pi_i = \{\pi_{i,1}, \pi_{i,2}, \dots, \pi_{i,M_i}\}$  właściwych dla  $V_i$ .
5:     Ustal funkcje normalizujące  $(g_{i,1}, g_{i,2}, \dots, g_{i,M_i})$  dla cech  $(\pi_{i,1}, \pi_{i,2}, \dots, \pi_{i,M_i})$  przynależnych do  $V_i$ .
6:     Wyznacz wektor wartości progowych  $\Phi_i = [\phi_{i,1}, \phi_{i,2}, \dots, \phi_{i,M_i}]$ .
7:     Ustal funkcje  $F_1, F_2, \dots, F_{|\Pi|}$  łączące izolację dla poszczególnych cech ze zbioru  $\Pi = \cup_{i=1}^N \Pi_i = \{\pi_1, \pi_2, \dots, \pi_{|\Pi|}\}$ . //w naszym przypadku będzie to funkcja MergeFunctionProb
8:     for  $V_i \in V$  do
9:       for  $j = 1, 2, \dots, M_i$  do
10:        Zbierz wartości  $(x_1, x_2, \dots, x_k)$  cechy  $\pi_{i,j}$  dla wierzchołka  $V_i$ .
11:         $(x'_1, x'_2, \dots, x'_k) = (g_{i,j}(x_1), g_{i,j}(x_2), \dots, g_{i,j}(x_k))$  - normalizacja wartości cech, korzystając z funkcji normalizujących ustalonych w kroku 4.
12:        for  $k = 1, 2, \dots, K$  do
13:           $s_{i,j,k} = h_{i,j,k}((x'_1, x'_2, \dots, x'_k)) \in P_{i,j,k}$ 
14:           $I[V_i, \pi_j] = \sum_{k=1}^K s_{i,j,k} / K$ 
15:        for  $G_s \in \{G_1, G_2, \dots, G_S\}$  do
16:          for  $\pi_j \in \Pi$  do
17:             $W = \{V_u \in V : \pi_j \in \Pi_u\}$  //wierzchołki z  $G_s$  z daną cechą
18:             $tmp = W.Select(v \Rightarrow I[v, \pi_j])$  //lista wartości izolacji dla wszystkich wierzchołków z  $G_s$  z daną cechą
19:             $weights = repeat(1.0 / len(tmp), len(tmp))$  //wektor takich samych wag sumujących się do 1
20:             $r_{j,s} = MergeFunctionProb(tmp, weights)$ 
21:             $globalWeights[\pi_j].Add(len(tmp))$  //dodaje skalar do listy wag
22:             $mergedResults[\pi_j].Add(r_{j,s})$ 
23:          for  $\pi_i \in \Pi$  do
24:             $weights = globalWeights[\pi_i].Select(v \Rightarrow v/globalWeights[\pi_i].Sum())$  // wagi muszą się sumować do 1
25:             $isolationVector[i] = MergeFunctionProb(mergedResults[\pi_i], weights)$ 
26:             $q = 2$  // na potrzeby obliczania izolacji końcowej
27:             $I = 1 - \left( \frac{1}{|\Pi|} \sum_{i=1}^{|\Pi|} (1 - isolationVector[i])^q \right)^{\frac{1}{q}}$  // izolacja końcowa
```

Algorytm 11 Przykładowa funkcja łącząca - wariant do modelu probabilistycznego

```

1: procedure MERGEFUNCTIONPROB(X, weights)
2:   q=1 //parametr zależny od cechy, tutaj przykładowa wartość. Dla q = 0 wymagana jest
   osobna implementacja - średnia geometryczna
3:   res = 0
4:   for i = 0; i < len(X); ++ i do
5:     res += pow(X[i], q) * weights[i]
6:   return pow(res, 1/q)
  
```

 Tablica 8.3: Wartości estymatorów i zbiory $P_{i,j,k}$

Cecha V_i / r_j	$h_{i,j,1}$	$h_{i,j,2}$	$h_{i,j,3}$	$\phi_{i,j}$	$P_{i,j,1}$	$P_{i,j,2}$	$P_{i,j,3}$	$I_{i,j}$
T1 V_1/r_1	0,0000	0,4438	0,1206	0,6000	[0,6000; 1]	[0,6000; 1]	[0; 0,1707]	0,3333
T2 V_2/r_1	0,0000	0,6035	0,0997	0,7500	[0,7500; 1]	[0,7500; 1]	[0; 0,0667]	0,0000
T3 V_3/r_1	0,0405	0,2971	0,0588	0,5946	[0,5946; 1]	[0,5946; 1]	[0; 0,1753]	0,3333
T4 V_4/r_1	0,1563	0,5078	0,0751	0,7500	[0,7500; 1]	[0,7500; 1]	[0; 0,0667]	0,0000
T5 V_4/r_2	0,1072	0,3451	0,0293	0,5597	[0,5597; 1]	[0,5597; 1]	[0; 0,2068]	0,3333
T6 V_5/r_1	0,0405	0,4018	0,0888	0,5946	[0,5946; 1]	[0,5946; 1]	[0; 0,1753]	0,3333
T7 V_6/r_1	0,0313	0,5859	0,0649	0,7500	[0,7500; 1]	[0,7500; 1]	[0; 0,0667]	0,3333
T8 V_7/r_1	0,0341	0,2794	0,0606	0,5946	[0,5946; 1]	[0,5946; 1]	[0; 0,1753]	0,3333

Tablica 8.4: Izolacja wewnątrz sieci RAN i CN - model probabilistyczny

Część	Parametr	Wartości	Złączone wartości
RAN	Zastosowany język programowania	T1	0,3333
	Siła użytego szyfrowania symetrycznego	T2	0,0
	Średni czas pomiędzy analizami podatności	T3	0,3333
CN	Siła użytego szyfrowania symetrycznego	T4, T7	$\frac{0,0+0,3333}{2}$
	Emisja elektromagnetyczna	T5	0,3333
	Średni czas pomiędzy analizami podatności	T3, T6, T8	$\frac{0,3333+0,3333+0,3333}{3}$

Tablica 8.5: Wektor izolacji - model probabilistyczny

Cecha	Wartości	Wagi	Wartości wektora izolacji - model probabilistyczny
Zastosowany język programowania	T1	1	0,3333
Siła użytego szyfrowania symetrycznego	T2, T4 T7	1/3, 2/3	$\frac{1}{3}T2 + \frac{2}{3}T4 T7 = \frac{1}{3} \cdot 0,0 + \frac{2}{3} \cdot \frac{0,0+0,3333}{2} = 0,1111$
Średni czas pomiędzy analizami podatności	T3, T3 T6 T8	1/4, 3/4	$\frac{1}{4} \cdot 0,3333 + \frac{3}{4} \cdot \frac{0,3333+0,3333+0,3333}{3} = 0,3333$
Emisja elektromagnetyczna	T5	1	0,3333

9. Izolacja usług

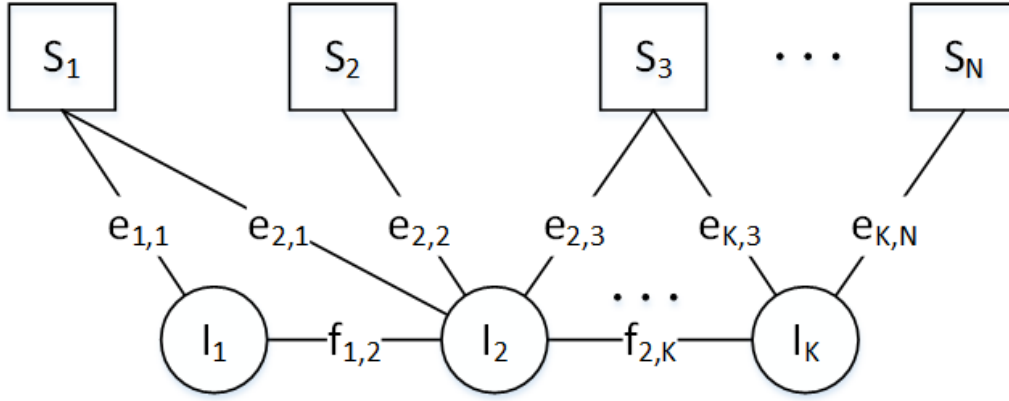
9.1. PRZEDSTAWIENIE ZAGADNIENIA

Izolację pomiędzy plastrami można interpretować jako izolację pomiędzy usługami. W pracy [88] opisano matematyczny model izolacji usług w sieciach plastrowych, której najważniejsze fragmenty przedstawiamy w niniejszym rozdziale.

W modelu izolacji usług, sieć dzielona jest na plastry, w których możliwe jest realizowanie zadanych usług. Przypisanie usług do plastrów można zrealizować na wiele sposobów, w pracy [88] podano następujące kryteria.

- *Kryterium osiągalności* – w ramach danej instancji plastra możliwe jest zrealizowanie danej instancji usługi (odpowiednie węzły sieci wymagane do realizacji usługi do niej należą).
- *Kryterium zasobów* – w ramach danej instancji plastra istnieją zasoby, które można wykorzystać w celu realizacji usługi.
- *Kryterium jakościowe* – w ramach danej instancji plastra realizujemy usługi, które mają zbliżone do siebie wymagania QoS i/lub QoE, aby możliwe było efektywniejsze alokowanie zasobów dla tych usług.
- *Kryterium izolacji* – w ramach danej instancji plastra realizujemy instancje usług na jakie zezwalają polityki biznesowe i polityki bezpieczeństwa (klienta usługi, serwera usług, operatora, itd.).

Warto podkreślić, że przydział usług do plastrów może zmieniać się w czasie, na skutek incydentów w systemie informatycznym, bądź też planowanych zmian w systemie. Możliwe jest również dynamiczne przenoszenie usługi pomiędzy plastrami, aby reagować na zwiększone zapotrzebowanie na zasoby w danym fragmencie systemu, bądź też w celu równoważenia obciążenia (*load balancing*). Do decyzji architekta systemu pozostaje zdefiniowanie odpowiednich polityk opisujących zasady migracji usług pomiędzy plastrami, aby bezpieczeństwo systemu, wymagania QoS/QoE, oraz wymagania kontraktowe nie zostały naruszone.



Rysunek 9.1: Graf przedstawiający instance usług, instance plastrów i parametry z nimi powiązane. Rysunek pochodzi z pracy [88].

