

Rozprawa doktorska

Modelowanie złożonych systemów niepełnodostępnych

mgr inż. Michał Weissenberg

Promotor:
dr hab. inż. Sławomir Hanczewski

Wydział Informatyki i Telekomunikacji
Politechnika Poznańska
03.06.2024

Streszczenie

Rozprawa doktorska podejmuje tematykę modelowania złożonych wielousługowych (heterogenicznych) systemów teleinformatycznych. Osiągnięcie naukowe, stanowiące podstawę rozprawy, obejmuje cykl powiązanych tematycznie: pięciu artykułów naukowych, opublikowanych w czasopismach naukowych, oraz trzech referatów, opublikowanych w recenzowanych materiałach z konferencji międzynarodowych. W pracy zaprezentowano modele analityczne umożliwiające wyznaczenie charakterystyk jakościowych złożonych systemów teleinformatycznych, w szczególności systemów chmurowych działających w modelu infrastruktura jako usługa (ang. *Infrastructure as a Service - IaaS*).

W pierwszej części pracy zawarto omówienie osiągnięć, o których mowa w art. 186 ust. 1 pkt. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce. Przedstawiony został wstęp do problematyki badawczej oraz cel i zakres badań. Następnie doktorant opisał podstawowe modele analityczne systemów wielousługowych, które posłużyły do opracowania własnych modeli systemów chmurowych IaaS. W dalszej części pracy zaprezentowano opracowane modele jednoparametrycznych oraz wieloparametrycznych złożonych systemów teleinformatycznych z ruchem strumieniowym. Doktorant opisał także kolejkowy model systemu teleinformatycznego z ruchem Erlanga, Engseta oraz Pascala typu elastycznego i adaptacyjnego. Wszystkie z zaproponowanych modeli mają charakter przybliżony, w związku z tym w pracy został opisany modułowy model symulacyjny opracowany przez doktoranta, który posłużył do weryfikacji dokładności opracowanych modeli analitycznych.

W dalszej części pracy zaprezentowany został wykaz pozostałych osiągnięć doktoranta, niewchodzących w skład dzieła, a składających się na całokształt jego dorobku. W ostatniej części rozprawy dołączone zostały załączniki w postaci artykułów stanowiących dorobek doktoranta oraz oświadczenia współautorów artykułów.

Opracowane w procesie realizacji badań i wchodzące w skład dorobku modele umożliwiły określenie parametrów jakościowych dla omawianych systemów, takich jak prawdopodobieństwo strat, prawdopodobieństwo blokady, a także wykorzystanie poszczególnych zasobów systemu. Dzięki temu, metody te mogą posłużyć w procesie wymiarowania systemów chmurowych do oszacowania wielkości wykorzystywanych w nich zasobów (serwerów), przy zachowaniu wymaganej jakości usług lub poziomu eksploatacji urządzeń. Wyniki przeprowadzonych eksperymentów dla założonych scenariuszy badawczych, potwierdziły dokładność zaproponowanych metod.

Abstract

The doctoral dissertation concerns the modeling of complex multi-service (heterogeneous) ICT systems. The presented scientific achievement includes: a series of five thematically related scientific articles published in academic journals and three articles published in peer-reviewed proceedings of international conferences. The dissertation presents analytical models used to determine the qualitative characteristics of complex ICT systems, particularly cloud systems operating in the Infrastructure as a Service (IaaS) model.

The first part of the article includes a discussion of the achievements mentioned in Article 186, paragraph 1, item 1 of the Act of July 20, 2018, Law on Higher Education and Science. It presents an introduction to the research problem as well as the aim and scope of the research. Subsequently, the doctoral candidate described the basic analytical models of multi-service systems, which were used to develop their own models of IaaS cloud systems. The following section of the work presents the developed singleparameter and multiparameter models of complex ICT systems with stream traffic. The doctoral candidate also described the model of queuing ICT system with Erlang, Engset, Pascal (BPP) elastic and adaptive traffic. All the proposed models are approximate; therefore, the dissertation describes the modular simulation model developed by the doctoral candidate, which was used to verify the accuracy of the developed analytical models.

The subsequent part of the dissertation presents a list of other achievements by the doctoral candidate that are not comprising the main scientific achievement but contribute to their overall body of work. The final part of the dissertation includes appendices in the form of articles constituting the doctoral candidate's main scientific achievement and statements from the co-authors of the articles.