9.2. MODEL MATEMATYCZNY IZOLACJI USŁUG

Przejdźmy teraz do modelu matematycznego zaproponowanego w [88], do którego podstawą jest graf dwudzielny kojarzący usługi i plastry przedstawiony na rysunku 9.1. Oznaczmy jako $e_{i,j} \in \{0; 1\}$ stałą, która określa możliwość realizacji instancji usługi I_i (patrz rys. 9.1) w instancji plastra S_j . Załóżmy, że zmienna binarna $x_{i,j}$ oznacza, że instancja usługi I_i jest realizowana w ramach instancji plastra S_j . Mamy w takiej sytuacji następujący układ nierówności:

$$\left(\bigwedge_{i=1}^K \right) \sum_{j=1}^N e_{i,j} x_{i,j} \leq 1. \quad (9.1)$$

Oznaczmy jako $p_j = (p_{j,1}, p_{j,2}, \dots, p_{j,M}) \in \mathbb{R}_{\geq 0}^M$ ciąg parametrów, które opisują gwarantowane przez instancję plastra S_j wielkości sieciowe, wymagania techniczne, itp.; są to analogiczne parametry jak w przypadku innych modeli izolacji rozpatrywanej w tej pracy. Parametr $p_{j,t}$ w ramach instancji plastra S_j spełnia wymaganie $R_{i,t}$ stawiane przez instancję usługi I_i gdy $R_{i,t} \leq p_{j,t}$; w przypadku gdy instancja plastra nie gwarantuje spełnienia danego parametru, to przyjmuje on wartość 0. Realne pomiary (lub wielkości otrzymane w dowolny inny sposób) można znormalizować do konwencji nieujemnych parametrów sieciowych korzystając z funkcji normalizujących opisanych w tabeli 4.1 w rozdziale 4. Dostajemy poniższy układ nierówności [88]:

$$\left(\bigwedge_{i=1}^K \right) \left(\bigwedge_{t=1}^M \right) \sum_{j=1}^N (R_{i,t} - p_{j,t}) e_{i,j} x_{i,j} \leq 0. \quad (9.2)$$

Na podstawie (9.1) zauważamy, że w ramach powyższej sumy będzie występować co najwyżej jeden niezerowy składnik [88]. Autor w pracy [88] zauważa, że nierówność podana w (9.2) może być spełniona dla więcej niż jednego składnika sumy, gdy dana instancja usługi w kontekście danego parametru spełnia zasadę superpozycji. Za przykład takiego parametru podano przepływność binarną - możliwe jest podzielenie strumienia danych na wiele strumieni cząstkowych, realizowanych za pomocą różnych plastrów. Innym typem parametru jest dostępność albo siła szyfrowania (w bitach klucza) kanału komunikacyjnego - dla takich parametrów zasada superpozycji nie zachodzi, więc niemożliwe jest rozbięcie realizacji usługi na wiele plastrów, aby sumarycznie spełnić (w taki prosty sposób) wymagania jakościowe. Na tym etapie można przeprowadzić wstępną analizę, które plastry mają teoretyczną zdolność do realizacji danej usługi, a które ze względu na niespełnianie wymagań jakościowych, należy wykluczyć z rozważań. Zadanie to polega na zamianie wielkości $e_{i,j}$ na wielkość $e'_{i,j}$, które zależą od siebie w następujący sposób [88]:

$$e'_{i,j} = e_{i,j}c_{i,j}, \quad (9.3)$$

gdzie $c_{i,j}$ to stała zawężająca zbiór potencjalnych plastrów w kontekście realizacji usług. W pracy [88] zaproponowano następującą formułę:

$$c_{i,j} = \begin{cases} 1, & \text{gdy } \sum_{t=1}^M \text{sgn}(p_{j,t} - R_{i,t}) = M, \\ 0, & \text{w p.p..} \end{cases} \quad (9.4)$$

Wyznaczenie wartości $c_{i,j}$ wymaga wykonania $O(NKM)$ porównań, do implementacji nie jest konieczne wykonywanie dokładnego sumowania, tylko porównywanie wymagań instancji usług z możliwościami instancji plastrów. W pracy [88] odnotowano, że uwzględnienie $e'_{i,j}$ w miejscu $e_{i,j}$ w nierównościach (9.2) powoduje, że są one zawsze spełnione. Mamy bowiem:

- gdy $c_{i,j} = 0$, to nierówność jest oczywiście spełniona, ponieważ $(R_{i,t} - p_{j,t})e_{i,j} \cdot 0 \cdot x_{i,j} \leq 0$.
- gdy $c_{i,j} = 1$, to dla każdego i, j, t mamy $R_{i,t} - p_{j,t} \leq 0$, a więc $(R_{i,t} - p_{j,t})e_{i,j}x_{i,j} \leq 0$.

Wprowadźmy teraz możliwość zdefiniowania izolacji pomiędzy usługami poprzez wskazanie, które instancje usług nie mogą występować razem ze sobą w ramach jednego plastra - jest to metoda zaczerpnięta z pracy [88]. Oznaczmy jako $f_{i,j}$ stałą binarną, której znaczenie jest następujące: wartość 1 oznacza brak możliwości koegzystencji instancji usługi I_i i I_j w jednej instancji plastra. Wartość 0 oznacza brak przeciwwskazań do koegzystencji tych instancji usług.

Wówczas mamy układ nierówności [88]:

$$\left(\bigwedge_{i=1}^K \right) \left(\bigwedge_{j=i+1}^K \right) \left(\bigwedge_{r=1}^N \right) e'_{i,r} x_{i,r} + e'_{j,r} x_{j,r} + f_{i,j} \leq 2. \quad (9.5)$$

Dla dopełnienia modelu optymalizacyjnego, należy zdefiniować funkcję celu. W pracy [88] podano przykładową funkcję w postaci:

$$\max(Z) = \max \left(\sum_{i=1}^K \left(\mu_i \sum_{j=1}^N e'_{i,j} x_{i,j} \right) \right), \quad (9.6)$$

gdzie μ_i to zysk z realizacji instancji usługi I_i . Możliwe jest bardziej precyzyjne określenie funkcji celu - w oparciu o przychód i koszty związane z realizacją usługi, w tym koszty stałe i zmienne.

9.3. WNIOSKI

Przedstawiony model pokazuje alternatywny sposób opisu izolacji, skupiając się na optymalnym przypisaniu usług do instancji plastrów, z wykorzystaniem pewnej funkcji celu. Widzimy, że techniki programowania matematycznego są zręcznym narzędziem do opisu zagadnień izolacji, zarówno wprost, jak i jako dodatkowe narzędzie, jak w przypadku modelu wykorzystującego skierowane liczby rozmyte. Ze względu na użycie zmiennych całkowitoliczbowych mamy do czynienia z programowaniem całkowitoliczbowym. Ogólne metody rozwiązywania takiego problemu optymalizacyjnego są przedstawione np. w [98]. W pracy [88] zaprezentowano algorytm heurystyczny do rozwiązywania przedstawionego problemu optymalizacyjnego; przyczynkiem do tego był fakt, że problem programowania całkowitoliczbowego jest problemem NP-zupełnym (NP - *Nondeterministic Polynomial*) [47].

Podczas przydziału usługi do instancji plastra, w sposób jawny lub niejawny, należy dokonać wyboru jak realizować usługę, za pomocą jakich zasobów, ukrytych za abstrakcją plastra sieci. Tego typu decyzje mogą być podjęte wcześniej, podczas fazy planowania architektury plastrów i usług, bądź też w sposób dynamiczny, w zależności od aktualnego obciążenia sieci i wykorzystania zasobów. Do wykonywania tego typu zadań może być wykorzystany orkiestrator, który ma wiedzę na temat stanu sieci.

10. Podsumowanie i wnioski

10.1. PODSUMOWANIE WYNIKÓW BADAŃ

W niniejszej rozprawie uzyskano następujące wyniki badań, których celem było udowodnienie postawionej tezy.

- W podrozdziale 3.5. przedstawiono argumenty za stwierdzeniem, że zarządzanie izolacją i jej opis ilościowy jest ważnym elementem zapewniania bezpieczeństwa usług w sieciach plastrowych. Odwołano się do zdefiniowanej w pracy definicji izolacji i wyzwań związanych z bezpieczeństwem usług w sieciach plastrowych. Uzasadniono, dlaczego opis ilościowy jest istotny z punktu widzenia zapewniania bezpieczeństwa.
- W podrozdziale 3.4. zaproponowano definicję izolacji w sieciach plastrowych. W rozdziale 6 zaprezentowano model rozmyty wykorzystujący skierowane liczby rozmyte, który służy do opisu poziomu izolacji w sieci plastrowej. Opis ilościowy izolacji był przedmiotem modeli opisanych w rozdziałach 4 (model bazowy, operujący na liczbach rzeczywistych), 6 i 7 (model rozmyty) oraz 8 (model probabilistyczny).
- Z powodzeniem zastosowano skierowane liczby rozmyte do procedury wyznaczania izolacji, która była określona dla liczb rzeczywistych (ostrych). Oryginalna procedura musiała zostać rozszerzona w niniejszej pracy o następujące elementy.
 - Algorytm wyznaczania właściwej skierowanej liczby rozmytej najbliższej zadanej skierowanej liczbie rozmytej niewłaściwej - opisany w podrozdziale 6.2.. Zaproponowany algorytm może być zastosowany do rozwiązywania innych problemów niż wyznaczenie poziomu izolacji. Przedstawiony pomysł na wyznaczenie liczby właściwej polega na zdefiniowaniu odpowiedniej metryki pomiędzy liczbą właściwą i niewłaściwą, która jest następnie minimalizowana z wykorzystaniem technik optymalizacyjnych, w szczególności programowania liniowego. Ograniczeniami w postawionym zagadnieniu są zależności pomiędzy punktami należącymi do funkcji przynależności do poszukiwanej

Tablica 10.1: Porównanie modeli izolacji

Właściwość	Model rozmyty	Model probabilistyczny	Model izolacji usług
Złożoność modelu	średnia	niska	niska
Uwzględnienie topologii sieci	częściowe	częściowe	brak
Uwzględnienie polityk usług	brak	brak	jest
Opis wymagań usług	za pomocą cech	za pomocą cech	za pomocą cech
Standaryzacja wymagań	poprzez normalizację	poprzez normalizację	poprzez normalizację
Główne wyzwania	Wybór wierzchołków podstawowych, porównywanie wyników rozmytych, interpretacja niewłaściwej SLR, normalizacja parametrów	Wybór wierzchołków podstawowych, dobór zbiorów gdzie jest spełniona izolacja, normalizacja parametrów	normalizacja parametrów
Zastosowanie	Wyznaczenie poziomu izolacji w sieci lub systemie informacyjnym	Wyznaczenie poziomu izolacji w sieci lub systemie informacyjnym	Planowanie wykorzystania zasobów przez usługi
Możliwości rozszerzenia	Alternatywny porządek liniowy, inna metryka do wyznaczenia właściwej SLR, analiza kształtu SLR	Dodatkowe funkcje opisujące zestaw danych	Dodatkowe ograniczenia związane z zasobami, minimalną liczbą instancji danej klasy usług

skierowanej liczby rozmytej; ograniczenia te mają postać nierówności liniowych. Minimalizowana metryka ma postać funkcji nieliniowej, więc sprowadzono zdefiniowany problem nieliniowy do problemu programowania liniowego za pomocą relaksacji funkcji nieliniowej przy pomocy zestawu funkcji liniowych.

- Algorytm wyznaczania skierowania zadanej niewłaściwej liczby rozmytej - opisany w rozdziale 7. Algorytm nie wykorzystuje w tym celu operatora wyostrażania niewłaściwej

liczby rozmytej, tylko wzajemne położenie funkcji f i g składających się na skierowaną liczbę rozmytą wokół jądra skierowanej liczby rozmytej, czyli tam, gdzie funkcja przynależności do danej liczby rozmytej jest równa 1. Ciekawy jest przypadek liczb trójkątnych i innych, których jądro jest zbiorem jednopunktowym - w takiej sytuacji trzeba odwołać się do pozostałych fragmentów funkcji f i g , aby uzyskać skierowanie takiej liczby.

- Sposób określania porządku wśród skierowanych liczb rozmytych - w tym celu wprowadzono porządek liniowy w podrozdziale 6.4., uwzględniając wymagania specyficzne dla izolacji jakie są stawiane temu porządkowi. W szczególności, rozpatrzono sytuację porównywania (ustalania w jakim są porządku) dwóch skierowanych liczb rozmytych oraz porównywania liczby rzeczywistej (ostrej) i skierowanej liczby rozmytej. Są to dwie podstawowe operacje, jakie odbiorca rezultatu wyznaczania poziomu izolacji mógłby chcieć wykonać, w celu porównania aktualnego stanu z poprzednimi stanami oraz porównania aktualnego stanu z wartościami progowymi.

Wymienione elementy nie występowały w przypadku oryginalnego algorytmu, zdefiniowanego dla liczb rzeczywistych, ponieważ są one specyficzne dla liczb rozmytych lub skierowanych liczb rozmytych.

- Przeprowadzono walidację opracowanych algorytmów z wykorzystaniem przykładu obliczeniowego wspomaganego komputerowo - jest to główna zawartość rozdziału 7.
- Dodatkowo stworzono probabilistyczny model opisu poziomu izolacji, który został zaprezentowany w rozdziale 8.

Tablica 10.1 przedstawia porównanie modeli izolacji będących przedmiotem niniejszej rozprawy. Model rozmyty jest bardziej zaawansowany niż model probabilistyczny, chociaż wymaga jednocześnie wykonania bardziej skomplikowanych operacji, jak np. rozwiązanie problemu programowania liniowego w celu wyznaczenia poziomu izolacji. W modelu rozmytym możliwe jest wykorzystanie skierowania liczby rozmytej oraz wdrożenie analizy liczby rozmytej pod kątem jej kształtu. W przypadku modelu probabilistycznego do jego prawidłowego funkcjonowania potrzebne jest określenie zbioru wartości progowych; sam model probabilistyczny można natomiast rozszerzać poprzez dodawanie kolejnych funkcji operujących na danych wejściowych. Model izolacji usług może być wykorzystany przy wspomaganie planowania wykorzystania zasobów przez usługi, jest również otwarty na rozszerzenia o dodatkowe

kryteria związane z wykorzystaniem zasobów. Model probabilistyczny ze względu na niski poziom skomplikowania mógłby być zastosowany w ramach prac rozwojowych nad działaniem samego rdzenia algorytmu związanego z wyznaczaniem izolacji.