The models developed during the research and included in the dissertation allowed for the determination of qualitative parameters for the considered systems, such as loss probability, blocking probability, and the utilization of individual system resources. As a result, these methods can be used in the dimensioning process of cloud systems to estimate the size of the resources (servers) used in them while maintaining the required quality of services or level of device operation. The results of the experiments conducted for the assumed test scenarios confirmed the validity of the proposed methods.

Spis treści

1	Omówienie osiągnięć, o których mowa w art. 186 ust. 1 pkt. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce	6
1.1	Artykuły w czasopiśmie naukowych	6
1.2	Artykuły w materiałach konferencyjnych (recenzowane i opublikowane)	7
1.3	Wstęp do problematyki badawczej	7
1.4	Cel i zakres badań	10
1.5	Podstawowe modele analityczne systemów wielousługowych	11
1.5.1	Modele systemów niezależnych od stanu	13
1.5.2	Modele systemów zależnych od stanu	14
1.6	Opis głównych osiągnięć doktoranta	20
1.6.1	Złożona struktura systemu	20
1.6.2	Modele jednoparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego.	21
1.6.3	Modele wieloparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego	35
1.6.4	Model jednoparametrycznego kolejkowego systemu teleinformatycznego z ruchem Erlanga, Engseta i Pascala (BPP) typu elastycznego i adaptacyjnego.	47
1.6.5	Modułowy symulator umożliwiający analizę złożonych systemów teleinformatycznych z ruchem Erlanga, Engseta, Pacala typu strumieniowego, elastycznego oraz adaptacyjnego.	56
2	Wykaz pozostałych osiągnięć doktoranta	70
2.1	Spis pozostałych publikacji doktoranta	70
2.1.1	Publikacje w obszarze modelowania systemów teleinformatycznych	70
2.1.2	Pozostałe publikacje	71
2.2	Wykaz udziału w komitetach organizacyjnych i naukowych konferencji krajowych lub międzynarodowych, z podaniem pełnionej funkcji.	72
2.2.1	Członek komitetu organizacyjnego:	72
2.3	Wykaz uczestnictwa w pracach zespołów badawczych realizujących projekty, z podziałem na projekty zrealizowane i będące w toku realizacji oraz z uwzględnieniem informacji o pełnionej funkcji w ramach prac zespołów.	72
2.3.1	Projekty międzynarodowe	72
2.3.2	Projekty krajowe	74

2.4	Wykaz staży w instytucjach naukowych lub artystycznych, w tym zagranicznych, z podaniem miejsca, terminu, czasu trwania stażu i jego charakteru	75
2.5	Współpraca z otoczeniem społeczno-gospodarczym	75
2.5.1	Współpraca z sektorem gospodarczym.	75
2.5.2	Współpraca z otoczeniem społecznym	76
2.6	Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę	77
2.6.1	Prowadzone zajęcia dydaktyczne	77
2.6.2	Wybrane kursy szkoleniowe i inne zajęcia dydaktyczne	78
2.6.3	Odbyte kursy i szkolenia	78
2.7	Załącznik 9 - Oświadczenia współautorów artykułów	79

Dane osobowe doktoranta

1. Imię i nazwisko: Michał Weissenberg.
2. Data urodzenia: 24 kwiecień 1992
3. Miejsce urodzenia: Świecie, Polska.
4. Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania.

2017: dyplom magistra inżyniera elektroniki i telekomunikacji, w zakresie sieci teleinformatycznych, Politechnika Poznańska, Wydział Elektroniki i Telekomunikacji.

5. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych.

2014-2019: referent techniczny w Katedrze Sieci Telekomunikacyjnych i Komputerowych, Wydział Elektroniki i Telekomunikacji, Politechnika Poznańska;

2020-2020: referent techniczny w Instytucie Sieci Teleinformatycznych, Wydział Informatyki i Telekomunikacji, Politechnika Poznańska;

2020-obecnie: asystent w Instytucie Sieci Teleinformatycznych, Wydział Informatyki i Telekomunikacji, Politechnika Poznańska.

Rozdział 1

Omówienie osiągnięć, o których mowa w art. 186 ust. 1 pkt. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce

Zgłoszone osiągnięcie naukowe pod tytułem „**Modelowanie złożonych systemów niepełnodostępnych**” obejmuje poniższy cykl powiązanych tematycznie artykułów naukowych opublikowanych w czasopismach naukowych lub w recenzowanych materiałach z konferencji międzynarodowych.