10.2. WNIOSKI I KIERUNKI DALSZYCH BADAŃ

Zagadnienie oceny poziomu izolacji w sieciach plastrowych jest istotne z perspektywy Operatora Telekomunikacyjnego, aby prawidłowo realizować usługę dostarczania platformy dla innych dostawców usług. Ze względu na powiązanie izolacji zarówno z zagadnieniami bezpieczeństwa systemów informacyjnych, jak i z zagadnieniem spełniania wymagań kontraktowych wobec usług, możliwe jest rozpatrywanie poziomu izolacji w kontekście innych podmiotów i sieci, niż Operatorzy Telekomunikacyjni i ich sieci. W przypadku dużych przedsiębiorstw, zarządzanie izolacją w systemach informacyjnych jest zadaniem z zakresu architektury korporacyjnej (ang. *Enterprise Architecture*).

Skierowane liczby rozmyte są ciekawą ideą, która pozwala zapisywać informację o wielkościach, co do których nie ma pewności co do wartości, wartość zmienia się w czasie, bądź też jest obciążona z jakiegoś powodu niepewnością. Wykonywanie obliczeń na SLR jest możliwe, chociaż niesie za sobą ryzyko operowania na liczbach rozmytych, co do których nie ma jasnej interpretacji - tzw. niewłaściwych skierowanych liczb rozmytych. W niniejszej pracy zaproponowano metodę wyznaczania skierowanej liczby rozmytej właściwej dla zadanej SLR niewłaściwej. Jest to operacja w pewnym sensie podobna do operacji wyostrzania, która zadanej skierowanej liczby rozmytej przypisuje liczbę ostrą (liczbę rzeczywistą).

W toku realizacji niniejszej pracy zidentyfikowano kilka potencjalnych kierunków dalszych badań, z których warto wymienić następujące.

— Metoda wyznaczania wierzchołków podstawowych. Dobór zbioru wierzchołków podstawowych jest istotny z punktu widzenia zgodności modelu sieci ze stanem rzeczywistym. Zagadnienie jest wieloaspektowe, wymaga dokładnej znajomości topologii sieciowej systemu informatycznego, świadomości różnic pomiędzy założoną architekturą systemu a jego rzeczywistą implementacją. Dodatkową trudnością jest konieczność identyfikacji potencjalnych wektorów ataku, które są powiązane z modelem na poziomie wierzchołków i/lub cech przypisanych wierzchołkom. Dużym wyzwaniem może być odpowiednie ujęcie w

modelu nieznanym dotychczas wektorów ataku na dany system, w szczególności ataków typu *0-day*.

- Wykorzystanie technik AI do analizowania poziomu izolacji. Ze względu na liczebność parametrów, proces zbierania danych pomiarowych powinien być zautomatyzowany, a następnie przetwarzany w czasie zbliżonym do czasu rzeczywistego, aby możliwe było wykorzystanie uzyskanej wiedzy w bieżących zadaniach operacyjnych związanych z zarządzaniem usługami IT w danym podmiocie. Metody z obszaru AI mogą pomóc w analizie zdobytych danych i informacji, wskazać słabe strony aktualnego stanu sieci. Istnieje również możliwość przeprowadzenia procesu trenowania sieci neuronowych w kierunku wyznaczenia poziomu izolacji danej sieci teleinformatycznej lub systemu informacyjnego.
- Wykorzystanie zaawansowanych metod analizy danych do uzyskania informacji o aktualnym poziomie izolacji. Mogą to być metody z obszaru *data mining* lub eksploracji danych, a ich zastosowanie mogłoby uzupełnić model probabilistyczny o dodatkowe funkcje opisujące zbiór danych.
- Dalsze prace nad algorytmami: wyznaczenia skierowania niewłaściwej skierowanej liczby rozmytej, poszukiwania liczby właściwej odpowiadającej liczbie niewłaściwej oraz prace nad sposobem określania porządku liniowego wśród skierowanych liczb rozmytych. Szczególnie pomocne mogłoby być stworzenie algorytmu poszukiwania liczby właściwej, który nie wymagałby wstępnego wyznaczenia skierowania niewłaściwej liczby rozmytej.
- Badania nad interpretacją rozmytej izolacji, szczególnie w formie skierowanej liczby rozmytej niewłaściwej.
- Stworzenie zestawu funkcji normalizujących, który minimalizuje konieczność dostosowywania histogramu wykorzystywanego do stworzenia skierowanej liczby rozmytej.

10.3. PRZEGLĄD OPUBLIKOWANYCH WYNIKÓW BADAŃ

Autor niniejszej publikacji był autorem lub współautorem następujących artykułów naukowych, materiałów pokonferencyjnych i rozdziałów w monografiach. Część z wyników uzyskanych w ramach poniższych publikacji była wykorzystana do stworzenia niniejszej rozprawy. Główne wyniki związane ze skierowanymi liczbami rozmytymi będą przedmiotem nowej publikacji, niewymienionej poniżej.

1. Niewolski, W., Nowak, T.W., Sepczuk, M., Kotulski, Z., Artych, R., Bocianiak, K., Wary, J.-P., **Security Context Migration in MEC: Challenges and Use Cases**, *Electronics*, 2022, 11, 3512, <https://doi.org/10.3390/electronics11213512> . Praca skupia się na zagadnieniu migracji kontekstu bezpieczeństwa pomiędzy sieciami 5G MEC, definiuje wymagania co do tej operacji w kontekście przykładów użycia zaproponowanych w pracy [92]. Zagadnienie to wiąże się z tematem izolacji poprzez fakt, że spełnienie wymagań w dziedzinie izolacji może być warunkiem koniecznym do prawidłowego przeprowadzenia migracji usługi pomiędzy dwoma środowiskami 5G MEC. Praca jest cytowana w rozdziale 1.
2. Sepczuk, M., Kotulski, Z., Niewolski, W., Nowak, T.W., **Low-complexity access control scheme for MEC-based services**, 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS), 2022, s. 673-681, <https://doi.org/10.15439/2022F55> . W pracy zaproponowano nowy protokół kontroli dostępu do usług realizowanych za pomocą sieci 5G MEC. W pracy przedstawiana jest koncepcja *MEC Enablera*, który w ramach swojej odpowiedzialności ma również obszar zarządzania plastrami (*Slice MANO*). Praca jest cytowana w rozdziale 1.
3. Kotulski, Z., Niewolski, W., Nowak, T., Sepczuk, S., Hsu R.-H., Cheng, S.-M., Chang S.-C., Yu, T.-Y., Hung, C.-F., **ACMEC - ACCESS CONTROL IN 5G MEC. JOINT RESEARCH PROJECT OF NTUST (TAIWAN) AND WUT (POLAND)**, *KSTiT 2022, PRZEGLĄD TELEKOMUNIKACYJNY - WIADOMOŚCI TELEKOMUNIKACYJNE*, Rocznik 2022, zeszyt 4, <https://doi.org/10.15199/59.2022.4.15> . Praca przedstawia wyniki badań zrealizowanych w ramach międzynarodowego projektu polsko-tajwańskiego finansowanego przez NCBiR. W szczególności, w pracy zaprezentowano wysokopoziomowy opis *MEC Enablera*, który był kluczowym komponentem w architekturze bezpieczeństwa sieci 5G MEC stworzonej na potrzeby tego projektu. Praca została zacytowana w rozdziale 1.
4. Niewolski, W., Nowak, T.W., Sepczuk, M., Kotulski, Z., **Token-Based Authentication Framework for 5G MEC Mobile Networks**, *Electronics*, 2021, 10, 1724, <https://doi.org/10.3390/electronics10141724> . Celem publikacji było przedstawienie idei uwierzytelnienia za pomocą *MEC Enablera* i tokenów JMAT w środowisku 5G

MEC. Praca zawiera wyniki badania wydajności generowania tokenów wykorzystywanych do uwierzytelnienia. Praca została zacytowana w rozdziale 1.

5. Nowak, T. W., Sepczuk, M., Kotulski, Z. i in., **Verticals in 5G MEC-Use Cases and Security Challenges**, *IEEE Access*, 2021, s.87251-87298, <https://doi.org/10.1109/ACCESS.2021.3088374> Praca zawiera szeroką analizę scenariuszy użycia sieci 5G i 5G MEC w odniesieniu do sektorów wertykalnych, co przyczyniło się do głębszego zrozumienia potrzeb związanych z zapewnianiem izolacji w tego typu sieciach, aby stworzyć przestrzeń do realizacji nowoczesnych usług telekomunikacyjnych. Odniesienia do pracy znajdują się w rozdziale 1.
6. Kotulski, Z., Niewolski, W., Nowak, T.W., Sepczuk, M., **New Security Architecture of Access Control in 5G MEC**, *Thampi, S.M., Wang, G., Rawat, D.B., Ko, R., Fan, C.I. (eds) Security in Computing and Communications. SSCC 2020. Communications in Computer and Information Science*, vol 1364, Springer, Singapore, 10.01.2021, https://doi.org/10.1007/978-981-16-0422-5_6 . Publikacja zawiera propozycję nowego pomysłu realizacji uwierzytelnienia w sieci 5G MEC, analizując potrzeby poszczególnych interesariuszy takiego systemu i ich wymagania co do bezpieczeństwa. Przedstawione są przykładowe przypadki użycia usług w sieci MEC oraz wstępne schematy przepływu w mechanizmie kontroli dostępu. Praca została zacytowana w rozdziale 1.
7. Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M. A., **5G networks: Types of isolation and their parameters in RAN and CN slices**, *Computer Networks (Elsevier)*, 22.04.2020, <https://doi.org/10.1016/j.comnet.2020.107135>. Praca zawiera przegląd technik izolacji stosowanych w systemach teleinformatycznych, który został ujęty w rozdziale 3. W pracy zaprezentowano również model wyznaczania izolacji wykorzystujący liczby rzeczywiste, będący przedmiotem opisu w rozdziale 4 i podstawą do utworzenia modelu wykorzystującego skierowane liczby rozmyte.
8. Nowak, T., **Sieci o strukturze plastrowej**, rozdział w monografii *Prace Seminarium Naukowego Instytutu Telekomunikacji Politechniki Warszawskiej. Praca zbiorowa pod redakcją A. Jakubiaka*, tom 3, 2018. Praca opisuje budowę sieci o strukturze plastrowej, porusza temat izolacji w takich sieciach, oraz przedstawia istotne problemy implementacyjnej takiej sieci. Praca jest cytowana w rozdziale 2.

9. Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M. A., **Graph-based quantitative description of networks' slices isolation**, *The Federated Conference on Computer Science and Information Systems (FedCSIS)*, Poznań, Polska, 9-12.09.2018, https://annals-csis.org/Volume_15/pliks/fedcsis.pdf, <https://doi.org/10.15439/2018F322>. W pracy przedstawiono algorytm wyznaczania izolacji za pomocą liczb rzeczywistych i cech sieciowych powiązanych z hipergrafem badanej sieci. Zaproponowano sposób wyznaczania izolacji dla wierzchołków heterogenicznych. Rozważania podjęte w tym artykule zostały użyte w rozdziale 4.
10. Nowak, T.W., **Przedstawianie parametrów sieciowych z wykorzystaniem skierowanych liczb rozmytych**, *Przegląd Telekomunikacyjny + Wiadomości Telekomunikacyjne*, s.726-729, DOI:10.15199/59.2018.8-9.37. Publikacja naukowa przedstawiona również na konferencji KSTiT w 2018 r.; w pracy poruszono zagadnienie interpretacji parametrów sieciowych za pomocą skierowanych liczb rozmytych. Opisano sposób tworzenia skierowanych liczb rozmytych na podstawie zestawu danych pomiarowych, co zostało wykorzystane w rozdziale 6 do utworzenia modelu rozmytego. Sposób obejmuje zarówno stworzenie funkcji przynależności do skierowanej liczby rozmytej, jak i otrzymania funkcji składowych skierowanej liczby rozmytej oraz wyznaczenia jej skierowania w oparciu o dane wejściowe.
11. Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M. A., Artych, R., Bocianiak, K., Osko, T., Wary, J.-P., **Towards constructive approach to end-to-end slice isolation in 5G networks**, *EURASIP Journal of Information Security*, 2018: 2, s.1-23, 2018. <https://doi.org/10.1186/s13635-018-0072-0>. W pracy skupiono się na opisanu zagadnień związanych z sieciami plastrowymi, pełnym połączeniem plastrowym, izolacją w sieciach plastrowych włącznie ze scenariuszem E2E oraz wypisaniem zagadnień wymagających rozwiązania w kontekście sieci plastrowych w ich implementacji w sieciach typu 5G. Materiał z tej publikacji był podstawą do opisu stanu wiedzy w rozdziałach 2 i 3.
12. Nowak, T., **Matematyczny model izolacji usług w sieciach plastrowych**, *Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne*, vol. 8-9, 09.2017. DOI: 10.15199/59.2017.8-9.50. Indywidualna publikacja naukowa przedstawiona również na konferencji KSTiT w 2017 r.; w pracy poruszono zagadnienie przydziału zasobów sieciowych do usług w sposób zapewniający izolację. Koncepcja ta została zaprezentowana w rozdziale 9.

13. Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M., Artych, R., Bocianiak, K., Osko, T., Wary, J.-P., **On end-to-end approach for slice isolation in 5G networks. Fundamental challenges**, *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems*, w: M. Ganzha, L. Maciaszek, M. Paprzycki (Eds). ACSIS, vol.11, s.783-792, 2017, <https://doi.org/10.15439/2017F228>. Tematem pracy było przedstawienie zagadnienia izolacji jako wyzwania stojącego przed podziałem sieci 5G na plastry. Zaprezentowano problemy i wyzwania związane z bezpieczeństwem w kontekście izolacji oraz towarzyszące im wyzwania stojące przed sieciami 5G. Praca jest cytowana w rozdziale 1.
14. Nowak, T., **Przegląd metod realizacji generatorów losowych ciągów binarnych**, *rozdział w monografii Prace Seminarium Naukowego Instytutu Telekomunikacji Politechniki Warszawskiej. Praca zbiorowa pod redakcją A. Jakubiaka*, tom 2, 12.2016. W pracy opisano generatory losowych ciągów binarnych, w tym metody sprzętowe. Efektem pracy było poszerzenie wiedzy w zakresie generacji losowych ciągów binarnych i metodach wpływu na tego typu generatory. Na podstawie tej pracy można wysnuć wniosek, że kryptograficznie bezpieczne rozwiązania sprzętowe wprowadzają do strumienia danych pewną zależność od świata zewnętrznego, co oznacza, że idealnie zaizolowany strumień danych może mieć techniczne trudności w implementacji wielu algorytmów i protokołów kryptograficznych ze względu na brak źródła entropii. Praca była zacytowana w rozdziale 3.

BIBLIOGRAFIA

- [1] 5G Americas White Paper, **Network Slicing for 5G Networks & Services**, 11.2016, https://www.5gamericas.org/wp-content/uploads/2019/07/5G_Americas_Network_Slicing_11.21_Final.pdf
- [2] 5G ENSURE, 2016, www.5gensure.eu
- [3] 5G-NORMA, **Functional Network Architecture and Security Requirements**, *5G-NORMA Deliverable D3.1*, 31.12.2015, https://www.it.uc3m.es/wnl/5gnorma/pdf/5g_norma_d3-1.pdf
- [4] 5G PPP Architecture Working Group, **View on 5G Architecture (Version 4.0)**, 10.2021, <https://doi.org/10.5281/zenodo.5155657>
- [5] Afolabi, I., Prados-Garzon, J., Baga, M., Taleb, T., Ameigeiras, P., **Dynamic Resource Provisioning of a Scalable E2E Network Slicing Orchestration System**, *IEEE Transactions on Mobile Computing*, vol. 19, no. 11, s. 2594-2608, 01.11.2020, <https://arxiv.org/ftp/arxiv/papers/2201/2201.03997.pdf>, DOI: 10.1109/TMC.2019.2930059
- [6] Agencja Bezpieczeństwa Wewnętrznego, **Bezpieczeństwo fizyczne**, <https://bip.abw.gov.pl/bip/informacje-niejawne-1/nadzor-nad-systemem-oc/bezpieczenstwo-fizyczn/150,Kancelarie-tajne-Bezpieczenstwo-fizyczne.html#2>
- [7] Arfaoui, G., Vilchez, J. M. S., Wary, J.-P., **Security and Resilience in 5G: Current Challenges and Future Directions**, *IEEE Trustcom/BigDataSE/ICSS*, Sydney, Australia, s. 1010 - 1015, 2017, DOI: 10.1109/Trustcom/BigDataSE/ICSS.2017.345
- [8] Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M., **Above the Clouds: A Berkeley View of Cloud Computing**, Technical Report No. UCB/EECS-2009-28, <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html>, 10.02.2009
- [9] Aso, O., Tadakuma, M., Namiki, S., **Four-Wave Mixing in Optical Fibers and Its Applications**, *Furukawa Review*, nr 19, 2000, https://www.furukawa.co.jp/review/fr019/fr19_12.pdf

- [10] Blenk, A., Basta, A., Reisslein, M., **Survey on network virtualization hypervisors for software defined networking**, *IEEE Communications Surveys and Tutorials*, vol. 18, no. 1, 2016, s. 655-685
- [11] Bulakci, O., **Towards sustainable 5G Networks. Vision & Design Principles for New Horizons**, *IEEE Vehicular Technology Conf.*, Boston, 2015
- [12] Cabaj, K., Gawkowski, P., **HoneyPot systems in practice**, <http://pe.org.pl/articles/2015/2/16.pdf>, *Przegląd Elektrotechniczny*, ISSN 0033-2097, r. 91, nr 2/2015, DOI:10.15199/48.2015.02.16
- [13] Chai W., **What is the CIA triad (confidentiality, integrity and availability)?**, <https://www.techtarget.com/whatis/definition/Confidentiality-integrity-and-availability-CIA>
- [14] Chapman, C., Ward, S., **Description of Network Slicing Concept**, https://ngmn.org/wp-content/uploads/Publications/2016/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf, *NGMN Alliance*, 2016
- [15] CHARISMA, 2016, <http://www.charisma5g.eu/>
- [16] Cherrier, S., Ghamri-Doudane, Y.M., Lohier, S., Roussel, G., **Fault-recovery and Coherence in Internet of Things Choreographies**, *IEEE World Forum on Internet of Things (WF-IoT)*, s.532-537, Seoul 2014, DOI: 10.1109/wf-iot.2014.6803224
- [17] Chirivella-Perez, E., Salva-Garcia, P., Sanchez-Navarro, I., Alcaraz-Calero, J. M., Wang, Q., **E2E network slice management framework for 5G multi-tenant networks**, *Journal of Communications and Networks*, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10146576>, DOI: 10.23919/JCN.2023.000019
- [18] Chwastyk, A., Kosinski, W., **Fuzzy calculus with applications**, *Mathematica applicanda*, vol. 41(1), p. 47-96, 2013, DOI:10.14708/ma.v41i1.380
- [19] Del Piccolo, V., Amamou, A., Haddadou, K., Pujolle, G., **A Survey of Network Isolation Solutions for Multi-Tenant Data Centers**, *IEEE Comm. Surveys and Tutorials*, 18(4), s.2787-2821, 2016, DOI:10.1109/COMST.2016.2556979
- [20] Docker Inc., **Reference documentation**, <https://docs.docker.com/reference/>
- [21] Dubois, D., Prade, H., **Operations on fuzzy numbers**, *International Journal of Systems Science*, 9:6, 613-626, 1978, DOI:10.1080/00207727808941724
- [22] Elgala, H., Little, T. D. C., **P-OFDM: Spectrally Efficient Unipolar OFDM**, *Optical Fiber Communication Conference, OSA Technical Digest (online)* (Optica Publishing Group, 2014), paper Th3G.7. <https://opg.optica.org/abstract.cfm?URI=OFC-2014-Th3G.7>, DOI:10.1364/OFC.2014.Th3G.7

- [23] Ericsson, **5G systems - Enabling industry and society transformation**, *Ericsson White Paper*, UEN 284 23-3244, 2015
- [24] ETSI, **Why do we need 5G?** <https://www.etsi.org/technologies/5g>
- [25] ETSI, **ETSI GR NFV-MAN 001 V1.2.1 (2021-12) - Network Functions Virtualisation (NFV); Management and Orchestration; Report on Management and Orchestration Framework**, 2021, https://www.etsi.org/deliver/etsi_gr/NFV-MAN/001_099/001/01.02.01_60/gr_NFV-MAN001v010201p.pdf
- [26] ETSI GS NFV-IFA 009 V1.1.1 (2016-07) **Network Functions Virtualisation (NFV); Management and Orchestration; Report on Architectural Options**, https://www.etsi.org/deliver/etsi_gs/nfv-ifa/001_099/009/01.01.01_60/gv_nfv-ifa009v010101p.pdf
- [27] European Parliament, **Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)** <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022L2555&qid=1685906926013>
- [28] European Parliament, **Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)**, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02016R0679-20160504>
- [29] Evans, M. J., Rosenthal, J., S., **Probability and Statistics. The Science of Uncertainty. Second Edition.**, University of Toronto, <https://www.utstat.toronto.edu/mikevans/jeffrosenthal/book.pdf>
- [30] Fournet, C., Keller, Ch., Laporte, V., **A Certified Compiler for Verifiable Computing**, *2016 IEEE 29th Computer Security Foundations Symposium (CSF)*, s. 268-280, DOI:10.1109/CSF.2016.26, 2016
- [31] Furuhashi, R., Nakao, A., **Opentag: Tag-based network slicing for wide-area coordinated in-network packet processing**, *2011 IEEE International Conference on Communications Workshops (ICC)*, IEEE, 2011. DOI:10.1109/iccw.2011.5963588
- [32] Ghormley, D.P., Rodrigues, S.H., Petrou, D., Anderson, T.E., **SLIC: An Extensibility System for Commodity Operating Systems**, *USENIX Annual Technical Conference*, s. 39-52, 1998

- [33] Google, IPv6 - Google, <https://www.google.com/intl/en/ipv6/statistics.html>, dostę 21.05.2023
- [34] Gutz, S., Story, A., Schlesinger, C., Foster, N., **Splendid Isolation: A Slice Abstraction for Software-Defined Networks**, *HotSDN'12*, 13.08.2012, Helsinki, Finlandia, <https://doi.org/10.1145/2342441.2342458>.
- [35] Harel, R., Babbage, S., **5G security recommendations Package 2: Network Slicing**, *NGMN Alliance*, 2016.
- [36] Henry, S., Alsohaily, A., Sousa, E. S., **5G is Real: Evaluating the Compliance of the 3GPP 5G New Radio System With the ITU IMT-2020 Requirements**, *IEEE Access*, vol. 8, s. 42828-42840, 2020, DOI:10.1109/ACCESS.2020.2977406
- [37] Herzog, U., i in., **Quality of service provision and capacity expansion through extended-DSA for 5G**, *Transactions of Emerging Telecommunications Technologies*, 27(9), s. 1250 - 1261, 2016, DOI:10.1109/EuCNC.2016.7561032
- [38] Hoopes, J., **Virtualization for Security including Sandboxing, Disaster Recovery, High Availability, Forensic Analysis, and Honeypotting**, Burlington, 2009
- [39] Huawei, **6G: The Next Horizon. From Connected People and Things to Connected Intelligence**, 2021, <https://www-file.huawei.com/-/media/corp2020/pdf/tech-insights/1/6g-white-paper-en.pdf>
- [40] IEEE, **IEEE Standard for Floating-Point** - IEEE Std 754-2008. 2008.
- [41] IEEE Standards Board, **Information technology- Telecommunications and information exchange between systems- Local and metropolitan area networks- Specific requirements- Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications**, 06.1997, DOI: 10.1109/IEEESTD.1997.85951
- [42] International Organization for Standardization, **ISO/IEC 11889-1:2009 – Information technology – Trusted Platform Module – Part 1: Overview**, 05.2009
- [43] Iovanna, P. i in., **End-to-end network slicing orchestration – A Key Enabler for Industry-Vertical use Cases**, *Ericsson Technology Review*, vol. 2022, no. 2, s. 2-10, 2022, <https://www.ericsson.com/49475b/assets/local/reports-papers/ericsson-technology-review/docs/2022/end-to-end-network-slicing-orchestration.pdf>, DOI: 10.23919/ETR.2022.9904695
- [44] Clarkson Virtualization Center, **Isolation Benchmark Suite**, 2012, <http://web2.clarkson.edu/class/cs644/isolation>

- [45] Jindal, K., Dalal, S., Sharma, K. K., **Analyzing Spoofing Attacks in Wireless Networks**, *2014 Fourth International Conference on Advanced Computing & Communication Technologies*, Rohtak, India, 2014, s. 398-402, DOI:10.1109/ACCT.2014.46.
- [46] Kalil, M., i in., **Wireless resource virtualization: opportunities, challenges, and solutions**, *Wireless Communications and Mobile Computing*, 2016; 16:2690-2699, DOI:10.1002/wcm.2717
- [47] Karp, R. M., **Reducibility among combinatorial problems**, *Complexity of Computer Computations*, 1972, 85–103
- [48] Kokku, R., Mahindra, R., Zhang, H., Rangarajan, S., **CellSlice: Cellular Wireless Resource Slicing for Active RAN Sharing**, *IEEE*, 2013
- [49] Korovkin, P.P., **Inequalities**, *Moscow : Mir Publishers*, 1975
- [50] Kosinski, W., **On fuzzy number calculus**, *Int. J. Appl. Math. Comput. Sci.*, vol. 16, no. 1, p. 51-57, 2006
- [51] Kosinski, W., Prokopowicz, P., **Algebra of fuzzy numbers**, *Mathematica Applicanda*, Vol. 5/46, s. 37-63, 2004, DOI:10.14708/ma.v32i46/05.1237
- [52] Kosinski, W., Prokopowicz, P., Slezak, D., **Ordered fuzzy numbers**, *Bulletin of the Polish Academy of Sciences*, Ser. Sci. Math., Vol. 51(3), s. 327-338, 2003
- [53] Kosinski, W., Prokopowicz, P., Slezak, D., **On Algebraic Operations on Fuzzy Numbers**, *Intelligent Information Processing and Web Mining*, 22 : 353-362, 2003, DOI:10.1007/978-3-540-36562-4_37
- [54] Kotulski, Z., Niewolski, W., Nowak, T.W., Sepczuk, M., **New Security Architecture of Access Control in 5G MEC**, *Thampi, S.M., Wang, G., Rawat, D.B., Ko, R., Fan, C.I. (eds) Security in Computing and Communications. SSCC 2020. Communications in Computer and Information Science*, vol 1364, Springer, Singapore, 10.01.2021, https://doi.org/10.1007/978-981-16-0422-5_6
- [55] Kotulski, Z., Niewolski, W., Nowak, T., Sepczuk, S., Hsu R.-H., Cheng, S.-M., Chang S.-C., Yu, T.-Y., Hung, C.-F., **ACMEC - ACCESS CONTROL IN 5G MEC. JOINT RESEARCH PROJECT OF NTUST (TAIWAN) AND WUT (POLAND)**, *KSTiT 2022, PRZEGLĄD TELEKOMUNIKACYJNY - WIADOMOŚCI TELEKOMUNIKACYJNE*, Rocznik 2022, zeszyt 4, <https://doi.org/10.15199/59.2022.4.15>
- [56] Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M. A., **5G networks: Types of isolation and their parameters in RAN and CN slices**, *Computer Networks (Elsevier)*, 22.04.2020, <https://doi.org/10.1016/j.comnet.2020.107135>

- [57] Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M. A., **Graph-based quantitative description of networks' slices isolation**, *The Federated Conference on Computer Science and Information Systems (FedCSIS)*, Poznań, Polska, 9-12.09.2018, , https://annals-csis.org/Volume_15/pliki/fedcsis.pdf, <https://doi.org/10.15439/2018F322>
- [58] Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M., Artych, R., Bocianiak, K., Osko, T., Wary, J.-P., **On end-to-end approach for slice isolation in 5G networks. Fundamental challenges**, *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems*, w: M. Ganzha, L. Maciaszek, M. Paprzycki (Eds). ACSIS, vol.11, s.783-792, 2017, <https://doi.org/10.15439/2017F228>
- [59] Kotulski, Z., Nowak, T., Sepczuk, M., Tunia, M. A., Artych, R., Bocianiak, K., Osko, T., Wary, J.-P., **Towards constructive approach to end-to-end slice isolation in 5G networks**, *EURASIP Journal of Information Security*, 2018: 2, s.1-23, 2018. <https://doi.org/10.1186/s13635-018-0072-0>
- [60] Kotulski, Z., Zwierko, A., **Security of mobile code** w D. Taniar (Ed.), *Mobile Computing: Concepts, Methodologies, Tools, and Applications*, Idea Group Inc. Hershey, PA 17033, USA, 2008, s. 2583–2599, DOI:10.4018/978-1-60566-054-7.ch197
- [61] Lachtar, N., Elkhail, A. A., Bacha, A., Malik, H., **A Cross-Stack Approach Towards Defending Against Cryptojacking**, *IEEE Computer Architecture Letters*, vol. 19, no. 2, s. 126-129, 2020, <https://par.nsf.gov/servlets/purl/10220294>, DOI:10.1109/LCA.2020.3017457.
- [62] Leśniewicz, M., **Sprzętowa generacja losowych ciągów binarnych**, *Wojskowa Akademia Techniczna*, 2009
- [63] Li, K., Wang, Ch., Shen, Q., Wang, J., Qi, Q., **MVNC: A SDN-based Multi-tenant Virtual Network Customization Mechanism in Cloud Data Center**, *2016 International Conference on Networking and Network Applications*, s.239-244, DOI:10.1109/NaNA.2016.51
- [64] Li, Q., Wu G., Papathanassiou, A., Wei, L., **End-to-end Network Slicing in 5G Wireless Communication Systems**, *Proc. ETSI Workshop on Future Radio Technologies and Air Interfaces*, s.1-4, 2016, https://docbox.etsi.org/Workshop/2016/201601_FUTURE_RADIO_TECHNOL_WORKSHOP/S06_ADVANCED_TOPIC_FUTURE_AIR_INTERFACES/END2END_NWK_SLICING_5G_WIRELESS_COMM_SYSTEMS_paper.pdf
- [65] Łachwa, A., **Rozmyty świat zbiorów, liczb, relacji, faktów reguł i decyzji**, *Akademicka Oficyna Wydawnicza EXIT*, 2001, ISBN 83-87674-21-4

- [66] Mamatas, L., Clayman, S., Galis, A., **Information Exchange Management as a Service for Network Function Virtualization Environments**, *IEEE Trans. Network and Service Management*, 13(3), s.564-577, 2016, DOI:10.1109/TNSM.2016.2587664
- [67] Manes, V.J.M., Jang, D., Ryu, Ch., Kang, B.B., **Domain isolated kernel: a lightweight sandbox for untrusted kernel extensions**, *Computers & Security*, 2018, DOI:10.1016/j.cose.2018.01.009
- [68] Marszałek, A., **Skierowane liczby rozmyte w modelowaniu i symulacji finansowych szeregów czasowych**, rozprawa doktorska, *Polska Akademia Nauk*, 2017, https://www.ippt.pan.pl/repository/doktoraty/open/2017marszalek_a_doktorat.pdf
- [69] Marszałek, A., Burczyński Tadeusz, **Modeling and forecasting financial time series with ordered fuzzy candlesticks**, *Information Sciences*, 2014, 273 : 144-155. DOI:10.1016/j.ins.2014.03.026
- [70] Matthew, H., **Intel SGX for Dummies (Intel SGX Design Objectives)**, 26.09.2013, zakt. 30.11.2015
- [71] Microsoft, **Azure for US Government**, <https://azure.microsoft.com/en-us/explore/global-infrastructure/government/#why-azure>, dostęp 12.04.2023
- [72] Microsoft, **Azure Government Top Secret now generally available for US national security missions**, <https://azure.microsoft.com/en-us/blog/azure-government-top-secret-now-generally-available-for-us-national-security-missions/>, dostęp 12.04.2023
- [73] Microsoft, **Cross-region replication in Azure: Business continuity and disaster recovery**, <https://learn.microsoft.com/en-us/azure/reliability/cross-region-replication-azure>, 25.02.2023
- [74] Microsoft, **Intel SGX based confidential computing VMs now available on Azure Dedicated Hosts**, <https://azure.microsoft.com/en-us/updates/intel-sgx-based-confidential-computing-vms-now-available-on-azure-dedicated-hosts/>, 11.12.2020
- [75] Microsoft, **What are Azure regions and availability zones?**, <https://learn.microsoft.com/en-us/azure/reliability/availability-zones-overview>, 15.03.2023
- [76] Microsoft, **What is infrastructure as code (IaC)?**, <https://learn.microsoft.com/en-us/devops/deliver/what-is-infrastructure-as-code>
- [77] Mohammedali, N. A., Kanakis, T. , Agyeman, M. O., Al-Sherbaz, A., **A Survey of Mobility Management as a Service in Real-Time Inter/Intra Slice Control**, *IEEE Access*, vol. 9, s. 62533-62552, 2021, DOI:10.1109/ACCESS.2021.3074024

- [78] Moreira, R., de Oliveira Silva, F., **Towards 6G Network Slicing**, <https://sol.sbc.org.br/index.php/w6g/article/download/17231/17069>
- [79] Munoz-Escoi, F.D., Bernabeu-Auban, J.M., **A survey on elasticity management in PaaS systems**, *Computing*, vol. 99, s. 617-656, 2017 DOI:10.1007/s00607-016-0507-8
- [80] Munoz, R., i in., **Integrated SDN/NFV Management and Orchestration Architecture for Dynamic Deployment of Virtual SDN Control Instances for Virtual Tenant Networks**, *J. Optical Comm. and Networking*, 7(11), s.B62-B70, 2015, DOI:10.1364/jocn.7.000b62
- [81] Nakao, A., i in., **End-to-end Network Slicing for 5G Mobile Networks**, *Journal of Information Processing*, vol.25, s. 153 - 163, 2017, DOI:10.2197/ipsjjip.25.153
- [82] National Institute of Information and Communications Technology (NICT), **Beyond 5G/6G White Paper (English version 2.0, 06.2022)**, https://beyond5g.nict.go.jp/images/download/NICT_B5G6G_WhitePaperEN_v2_0.pdf
- [83] Nejabati, R., Peng, S., Channegowda, M., Guo, B., Simeonidou, D., **SDN and NFV Convergence a Technology Enabler for Abstracting and Virtualising Hardware and Control of Optical Networks**, *Optical Fiber Comm. Conf. and Exhib. (OFC)*, 2015, DOI:10.1364/ofc.2015.w4j.6
- [84] Niewolski, W., Nowak, T.W., Sepczuk, M., Kotulski, Z., **Token-Based Authentication Framework for 5G MEC Mobile Networks**, *Electronics*, 2021, 10, 1724, <https://doi.org/10.3390/electronics10141724>
- [85] Niewolski, W., Nowak, T.W., Sepczuk, M., Kotulski, Z., Artych, R., Bocianiak, K., Wary, J.-P., **Security Context Migration in MEC: Challenges and Use Cases**, *Electronics*, 2022, 11, 3512, <https://doi.org/10.3390/electronics11213512>
- [86] Nocedal, J., Wright, S. J., **Numerical Optimization**, *Springer New York*, NY, ISBN 978-0-387-40065-5, 2006, <https://doi.org/10.1007/978-0-387-40065-5>
- [87] Nokia, **Dynamic end-to-end network slicing for 5G**, *Nokia White Paper*, 2016
- [88] Nowak, T., **Matematyczny model izolacji usług w sieciach plastrowych**, *Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne*, vol. 8-9, 09.2017. DOI: 10.15199/59.2017.8-9.50
- [89] Nowak, T.W., **Przedstawianie parametrów sieciowych z wykorzystaniem skierowanych liczb rozmytych**, *Przegląd Telekomunikacyjny + Wiadomości Telekomunikacyjne*, s.726-729, DOI:10.15199/59.2018.8-9.37
- [90] Nowak, T., **Przegląd metod realizacji generatorów losowych ciągów binarnych**, *rozdział w monografii Prace Seminarium Naukowego Instytutu Telekomunikacji Politechniki Warszawskiej. Praca zbiorowa pod redakcją A. Jakubiaka*, tom 2, 12.2016

- [91] Nowak, T., **Sieci o strukturze plastrowej**, rozdział w monografii *Prace Seminarium Naukowego Instytutu Telekomunikacji Politechniki Warszawskiej. Praca zbiorowa pod redakcją A. Jakubiaka*, tom 3, 2018
- [92] Nowak, T. W., Sepczuk, M., Kotulski, Z. i in., **Verticals in 5G MEC-Use Cases and Security Challenges**, *IEEE Access*, 2021, s.87251-87298, <https://doi.org/10.1109/ACCESS.2021.3088374>
- [93] Open Networking Foundation, ONF TR-521, **SDN Architecture**, Version 1.1, 2016
- [94] Oppenheim, A. V., Willsky, A. S., Nawab, S. H., **Signals & systems (2nd ed.)**, *Prentice-Hall, Inc.*, 1996, ISBN:978-0-13-814757-0
- [95] Pardalos, P. M., Vavasis, S. A., **Quadratic Programming with One Negative Eigenvalue is NP-Hard**, *Journal of Global Optimization*, 1, 15-22, 1991. DOI:10.1007/bf00120662
- [96] Peterson, L., Anderson, T., Culler, D., Roscoe, T., **A blueprint for introducing disruptive technology into the Internet**, *ACM SIGCOMM Computer Communication Review*, 33(1), s.59-64, 2003, <https://www.cs.princeton.edu/courses/archive/fall03/cs597B/handouts/pdn02-001.pdf>, DOI:10.1145/774763.774772
- [97] Pierce, B.C., **Type Systems for Programming Languages**, The MIT Press, 2002
- [98] Pióro, M., Medhi D., **Routing, Flow, and Capacity Design in Communication and Computer Networks**, *Morgan Kaufmann*, 2004, ISBN:9780125571890, DOI:10.1016/B978-0-12-557189-0.70005-X
- [99] Prokopowicz, P., i in., **Theory and Applications of Ordered Fuzzy Numbers. A Tribute to Professor Witold Kosinski**, *Studies in Fuzziness and Soft Computing* book series (STUDFUZZ, Vol. 356), Springer, Cham 2017. DOI:10.1007/978-3-319-59614-3
- [100] Radiocommunication Sector of International Telecommunication Union (ITU-R), **Recommendation ITU-R M.2083-0. IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond**, 09.2015, https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf
- [101] Rani, P., Chauhan, V., Kumar, S., Sharma, D., **A Review on Wireless Propagation Models**, *International Journal of Engineering and Innovative Technology (IJEIT)* Volume 3, Issue 11, 05.2014
- [102] Richart, M., Baliosian, J., Serrat, J., Gorricho, J.-L., **Resource Slicing in Virtual Wireless Networks: A Survey**, *IEEE Trans. Network and Service Management*, 13(3), s. 462 - 476, 2016, DOI:10.1109/TNSM.2016.2597295
- [103] **Rozporządzenie Prezesa Rady Ministrów z dnia z dnia 22 grudnia 2011 r. w sprawie sposobu oznaczania materiałów i umieszczania na nich klauzul tajności**, *Dziennik Ustaw*, nr 288, poz.