1.1 Artykuły w czasopismach naukowych

- [J1] Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg. High-accuracy analytical model for heterogeneous cloud systems with limited availability of physical machine resources based on Markov Chain. *Electronics*.13:11:2161. <https://doi.org/10.3390/electronics131121612024> [IF: 2.9, MEiN: 100]
- [J2] Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg. An analytical model of IaaS architecture for determining resource utilization. *Sensors*, 9:2758, 2024. <https://doi.org/10.3390/s24092758> [IF: 3.9, MEiN: 100]
- [J3] Joanna Weissenberg, Michał Weissenberg. Model of a queuing system with BPP elastic and adaptive traffic. *IEEE Access*. wolumen 10, strony 130771-130783, 2022. <https://doi.org/10.1109/ACCESS.2022.3229234> [IF: 3.476, MEiN: 100]
- [J4] Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg. A multiparameter analytical model of the physical infrastructure of a cloud-based system. *IEEE Access*. wolumen 9, strony 100981-100990, 2021. <https://doi.org/10.1109/ACCESS.2021.3097157> [IF: 3.476, MEiN: 100]

- [J5] Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg. Modeling of complex non-full-availability systems. *Journal of Telecommunications and Information Technology (JTIT)*. numer 2, strony 70–77, Styczeń 2020.
<https://doi.org/10.26636/jtit.2020.143320> [MEiN: 40]

1.2 Artykuły w materiałach konferencyjnych (recenzowane i opublikowane)

- [C1] Sławomir Hanczewski, Michał Weissenberg. A Study on the properties of complex non-full-availability systems. *39th International Business Information Management Association Conference IBIMA*. 30-31 May 2022, Granada, Spain. [MEiN: 70]
- [C2] Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg. The analytical model of complex non-full-availability system. *Choraś, M., Choraś, R. (eds) Image Processing and Communications. IP&C 2019. Advances in Intelligent Systems and Computing*. wolumen 1062, 2019. Springer, Cham.
<https://doi.org/10.1007/978-3-030-31254-1> [MEiN: 20]
- [C3] Sławomir Hanczewski, Michał Weissenberg. Concept of an analytical model for cloud computing infrastructure. *2018 11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP)*. strony 1-4, Budapeszt, Węgry, 2018.
<https://doi.org/10.1109/CSNDSP.2018.8471814> [MEiN: 20]

Pozostałe osiągnięcia naukowe doktoranta, składające się na całokształt dorobku, zostały wskazane w rozdziale 2.1 rozprawy doktorskiej.

1.3 Wstęp do problematyki badawczej

Współczesne systemy teleinformatyczne charakteryzują się, z punktu widzenia użytkownika końcowego, niemal nieograniczoną przepustowością oraz mnogością oferowanych usług. Przekłada się to na różnorodność strumieni ruchu, generowanych przez różne typy aplikacji, oraz konieczność ich wspólnej obsługi w zintegrowanych systemach. Przykładami takich strumieni mogą być m.in. strumienie wideo na żądanie, telefonii internetowej, wydarzeń transmitowanych w czasie rzeczywistym, poczta elektroniczna, a także żądania utworzenia maszyn wirtualnych w systemach chmurowych. Każda z tych usług może podlegać różnym metodom zarządzania ruchem, zależnym od oczekiwań użytkownika końcowego. Problem zapewnienia wymaganych parametrów jakościowych dla obsługiwanych strumieni jest zwielokrotniany poprzez gwałtowny wzrost wolumenu przesyłanych danych (zjawisko wzrostu natężenia ruchu w sieci Internet zostało szczegółowo opisane w raportach [1, 2]). Rzeczywistość ostatnich lat pokazała, że choć dostępna przepustowość sieci jest rzeczywiście bardzo duża, to wciąż ma pewne ograniczenia. Na przełomie lutego i marca 2020 roku, kiedy kraje europejskie zamroziły

swoje gospodarki w całkowitym lub częściowym lockdownie podczas pandemii Covid-19, telepraca i zdalna edukacja, a także wszelkiego rodzaju rozrywka online spowodowały, że zaobserwowany został znaczący, dochodzący nawet do 30%, wzrost natężenia ruchu w sieci [3, 4, 5]. Bezpośrednią konsekwencją powyższego było m.in. wystosowanie przez Komisję Europejską żądania do operatora Netflix oraz operatorów innych platform VoD (ang. *Video on Demand*) o ograniczenie jakości oferowanych materiałów wideo w celu zmniejszenia obciążenia sieci [6]. Ponadto obserwowany był ogromny wzrost użytkowników rozwiązań wideokonferencyjnych, jak np. ZOOM [4].