- 1692, <https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20112881692/O/D20111692.pdf>
- [104] Sahni, S., **Computationally Related Problems**, *SIAM J. Comput.*, vol. 3, no. 4, 1974, DOI:10.1137/0203021
- [105] Santos, J. F. i in., **Breaking Down Network Slicing: Hierarchical Orchestration of End-to-End Networks**, *IEEE Communications Magazine*, vol. 58, no. 10, s. 16-22, 10.2020, <https://www.marquez-barja.com/images/papers/breaking20-authorversion.pdf>, DOI:10.1109/MCOM.001.2000406
- [106] Sarkar, T.K., i in. **A Survey of Various Propagation Models for Mobile Communication**, *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 3, 06.2003
- [107] Schneider, F.B., Morrisett, G., Harper, R., **A language-based approach to security. Informatics: 10 Years Back, 10 Years Ahead**, *Lecture Notes in Computer Science*, Vol. 2000, Springer-Verlag, Heidelberg, 86-101
- [108] Sepczuk, M., Kotulski, Z., Niewolski, W., Nowak, T.W., **Low-complexity access control scheme for MEC-based services**, 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS), 2022, s. 673-681, <https://doi.org/10.15439/2022F55>
- [109] ServiceNow, **Segregating and securing data with domain separation**, <https://docs.servicenow.com/bundle/utah-platform-security/page/administer/company-and-domain-separation/reference/bp-segregate-secure.html>, 02.02.2023
- [110] ServiceNow, **Understanding domain separation**, 06.01.2023 https://docs.servicenow.com/bundle/tokyo-platform-security/page/administer/company-and-domain-separation/concept/c_DomainSeparation.html
- [111] Shannon, C.E., **A Mathematical Theory of Communication**, The Bell System Technical Journal, Vol. 27, s. 379–423, 623–656, July, October, 1948. <https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>
- [112] Shimojo, T., Takano, Y., Khan, A., Kaptchouang, S., Tamura, M., Iwashina, Sh., **Future mobile core network for efficient service operation**, *Proc. 1st IEEE Conference on Network Softwarization (NetSoft)*, s.1-6, 2015, DOI:10.1109/NETSOFT.2015.7116190
- [113] SliceNet, 2017, <https://5g-ps.eu/slicenet/>
- [114] Stanik, A., Koerner, M., Kao, O., **Service-level agreement aggregation for quality of service-aware federated cloud networking**, *IET Networks*, 4(5), s.264-269, 2015, DOI:10.1049/iet-net.2014.0104

- [115] Stirzaker D., **Elementary Probability**, *Cambridge University Press*, 2003, ISBN-13 978-0-511-07574-2, https://www.ctanujit.org/uploads/2/5/3/9/25393293/_elementary_probability.pdf
- [116] The Linux Foundation, **Kubernetes documentation**, <https://kubernetes.io/docs/home/>
- [117] **Ustawa z dnia 5 sierpnia 2010 r. o ochronie informacji niejawnych**, <https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20101821228/U/D20101228Lj.pdf>
- [118] Viswanathan, A., Neuman, B.C., **A survey of isolation techniques**, *Univ. Southern California, Inf. Sc. Ins.*, 2009
- [119] Wahbe, R., Lucco, S., Anderson, T.E., Graham, S.L., **Efficient software-based fault isolation**, *Proceedings of the 14th ACM Symposium on Operating Systems Principles*, s. 203-216, 12.1993
- [120] Wang, W., Wang, Z., **Total orderings defined on the set of all fuzzy numbers**, *Fuzzy Sets and Systems*, vol. 243, 2014, s. 131-141, ISSN 0165-0114, <https://doi.org/10.1016/j.fss.2013.09.005>
- [121] Wang, Z., Klir, G. J., *Fuzzy Measure Theory*, 1992, ISBN 978-1-4757-5303-5.
- [122] Wu, W., Zhou, C., Li, M., Wu, H., Zhou, H., Zhang, N., Shen, X., Zhuang, W., **AI-Native Network Slicing for 6G Networks**, <https://arxiv.org/pdf/2105.08576.pdf>
- [123] Zadeh, L., **Fuzzy sets**, *Information and Control*, Vol. 8, s. 338-353, 1965
- [124] Zegzhda, D.P., i in., **Use of Intel SGX to Ensure the Confidentiality of Data of Cloud Users**, *Automatic Control and Computer Sciences*, 2017, Vol. 51, No. 8, s. 848–854
- [125] Zhang, X., Jia, M., Chen, L., Ma, J., Qiu, J., **Filtered-OFDM - Enabler for Flexible Waveform in the 5th Generation Cellular Networks**, *2015 IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, USA, 2015, s. 1-6, DOI:10.1109/GLOCOM.2015.7417854
- [126] Zhao, C., i in., **Quantifying the Isolation Characteristics in Container Environments**, X. Shi i in.: *Network and Parallel Computing. NPC 2017. Lecture Notes in Computer Science*, vol.10578, Springer, https://doi.org/10.1007/978-3-319-68210-5_17
- [127] Zhou, Z., Yu, M., Gligor, V.D., **Dancing with Giants: Wimpy Kernels for On-Demand I/O Isolation**, *IEEE Security & Privacy*, Volume 13, Issue 2, s. 38-46, 2015
- [128] Zumelzu, N., Bedregal, B., Mansilla, E., Bustince, H., Diaz, R., **Admissible orders on fuzzy numbers**, *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 11, str. 4788-4799, 11.2022, <https://arxiv.org/pdf/2003.01530.pdf>, DOI:10.1109/TFUZZ.2022.3160326

WYKAZ SKRÓTÓW

5G	5 th generation
6G	6 th generation
BER	Bit Error Rate
BFSI	Banking, Financial Services, and Insurance
BSS	Business Support System
CDM	Code Division Multiplexing
CDMA	Code-Division Multiple Access
CDN	Content Delivery Network
CIA	Confidentiality, Integrity, Availability
CN	Core Network
COG	the Center Of Gravity
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
(D)DoS	Distributed Denial of Service
DECT	Digital Enhanced Cordless Telephony
DNS	Domain Name System
DTLS	Datagram Transport Layer Security
E2E	End-to-End
EM	Element Management
ETSI	European Telecommunications Standards Institute
FDM	Frequency Division Multiplexing
FDMA	Frequency-Division Multiple Access
f-OFDM	filtered-Orthogonal Frequency Division Multiplexing
GDPR	General Data Protection Regulation
gNB	gNodeB
IaaS	Infrastructure-as-a-Service
ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IMT	International Mobile Telecommunications-2020

IP	Internet Protocol
IPSec	Internet Protocol Security, IP Security
ISO/OSI RM	ISO Open Systems Interconnection Reference Model
IT	Information Technology
ITSM	Information Technology Service Management
ITU	International Telecommunication Union
KKT	warunki Karush-Kuhn-Tuckera
LAN	Local Area Network
LP	Linear Programming problem
MAC	Medium Access Control
MANO	Management and Orchestration
MEC	Mobile Edge Computing/Multi-access Edge Computing
MIMO	Multiple Input, Multiple Output
mMTC	massive Machine Type Communication
MMS	Multimedia Messaging Service
MNVO	Mobile Virtual Network Operator
MPLS	Multiprotocol Label Switching
MPPE	Microsoft Point-to-Point Encryption
NBI	Northbound Interface
NFVI	Network Function Virtualization Infrastructure
NFV	Network Function Virtualization
NFVO	Network Function Virtualization Orchestrator
NGMN	Next Generation Mobile Networks
NIS2	Network and Information Security 2 Directive
NO	Network Operator
NP	Nondeterministic Polynomial
OFDMA	Orthogonal Frequency-Division Multiple Access
OSS	Operation Support System
PaaS	Platform-as-a-Service
P-OFDM	Polar-Orthogonal Frequency Division Multiplexing
PKI	Public Key Infrastructure

PNF	Physical Network Functions
PSTN	Public Switched Telephone Network
QP	Quadratic Programming problem
QoE	Quality of Experience
QoS	Quality of Service
RAM	Random Access Memory
RAN	Radio Access Network
RODO	Rozporządzenie 2016/679 Parlamentu Europejskiego o ochronie danych osobowych
RTS	Ready To Send
RTT	Round Trip Time
SaaS	Software-as-a-Service
S.M.A.R.T.	Specific, Measurable, Achievable, Relevant, Time-bound
SBI	Southbound Interface
SDN	Software Defined Network
SDR	Software Defined Radio
SGX	Software Guard Extensions
SLA	Service Level Agreement
SLR	Skierowana Liczba Rozmyta
SMS	Short Message Service
SNR	Signal to Noise Ratio
SQL	Structured Query Language
SSD	Solid State Drive
SSH	Secure Shell
SSTP	Secure Socket Tunneling Protocol
TCP	Transmission Control Protocol
TDM	Time-Division Multiplexing
TDMA	Time-Division Multiple Access
TLS	Transport Layer Security
TOS	Type of Service
TPM	Trusted Platform Module

UDP	User Datagram Protocol
UE	User Equipment
URLLC	Ultra-Reliable and Low-Latency Communication
V2X	Vehicle-to-Everything
VIM	Virtualized Infrastructure Manager
VLAN	Virtual Local Area Network
VM	Virtual Machine
VNF	Virtual Network Functions
VNFM	Virtual Network Function Manager
VoIP	Voice over IP
VPN	Virtual Private Network
WDM	Wavelength Division Multiplexing
Wi-Fi	Wireless Fidelity
WWW	World Wide Web

Spis rysunków

1.1	Wymagania stawiane sieci 5G (IMT-2020 w nomenklaturze ETSI - European Telecommunications Standards Institute i ITU - <i>International Telecommunication Union</i>), w odniesieniu do wymagań wobec sieci 4G (<i>IMT-advanced</i> według ETSI i ITU); rysunek pochodzi z [100], zob. też [24], [36].	12
1.2	Ważność kluczowych zdolności sieci w kontekście trzech podstawowych kategorii komunikacji. rysunek pochodzi z [100], zob. też [24], [36].	13
1.3	Model NFV-MANO (<i>Network Function Virtualization-Management and Orchestration</i>) zdefiniowany w [26], patrz też [25].	15
1.4	Model NFV-MANO w uzupełnieniu 5G NORMA, zdefiniowany w [3].	16
2.1	Koncepcja sieci plasterkowej według NGMN - rysunek pochodzi z pracy [91], stworzony na podstawie [14].	23
2.2	Koncepcja łańcucha plasterków - rysunek pochodzi z pracy [91].	26
3.1	Przedstawienie kompleksowości potrzeb biznesowych w zakresie separacji strumieni danych i jezior danych na platformie ServiceNow. Szerokość poziomu piramidy symbolizuje częstotliwość z jaką należy wybierać poszczególne rozwiązania. Poziomy wyższe potrafią rozwiązywać bardziej skomplikowane problemy niż poziomy niższe. Rysunek utworzony na podstawie [109].	48