Kolejnym ogromnym wyzwaniem dla dostawców usług jest coraz bardziej złożona struktura oferowanych usług. Obecnie użytkownik oczekuje od systemu nie tylko przepływności czy też dostępu do łącza telefonicznego, ale także innych zasobów, w tym zasobów obliczeniowych, dostępu do maszyn wirtualnych lub aplikacji. Przykładem takich systemów są systemy chmurowe, które stanowią jeden z kluczowych obszarów na rynku ICT. Jak pokazuje raport Grand View Research z września 2023 roku [7], globalna wielkość samego rynku usług chmurowych do roku 2030 może osiągnąć wartość ponad 1 550 mld USD, co przekłada się na współczynnik CAGR (ang. *Compound Annual Growth Rate*) o wartości na poziomie 14,1%. Przyczynia się do tego zarówno przenoszenie aplikacji natywnych do chmury oraz zapotrzebowanie na przestrzeń dyskową, jak i migracja maszyn fizycznych z siedzib firm na rzecz maszyn wirtualnych oferowanych przez dostawców, na przykład w ramach usługi IaaS (ang. *Infrastructure as a Service*). IaaS jest rozwiązaniem, w którym dostawcy usług przydzielają użytkownikom końcowym moc obliczeniową w postaci maszyn wirtualnych (MW), uruchamianych na posiadanych przez dostawcę usługi maszynach fizycznych (MF). Rozwiązanie to zostało zdefiniowane przez NIST (ang. *National Institute of Standards and Technology*) jako zdolność do dostarczenia użytkownikowi końcowemu możliwości przetwarzania danych, dostępu do pamięci masowej, sieci i innych podstawowych zasobów obliczeniowych, w których użytkownik końcowy może wdrażać i uruchamiać dowolne oprogramowanie, które może obejmować systemy operacyjne i aplikacje. Użytkownik końcowy nie zarządza ani nie kontroluje podstawowej infrastruktury chmury, ale ma kontrolę nad systemami operacyjnymi, pamięcią masową i wdrożonymi aplikacjami, a także ewentualnie ograniczoną kontrolę nad wybranymi komponentami sieciowymi (np. zaporami hosta) [8]. Obecnie na rynku występuje wielu operatorów świadczących takie usługi, z których – do globalnych liderów – można zaliczyć Amazon z usługą AWS oraz Microsoft z usługą Azure. Udział tych podmiotów w rynku usług chmurowych, w czwartym kwartale roku 2023, wynosił ponad 50% [9]. W ramach rozwiązań wymienionych firm, użytkownicy mają możliwość zakupu maszyn wirtualnych opisanych trzema głównymi parametrami, takimi jak procesor, tj. CPU (ang. *Central Processing Unit*), pamięć RAM (ang. *Random Access Memory*) oraz przestrzeń dyskowa [10, 11], a także dostęp do usługi przy pomocy łącza internetowego.

Wskazane powyżej wymagania użytkownika końcowego, a także pojawiająca się w systemach coraz bardziej skomplikowana architektura oraz mnogość aplikacji, stawia przed operatorami szereg całkowicie nowych wyzwań. Dotyczą one między innymi oceny wydajności sieci oraz możliwości świadczenia usług na odpowiednim poziomie jakości QoS (ang. *Quality of Service*), który ma być zgodny z podpisywaną z użytkownikami umową o poziomie świadczenia usług SLA (ang. *Service Level Agreement*). Taka ocena może być przeprowadzona za pomocą modeli analitycznych lub symulacyjnych,

a także na podstawie długotrwałych pomiarów. Najszybszym sposobem sprawdzenia, jaki wpływ na działanie systemu będzie miało przyjęcie do obsługi nowego żądania, są modele analityczne. Modele te pozwalają uwzględnić bardzo wiele wykorzystywanych przez operatorów mechanizmów zarządzania ruchem, których stosowanie wynika przede wszystkim z próby ograniczenia ciągłej rozbudowy sieci na rzecz lepszego wykorzystania istniejących w niej zasobów. Do mechanizmów tych możemy zaliczyć m.in.:

- dynamiczną rezerwację zasobów [12, 13, 14, 15, 16], w której wprowadzane są pewne poziomy obciążenia systemu, po przekroczeniu których w systemie mogą być obsługiwane tylko zgłoszenia ściśle określonych klas zgłoszeń,
- przelew ruchu [17, 18, 19] polegający na przekierowaniu zgłoszenia, które nie może być obsłużone w jednym systemie, do innego systemu posiadającego wystarczającą liczbę zasobów do obsługi tego zgłoszenia,
- priorytety [20, 21], które dla poszczególnych klas zgłoszeń pozwalają określić ich ważność dla operatora wyrażoną w pierwszeństwie przyznawanych zasobów,
- kompresję ruchu [22, 23, 24, 25] umożliwiającą zmianę przepływności obsługiwanych zgłoszeń w zależności od obciążenia systemu,
- kolejkowanie zgłoszeń [26, 27, 28], które pozwala na oczekiwanie zgłoszeń na dostęp do systemu w przypadku chwilowego braku zasobów.

W literaturze przedmiotu, analiza telekomunikacyjnych systemów wielousługowych odbywa się w znaczącej większości na poziomie zgłoszeń. Każde zgłoszenie zdefiniowane jest jako strumień pakietów (ang. *flow*), o zadanych wymaganiach dotyczących żądanych zasobów. Wymagania te określane są zazwyczaj przez stałe przepływności (CBR – ang. *Constant Bit Rate*), które najczęściej wyznaczane są na podstawie prędkości bitowych rzeczywistych strumieni pakietów o zmiennej przepływności (VBR – ang. *Variable Bit Rate*), lub też na podstawie tak zwanego pasma ekwiwalentnego (EB – ang. *Equivalent Bandwidth*). Pasma ekwiwalentne to pewna wartość liczbowa, która jest wyznaczana na podstawie zarówno wymagań jakościowych źródła ruchu, jak i charakterystyki samego strumienia przez nie generowanego [29, 30]. Sam dobór pasma ekwiwalentnego nie przekłada się jednak na sposób modelowania systemu, a jedynie na jego parametryzację. W przypadku innych typów systemów, takich jak systemy chmurowe, analiza odbywa się na poziomie zgłoszeń żądania dostępu do określonych zasobów, na przykład procesora, przestrzeni dyskowej oraz pamięci RAM.

W pracach [23, 31] wykazano, że ruch na poziomie zgłoszeń może być aproksymowany strumieniem Poissona. Przekłada się to na możliwość analizowania systemu wielousługowego na podstawie wielowymiarowych procesów Markowa lub jednowymiarowych łańcuchów Markowa [22]. Modelowanie za pomocą wielowymiarowych procesów Markowa opiera się na tzw. mikrostanach, z których każdy jest definiowany przez liczbę zgłoszeń poszczególnych klas ruchu obsługiwanych w systemie. Modelowanie z wykorzystaniem jednowymiarowych łańcuchów Markowa opiera się z kolei na tzw. makrostanach, definiowanych przez całkowitą liczbę zajętych zasobów w rozważanym stanie zajętości systemu, niezależnie od rozkładu tej zajętości pomiędzy poszczególne klasy

ruchu [22]. W swoich badaniach doktorant podjął próbę opracowania modeli analitycznych złożonych systemów teleinformatycznych, wykorzystując zarówno wielowymiarowe procesy Markowa, jak i jednowymiarowe łańcuchy Markowa.

Złożoność modelowanych systemów może wynikać z czynników takich jak ich struktura, charakter oferowanych im zgłoszeń oraz zaimplementowanych metod zarządzania ruchem. Przykładem złożonego systemu teleinformatycznego są systemy chmurowe, w których to zgłoszenia pojawiające się na wejściu wymagają jednocześnie więcej niż jednego zasobu (procesor, przestrzeń dyskowa, pamięć RAM oraz przepływność). W celu uwzględnienia różnych wymaganych zasobów oraz dodatkowo struktury systemu chmurowego, doktorant zdecydował się w swoich badaniach rozważać rozwiązania implementowane zgodnie z modelem IaaS (ang. *Infrastructure as a Service*). Model ten jest często rozważany w literaturze przedmiotu, zarówno pod kątem wyznaczania poziomu strat w systemie [32, 33, 34, 35, 36, 37, 38], jak i innych parametrów jakościowych związanych z wykorzystaniem zasobów [39, 40, 41, 42, 43, 44, 45].