4.1	Wielowarstwowy model sieci - perspektywa sieciowa. Rysunek pochodzi z pracy [57].	56
4.2	Wielowarstwowy model sieci - perspektywa grafowa. Rysunek pochodzi z pracy [57].	57

5.1	Przykład dodawania dwóch różnych SLR o identycznych, dodatnich skierowaniach.	71
5.2	Przykład dodawania dwóch różnych SLR o przeciwnych skierowaniach.	72
6.1	Przykład złączenia dwóch funkcji liniowych, który w rezultacie tworzy funkcję nieodwracalną.	78
6.2	Przykład aproksymacji liniowej.	88
7.1	Przykładowa sieć do obliczeń, schemat zgodny z przykładem podanym w pracy [57].	107
7.2	Przykładowa sieć do obliczeń - zestaw węzłów podstawowych w sieciach RAN i CN, schemat zgodny z przykładem podanym w pracy [57].	108
7.3	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T1.	113
7.4	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T2.	114
7.5	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T3.	115
7.6	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T4.	116
7.7	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T5.	117
7.8	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T6.	118
7.9	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T7.	119
7.10	Normalizacja histogramu, wyznaczanie otoczki na podstawie równań (6.6 - 6.9) i wyznaczona SLR dla cechy T8.	120
7.11	Wyznaczona SLR (po operacji złączenia) dla cechy T1.	121
7.12	Wyznaczona SLR (po operacji złączenia) dla cechy T2.	121
7.13	Wyznaczona SLR (po operacji złączenia) dla cechy T3.	122
7.14	Wyznaczona SLR (po operacji złączenia) dla cechy T3 T6 T8.	122

7.15	Wyznaczona SLR (po operacji złączenia) dla cechy $T4 T7$	123
7.16	Wyznaczona SLR (po operacji złączenia) dla cechy $T5$	123
7.17	Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy $T1$	124
7.18	Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy $T2 T4 T7$	124
7.19	Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy $T3 T3 T6 T8$	125
7.20	Wektor izolacji - wyznaczona SLR (po operacji złączenia) dla cechy $T5$	125
7.21	Wyznaczona SLR opisująca końcową izolację w sieci.	126
7.22	Wyznaczona SLR opisująca końcową izolację w sieci - wariant właściwy wyznaczony za pomocą programowania liniowego.	126
9.1	Graf przedstawiający instancje usług, instancje plastrów i parametry z nimi powiązane. Rysunek pochodzi z pracy [88].	144

Spis tablic

2.1	Porównanie koncepcji plastrów horyzontalnych i plastrów wertykalnych	24
3.1	Przykładowe techniki izolacji w sieciach RAN, tabela zbudowana na podstawie [56]	51
4.1	Przykłady funkcji normalizujących, podanych w pracach [56, 57]	59
7.1	Zbiór funkcji normalizujących i znormalizowanych wartości, zgodnie z przykładem podanym w pracy [57]	109
7.2	Zbiór przykładowych danych pomiarowych	110
7.3	Zbiór znormalizowanych danych pomiarowych	111
7.4	Izolacja wewnątrz sieci RAN i CN	112
7.5	Wektor izolacji	112
8.1	Funkcje statystyczne i odpowiadające im zbiory	137
8.2	Wartości progowe $\phi_{i,j}$	138
8.3	Wartości estymatorów i zbiory $P_{i,j,k}$	141
8.4	Izolacja wewnątrz sieci RAN i CN - model probabilistyczny	142
8.5	Wektor izolacji - model probabilistyczny	142
10.1	Porównanie modeli izolacji	148

Dodatek A. Uzupełnienie do modelu probabilistycznego

W niniejszym dodatku zebrano definicje, lematy i twierdzenia matematyczne związane z modelem probabilistycznym, uzupełniające model probabilistyczny.

Definicja A.1 (Symetryczny rozkład prawdopodobieństwa ciągłej zmiennej losowej)

Rozkład ciągłej zmiennej losowej \mathbf{X} nazywamy symetrycznym rozkładem prawdopodobieństwa (w skrócie rozkładem symetrycznym) wtedy i tylko wtedy, gdy jego gęstość prawdopodobieństwa spełnia zależność:

$$f_{\mathbf{X}}(t) = f_{\mathbf{X}}(-t). \quad (\text{A.1})$$

Zmienną losową o takim rozkładzie oznaczamy jako $\mathbf{X} \sim \text{Sym}$.

Lemat A.1 (o dystrybuancie ciągłego rozkładu symetrycznego)

Niech $\mathbf{X} \sim \text{Sym}$ będzie ciągłą zmienną losową. Wówczas korzystając z równości (A.1) mamy:

$$\begin{aligned} F_{\mathbf{X}}(t) &= \int_{-\infty}^t f_{\mathbf{X}}(\tau) d\tau = \int_{-\infty}^t f_{\mathbf{X}}(-\tau) d\tau = \int_{\infty}^{-t} f_{\mathbf{X}}(\phi) (-d\phi) = \\ &= \int_{-t}^{\infty} f_{\mathbf{X}}(\phi) d\phi = 1 - \int_{-\infty}^{-t} f_{\mathbf{X}}(\phi) d\phi = 1 - F_{\mathbf{X}}(-t). \end{aligned} \quad (\text{A.2})$$

Lemat A.2 (O przekształceniu liniowym funkcji minimum i maksimum)

Niech \mathbf{X} będzie rozkładem opisującym zmienne losowe $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$. Oznaczmy za:

$$\left\{ \begin{array}{l} \mathbf{Y} = \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n), \\ \mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n), \\ \mathbf{V}_i = a\mathbf{X}_i + b; a, b \in \mathbb{R}; a \neq 0, \\ \mathbf{G} = \min(\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_n), \\ \mathbf{H} = \max(\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_n). \end{array} \right. \quad (\text{A.3})$$

Wówczas mamy:

$$\mathbf{G} = \begin{cases} a\mathbf{Y} + b \iff a \geq 0, \\ a\mathbf{Z} + b \iff a < 0, \end{cases} \quad (\text{A.4})$$

$$\mathbf{H} = \begin{cases} a\mathbf{Z} + b \iff a \geq 0, \\ a\mathbf{Y} + b \iff a < 0. \end{cases} \quad (\text{A.5})$$

Dowód. Gdy $a \geq 0$, to mamy:

$$\begin{aligned} \mathbf{G} &= \min(a\mathbf{X}_1 + b, a\mathbf{X}_2 + b, \dots, a\mathbf{X}_n + b) = b + \min(a\mathbf{X}_1, a\mathbf{X}_2, \dots, a\mathbf{X}_n) = \\ &= b + a \cdot \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) = a\mathbf{Y} + b, \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned} \mathbf{H} &= \max(a\mathbf{X}_1 + b, a\mathbf{X}_2 + b, \dots, a\mathbf{X}_n + b) = b + \max(a\mathbf{X}_1, a\mathbf{X}_2, \dots, a\mathbf{X}_n) = \\ &= b + a \cdot \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) = a\mathbf{Z} + b. \end{aligned} \quad (\text{A.7})$$

Dla $a < 0$ mamy z kolei:

$$\begin{aligned} \mathbf{G} &= \min(a\mathbf{X}_1 + b, a\mathbf{X}_2 + b, \dots, a\mathbf{X}_n + b) = b + \min(a\mathbf{X}_1, a\mathbf{X}_2, \dots, a\mathbf{X}_n) = \\ &= b + a \cdot \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) = a\mathbf{Z} + b, \end{aligned} \quad (\text{A.8})$$

$$\begin{aligned} \mathbf{H} &= \max(a\mathbf{X}_1 + b, a\mathbf{X}_2 + b, \dots, a\mathbf{X}_n + b) = b + \max(a\mathbf{X}_1, a\mathbf{X}_2, \dots, a\mathbf{X}_n) = \\ &= b + a \cdot \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) = a\mathbf{Y} + b. \end{aligned} \quad (\text{A.9})$$

■

Definicja A.2 (Symetryczny rozkład prawdopodobieństwa dyskretnej zmiennej losowej)

Rozkład dyskretnej zmiennej losowej \mathbf{X} nazywamy symetrycznym rozkładem prawdopodobieństwa (w skrócie rozkładem symetrycznym) wtedy i tylko wtedy, gdy jego funkcja masy prawdopodobieństwa [115] spełnia zależność:

$$pmf_{\mathbf{X}}(t) = pmf_{\mathbf{X}}(-t). \quad (\text{A.10})$$

Zmienną losową o takim rozkładzie oznaczamy jako $\mathbf{X} \sim \text{Sym}$.

Twierdzenie A.1 (Właściwości funkcji parzystych i nieparzystych)

Niech $f(x) \in \mathcal{N}$ oznacza, że $f(x)$ jest funkcją nieparzystą, czyli spełnia równość $f(x) = -f(-x)$ dla dowolnego x z jej dziedziny. Oznaczmy za $g(x) \in \mathcal{P}$ funkcję parzystą, czyli spełniającą równość $g(x) = g(-x) \in \mathcal{P}$ dla dowolnego x z jej dziedziny. Wówczas zachodzą właściwości:

$$h(x) = f(x)g(x) \Rightarrow h(x) = -f(-x)g(-x) = -h(-x) \in \mathcal{N}, \quad (\text{A.11})$$

$$w(x) \in \mathcal{N}, h(x) = f(x)w(x) \Rightarrow h(x) = -f(-x)(-w(-x)) = h(-x) \in \mathcal{P}, \quad (\text{A.12})$$

$$w(x) \in \mathcal{P}, h(x) = g(x)w(x) \Rightarrow h(x)g(-x)(w(-x)) = h(-x) \in \mathcal{P}. \quad (\text{A.13})$$

Mamy również następującą właściwość funkcji nieparzystej. Niech $f(x) \in \mathcal{N}$ i jest całkowna w sensie Riemanna na przedziale $[-T, T]$. Wówczas:

$$\begin{aligned} \int_{-T}^T f(x)dx &= \int_{-T}^0 f(x)dx + \int_0^T f(x)dx = \\ &= \int_T^0 f(-y)(-dy) + \int_0^T f(x)dx = \int_T^0 f(y)dy + \int_0^T f(x)dx = 0. \end{aligned} \quad (\text{A.14})$$

Lemat A.3 (o momentach funkcji minimum dla symetrycznych ciągłych zmiennych losowych)
Niech $\mathbf{Y} = \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ i $\mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ będą zmiennymi losowymi zależnymi od ciągu niezależnych zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$, które mają taki sam ciągły rozkład prawdopodobieństwa $\mathbf{X} \sim \text{Sym}$. Wówczas dla $m \in \mathbb{N}$ zachodzi równość:

$$\mathbb{E}(\mathbf{Y}^m) = (-1)^m \mathbb{E}(\mathbf{Z}^m). \quad (\text{A.15})$$

Dowód. Rozpatrzmy wyrażenie:

$$\mathbb{E}(\mathbf{Y}^m) - (-1)^m \mathbb{E}(\mathbf{Z}^m) = N \int_{-\infty}^{\infty} t^m ((1 - F_{\mathbf{X}}(t))^{N-1} - (-1)^m F_{\mathbf{X}}(t)^{N-1}) f_{\mathbf{X}}(t) dt. \quad (\text{A.16})$$

Korzystając z wzoru (A.2) mamy $1 - F_{\mathbf{X}}(t) = F_{\mathbf{X}}(-t)$, a więc:

$$\mathbb{E}(\mathbf{Y}^m) - (-1)^m \mathbb{E}(\mathbf{Z}^m) = N \int_{-\infty}^{\infty} t^m ((F_{\mathbf{X}}(-t))^{N-1} - (-1)^m F_{\mathbf{X}}(t)^{N-1}) f_{\mathbf{X}}(t) dt. \quad (\text{A.17})$$

Rozpatrzmy teraz dwa przypadki parzystości m . Gdy m jest liczbą parzystą, to funkcja $w(t) = t^m f_{\mathbf{X}}(t)$ jest funkcją parzystą, a funkcja $v(t) = (F_{\mathbf{X}}(-t))^{N-1} - (-1)^m F_{\mathbf{X}}(t)^{N-1}$ jest funkcją nieparzystą. Na podstawie właściwości (A.14) funkcja podcałkowa jest funkcją nieparzystą, a więc na mocy wzoru (A.11) całka jest równa zero, co dowodzi tezy.