1.4 Cel i zakres badań

Celem przeprowadzonych badań było opracowanie analitycznych i symulacyjnych modeli wielousługowych (heterogenicznych) złożonych systemów teleinformatycznych. Dodatkowym celem było: (1) opracowanie metod umożliwiającą analizę wykorzystania poszczególnych zasobów w złożonym systemie teleinformatycznym, obsługującym strumienie żądające jednocześnie zasobów różnych typów, (2) analiza systemów teleinformatycznych z ruchem Erlanga, Engseta oraz Pascala typu elastycznego i adaptacyjnego, które w przyszłości posłużą do rozbudowy opracowanych modeli złożonych systemów teleinformatycznych z ruchem Erlanga typu strumieniowego.

Przeprowadzone badania umożliwiły sformułowanie głównej tezy badawczej w następującym brzmieniu:

Teza badawcza:

Możliwe jest opracowanie na poziomie zgłoszeń modeli analitycznych o wysokiej dokładności dla wielousługowych złożonych systemów teleinformatycznych.

W ramach opisywanego w punktach 1.1-1.2 cyklu prac, doktorant z powodzeniem zrealizował następujące zadania:

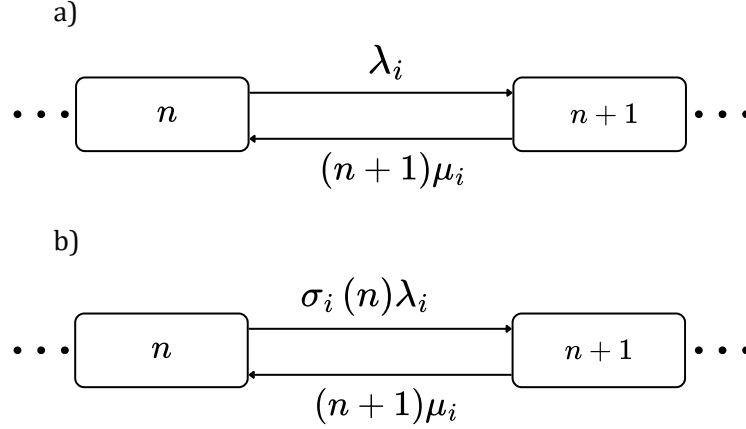
- Opracowanie koncepcji i modelu analitycznego wieloparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego. Wyniki badań zostały przedstawione w pracy [C3].
- Opracowanie i weryfikacja podstawowego modelu analitycznego jednoparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego. Wynik badań opublikowane zostały w pracy [C2].

- Opracowanie i weryfikacja rozszerzonego modelu analitycznego jednoparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego. Model charakteryzuje się większą dokładnością w porównaniu do modelu przedstawionego w pracy [C2]. Wyniki badań zaprezentowano w pracach [J5][C1].
- Opracowanie i weryfikacja zaawansowanego modelu analitycznego jednoparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego. Model charakteryzuje się większą dokładnością w porównaniu do modeli opublikowanych w pracach [C2,J5]. Wyniki badań zaprezentowano w pracy [J1].
- Opracowanie i weryfikacja modelu analitycznego wieloparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego. Wyniki badań zostały opublikowane w pracy [J4].
- Opracowanie i weryfikacja modelu analitycznego wieloparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego, umożliwiającego wyznaczenie stopnia wykorzystania poszczególnych zasobów systemu. Wyniki badań zaprezentowano w pracy [J2].
- Opracowanie i weryfikacja modelu analitycznego jednoparametrycznego kolejkowego systemu teleinformatycznego z ruchem Erlanga, Engseta i Pascala (BPP) typu elastycznego i adaptacyjnego. Wyniki badań opublikowano w pracy [J3].
- Opracowanie i implementacja modułowego symulatora, umożliwiającego analizę złożonych systemów teleinformatycznych z ruchem Erlanga, Engseta, Pascala typu strumieniowego, elastycznego oraz adaptacyjnego [J1,J2,J3,J4,J5,C1,C2,C3].