Gdy m jest liczbą nieparzystą, to funkcja $v(t) = (F_{\mathbf{X}}(-t))^{N-1} - (-1)^m F_{\mathbf{X}}(t)^{N-1}$ jest funkcją parzystą, natomiast funkcja $w(t) = t^m f_{\mathbf{X}}(t)$ jest funkcją nieparzystą. Stosując analogiczne rozumowanie jak w przypadku parzystego m , wnioskujemy, że rozpatrywana całka jest równa zero. ■

Lemat A.4 (o momentach funkcji minimum dla symetrycznych dyskretnych zmiennych losowych)

Niech $\mathbf{Y} = \min(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ i $\mathbf{Z} = \max(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ będą zmiennymi losowymi zależnymi od ciągu niezależnych zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$, które mają taki sam dyskretny rozkład prawdopodobieństwa $\mathbf{X} \sim \text{Sym}$. Wówczas dla $m \in \mathbb{N}$ zachodzi równość:

$$\mathbb{E}(\mathbf{Y}^m) = (-1)^m \mathbb{E}(\mathbf{Z}^m). \quad (\text{A.18})$$

Dowód. Załóżmy, że rozkład prawdopodobieństwa \mathbf{X} w punktach $c_1 < c_2 < \dots < c_K$ ma wartości odpowiednio p_1, p_2, \dots, p_K , $\sum_{i=1}^K p_i = 1$. Z faktu, że $\mathbf{X} \sim \text{Sym}$ mamy dla $i, j = 1, 2, \dots, K$:

$$p_i = p_{K-i+1}, \quad (\text{A.19})$$

$$\sum_{i=1}^j p_i = F_{\mathbf{X}}(c_j) = \sum_{i=1}^j p_{K-i+1} = 1 - F_{\mathbf{X}}(c_{K-j}), \quad (\text{A.20})$$

$$c_i = -c_{K-i+1}. \quad (\text{A.21})$$

Na podstawie niezależności zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ i (8.9) mamy

$$\mathbb{P}(\mathbf{Y} = c_i) = F_{\mathbf{Y}}(c_i) - F_{\mathbf{Y}}(c_{i-1}) = (1 - F_{\mathbf{X}}(c_{i-1}))^N - (1 - F_{\mathbf{X}}(c_i))^N. \quad (\text{A.22})$$

Przyjmujemy w powyższym wzorze, że $F_{\mathbf{X}}(c_0) = 0$. Z własności (A.20) mamy:

$$\mathbb{P}(\mathbf{Y} = c_i) = (1 - F_{\mathbf{X}}(c_{i-1}))^N - (1 - F_{\mathbf{X}}(c_i))^N = (F_{\mathbf{X}}(c_{K-i+1}))^N - (F_{\mathbf{X}}(c_{K-i}))^N. \quad (\text{A.23})$$

Na podstawie wzoru 8.11 i niezależności zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$ mamy:

$$\mathbb{P}(\mathbf{Z} = c_i) = (F_{\mathbf{X}}(c_i))^N - (F_{\mathbf{X}}(c_{i-1}))^N. \quad (\text{A.24})$$

Rozpatrzmy dwa warianty parzystości m . Niech $m = 2m' + 1, m' \in \mathbb{N}_{\geq 0}$. W takim przypadku zmienne losowe \mathbf{Y}^m i \mathbf{Z}^m przyjmują tylko wartości $c_1^m, c_2^m, \dots, c_K^m$. Dla nieparzystych m mamy:

$$\mathbb{P}(\mathbf{Y}^m = c_i^m) = \mathbb{P}(\mathbf{Y} = c_i), \quad (\text{A.25})$$

$$\mathbb{P}(\mathbf{Z}^m = c_i^m) = \mathbb{P}(\mathbf{Z} = c_i), \quad (\text{A.26})$$

ponieważ podnoszenie do m -tej potęgi jest bijekcją dla nieparzystych wartości m . Wówczas mamy:

$$\mathbb{E}(\mathbf{Y}^m) = \sum_{i=1}^K c_i^m \mathbb{P}(\mathbf{Y}^m = c_i^m) = \sum_{i=1}^K c_i^m \mathbb{P}(\mathbf{Y} = c_i), \quad (\text{A.27})$$

$$\mathbb{E}(\mathbf{Z}^m) = \sum_{i=1}^K c_i^m \mathbb{P}(\mathbf{Z}^m = c_i^m) = \sum_{i=1}^K c_i^m \mathbb{P}(\mathbf{Z} = c_i). \quad (\text{A.28})$$

Uwzględniając właściwość (A.21) daje nam to:

$$\begin{aligned} \mathbb{E}(\mathbf{Y}^m) &= \sum_{i=1}^K c_i^m ((F_{\mathbf{X}}(c_{K-i+1}))^N - (F_{\mathbf{X}}(c_{K-i}))^N) = \\ &= \sum_{i=1}^K (-c_{K-i+1})^m ((F_{\mathbf{X}}(c_{K-i+1}))^N - (F_{\mathbf{X}}(c_{K-i}))^N) = \\ &= (-1)^m \sum_{i=1}^K c_{K-i+1}^m ((F_{\mathbf{X}}(c_{K-i+1}))^N - (F_{\mathbf{X}}(c_{K-i}))^N). \end{aligned} \quad (\text{A.29})$$

Położmy teraz $w = K - i + 1$. Dostajemy wtedy:

$$\mathbb{E}(\mathbf{Y}^m) = (-1)^m \sum_{w=1}^K c_w^m ((F_{\mathbf{X}}(c_w))^N - (F_{\mathbf{X}}(c_{w-1}))^N). \quad (\text{A.30})$$

Wobec tego, ze wzorów (A.24, A.28, A.30) mamy:

$$\begin{aligned}\mathbb{E}(\mathbf{Y}^m) &= (-1)^m \sum_{w=1}^K c_w^m ((F_{\mathbf{X}}(c_w))^N - (F_{\mathbf{X}}(c_{w-1}))^N) = \\ &= (-1)^m \sum_{w=1}^K c_w^m \mathbb{P}(\mathbf{Z} = c_w) = (-1)^m \mathbb{E}(\mathbf{Z}^m).\end{aligned}\tag{A.31}$$

Rozpatrzmy teraz przypadek, gdy $m = 2m'$, $m' \in \mathbb{N}_{\geq 0}$. Wtedy mamy:

$$\mathbb{P}(\mathbf{Y}^m = c_i^m) = \mathbb{P}(\mathbf{Y} = c_i \vee \mathbf{Y} = -c_i),\tag{A.32}$$

$$\mathbb{P}(\mathbf{Z}^m = c_i^m) = \mathbb{P}(\mathbf{Z} = c_i \vee \mathbf{Z} = -c_i).\tag{A.33}$$

Założmy, że $c_i \neq 0$, wówczas mamy na podstawie (A.21):

$$\mathbb{P}(\mathbf{Y}^m = c_i^m) = \mathbb{P}(\mathbf{Y} = c_i) + \mathbb{P}(\mathbf{Y} = -c_i) = \mathbb{P}(\mathbf{Y} = c_i) + \mathbb{P}(\mathbf{Y} = c_{K-i+1}),\tag{A.34}$$

$$\mathbb{P}(\mathbf{Z}^m = c_i^m) = \mathbb{P}(\mathbf{Z} = c_i) + \mathbb{P}(\mathbf{Z} = -c_i) = \mathbb{P}(\mathbf{Z} = c_i) + \mathbb{P}(\mathbf{Z} = c_{K-i+1}).\tag{A.35}$$

Na podstawie niezależności zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$, właściwości (8.9) oraz równania (A.22) mamy:

$$\begin{aligned}\mathbb{P}(\mathbf{Y}^m = c_i^m) &= \\ &= (1 - F_{\mathbf{X}}(c_{i-1}))^N - (1 - F_{\mathbf{X}}(c_i))^N + (1 - F_{\mathbf{X}}(c_{K-i}))^N - (1 - F_{\mathbf{X}}(c_{K-i+1}))^N = \\ &= (1 - F_{\mathbf{X}}(c_{i-1}))^N - (1 - F_{\mathbf{X}}(c_i))^N + (F_{\mathbf{X}}(c_i))^N - (F_{\mathbf{X}}(c_{i-1}))^N.\end{aligned}\tag{A.36}$$

Na podstawie niezależności zmiennych losowych $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$, właściwości (8.11) oraz równania (A.24) mamy:

$$\begin{aligned}\mathbb{P}(\mathbf{Z}^m = c_i^m) &= \\ &= (F_{\mathbf{X}}(c_i))^N - (F_{\mathbf{X}}(c_{i-1}))^N + (F_{\mathbf{X}}(c_{K-i+1}))^N - (F_{\mathbf{X}}(c_{K-i}))^N = \\ &= (F_{\mathbf{X}}(c_i))^N - (F_{\mathbf{X}}(c_{i-1}))^N + (1 - F_{\mathbf{X}}(c_{i-1}))^N - (1 - F_{\mathbf{X}}(c_i))^N = \mathbb{P}(\mathbf{Y}^m = c_i^m).\end{aligned}\tag{A.37}$$

Dalej mamy:

$$\mathbb{E}(\mathbf{Y}^m) = \sum_{i=1}^K c_i^m \mathbb{P}(\mathbf{Y}^m = c_i^m) = \sum_{i=1}^K c_i^m \mathbb{P}(\mathbf{Z}^m = c_i^m) = \mathbb{E}(\mathbf{Z}^m). \quad (\text{A.38})$$

Dla parzystego m daje to dowodzoną tezę. Pozostał nam przypadek, gdy dla pewnego i mamy $c_i = 0$. Dla tego przypadku mamy:

$$\mathbb{P}(\mathbf{Y}^m = 0) = \mathbb{P}(\mathbf{Y} = 0) = (1 - F_{\mathbf{X}}(c_{i-1}))^N - (1 - F_{\mathbf{X}}(0))^N, \quad (\text{A.39})$$

$$\mathbb{P}(\mathbf{Z}^m = 0) = \mathbb{P}(\mathbf{Z} = 0) = (F_{\mathbf{X}}(0))^N - (F_{\mathbf{X}}(c_{i-1}))^N. \quad (\text{A.40})$$

Ponieważ ciąg c_1, c_2, \dots, c_K jest ściśle rosnący, to $c_i = 0 = -c_{K-i+1}$ a więc $i = \frac{K+1}{2}$. Zachodzi więc równość:

$$F_{\mathbf{X}}(c_i) = 1 - F_{\mathbf{X}}(c_{K-i}) = 1 - F_{\mathbf{X}}\left(c_{\frac{K-1}{2}}\right) = 1 - F_{\mathbf{X}}(c_{i-1}). \quad (\text{A.41})$$

Z powyższego wynika, że:

$$\mathbb{P}(\mathbf{Y}^m = 0) = \mathbb{P}(\mathbf{Z}^m = 0). \quad (\text{A.42})$$

Pozwala to zastosować analogiczne rozumowanie we wzorze (A.38) jak dla pozostałych punktów $c_1, c_2, \dots, c_{i-1}, c_{i+1}, c_{i+2}, \dots, c_K$. ■