1.5 Podstawowe modele analityczne systemów wielo-usługowych

W tym rozdziale przedstawione zostały podstawowe modele analityczne systemów teleinformatycznych, które posłużyły doktorantowi do opracowania własnych modeli systemów chmurowych IaaS.

W ramach rozważań nad modelowaniem systemów teleinformatycznych można wyróżnić modele dla systemów niezależnych od stanu oraz zależnych od stanu. W systemach niezależnych od stanu przejście pomiędzy poszczególnymi stanami reprezentującymi zajętość systemu zależy jedynie od parametrów λ_i i μ_i , gdzie dla każdej z m klas zgłoszeń, λ_i określa intensywność strumienia napływu zgłoszeń klasy i ($i \leq i \leq m$), natomiast μ_i intensywność strumienia obsługi dla klasy i ($i \leq i \leq m$). Iloraz tych dwóch wartości umożliwia wyznaczenie średniej wartości natężenia ruchu a_i klasy i ($a_i = \frac{\lambda_i}{\mu_i}$). W systemach niezależnych od stanu, pojawiające się na wejściu systemu zgłoszenie ma dostęp do wszystkich jego zasobów, o ile są one wolne. Natomiast w przypadku systemów zależnych od stanu, zgłoszenia mają dostęp do ograniczonej części zasobów. W związku z tym, przejście pomiędzy stanami zależy dodatkowo od tak zwanego warunkowego współczynnika przejścia pomiędzy stanami $\sigma_i(n)$, gdzie i jest numerem klasy



Rysunek 1.1: Przejście między stanami procesu Markowa: (a) system z niezależnym od stanu procesem napływania zgłoszeń i obsługi zgłoszeń, (b) system z zależnym od stanu procesem napływania i/lub obsługi zgłoszeń [46].

w systemie ($i \leq i \leq m$), natomiast n liczbą zajętych jednostek alokacji JA w systemie. Zależność ta w uproszczony sposób została przedstawiona na rysunku 1.1. JA są bezwymiarowymi jednostkami alokacji zasobów (ang. *Allocation Unit*), które umożliwiają wyrażenie żądań poszczególnych strumieni ruchu na potrzeby procesu modelowania. Są one wyznaczone jako najmniejszy wspólny dzielnik żądań wszystkich klas ruchu.

Jak zostało udowodnione w pracy [46], w przypadku systemów, w których zależność od stanu jest wynikiem wielu różnych i niezależnych czynników, całkowity warunkowy współczynnik przejścia można wyznaczyć następująco:

$$\sigma_i(n) = [\sigma_i(n)]^R [\sigma_i(n)]^A [\sigma_i(n)]^P, \quad (1.1)$$

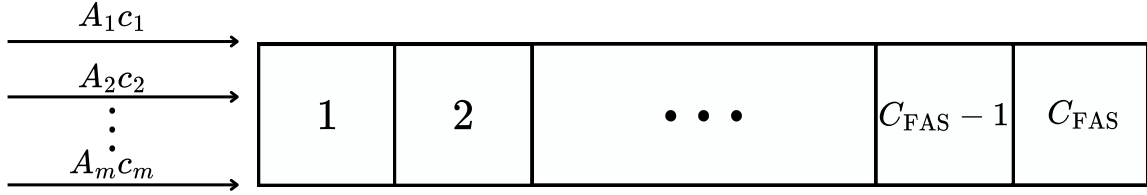
gdzie:

- $[\sigma_i(n)]^R$ jest warunkowym współczynnikiem przejścia dla zgłoszeń klasy i w stanie zajętości n , wynikającym z właściwości strumienia zgłoszeń tej klasy,
- $[\sigma_i(n)]^A$ jest warunkowym współczynnikiem przejścia dla zgłoszeń klasy i w stanie zajętości n , wynikającym ze struktury systemu,
- $[\sigma_i(n)]^P$ jest warunkowym współczynnikiem przejścia dla zgłoszeń klasy i w stanie zajętości n , wynikającym z przyjętego algorytmu przyjmowania nowych zgłoszeń w systemie.

Znajomość wartości wszystkich parametrów $[\sigma_i(n)]$ umożliwia określenie rozkładu zajętości w systemie za pomocą następującego wzoru [46]:

$$n[P(n)]_C = \sum_{i=1}^m A_i c_i [\sigma(n)_i] [P(n - c_i)]_C, \quad (1.2)$$

gdzie $[P(n)]_C$ jest prawdopodobieństwem zajętości n JA w systemie, natomiast C pojemnością systemu.



Rysunek 1.2: Ilustracja konstrukcji modelu FAS.

1.5.1 Modele systemów niezależnych od stanu

W ramach modeli systemów niezależnych od stanu można wyróżnić model systemu pełnodostępnego (ang. *Full Available System* – FAS). Model ten został zaproponowany niezależnie przez Kaufmana [22] i Roberta [47] do oceny prawdopodobieństwa blokady w łączu, w którym zgłoszenia o różnych przepływnościach są multipleksowane. W modelu tym zgłoszenia pojawiające się na wejściu systemu FAS posiadają dostęp do wszystkich wolnych zasobów tego systemu.

1.5.1.1 Model FAS

FAS to model zasobów o pełnej dostępności. Zakłada się w nim, że zasoby o pojemności C_{FAS} JA obsługują m klas zgłoszeń Erlanga, podczas gdy każde zgłoszenie klasy i ($i \in \{1, \dots, m\}$) wymaga c_i JA do obsługi. Nowe zgłoszenie klasy i zostanie obsłużone w systemie FAS wtedy i tylko wtedy, gdy model FAS ma c_i wolnych JA. Ilustracja takiego systemu została przedstawiona na rysunku 1.2.

Rozkład zajętości dla FAS jest opisany następującym wzorem [22, 47]:

$$n[P_{\text{FAS}}(n)]_{C_{\text{FAS}}} = \sum_{i=1}^m A_{\text{FAS},i} c_i [P(n - c_i)]_{C_{\text{FAS}}}, \quad (1.3)$$

gdzie $A_{\text{FAS},i}$ to natężenie ruchu klasy i , a $[P_{\text{FAS}}(n)]_{C_{\text{FAS}}}$ to prawdopodobieństwo zajętości n JA w FAS o pojemności C_{FAS} JA. Prawdopodobieństwo blokady zgłoszeń klasy i (definiowane jako prawdopodobieństwo zdarzenia w systemie, gdy aktualny stan zajętości zasobów uniemożliwia przyjęcie kolejnego zgłoszenia klasy) jest określane za pomocą następującego wzoru:

$$E_{\text{FAS},i} = \sum_{n=C_{\text{FAS}}-c_i+1}^{C_{\text{FAS}}} [P_{\text{FAS}}(n)]_{C_{\text{FAS}}}. \quad (1.4)$$

W przypadku systemów z ruchem Erlanga prawdopodobieństwo blokady E jest równe prawdopodobieństwu strat B , które definiowane jest jako stosunek liczby zgłoszeń traconych w systemie do wszystkich zgłoszeń oferowanych temu systemowi.

Aby uprościć prezentację modeli doktoranta, użyto następującej notacji dla systemu FAS:

$$\{\mathbf{P}_{\text{FAS}}, \mathbf{E}_{\text{FAS}}\} = \text{FAS}(\mathbf{A}_{\text{FAS}}, \mathbf{c}, C_{\text{FAS}}), \quad (1.5)$$

gdzie \mathbf{A}_{FAS} jest zbiorem natężenia oferowanego ruchu poszczególnych klas zgłoszeń:

$$\mathbf{A}_{\text{FAS}} = \{A_{\text{FAS},1}, A_{\text{FAS},2}, \dots, A_{\text{FAS},m}\}; \quad (1.6)$$

\mathbf{c} jest zbiorem żądań poszczególnych klas zgłoszeń wyrażonych w jednostkach alokacji:

$$\mathbf{c} = \{c_1, c_2, \dots, c_m\}; \quad (1.7)$$

i \mathbf{E}_{FAS} , \mathbf{B}_{FAS} są odpowiednio zbiorami prawdopodobieństw blokady poszczególnych klas oraz prawdopodobieństw strat poszczególnych klas zgłoszeń:

$$\mathbf{E}_{\text{FAS}} = \{E_{\text{FAS},1}, E_{\text{FAS},2}, \dots, E_{\text{FAS},m}\}, \quad (1.8)$$

$$\mathbf{B}_{\text{FAS}} = \{B_{\text{FAS},1}, B_{\text{FAS},2}, \dots, B_{\text{FAS},m}\}. \quad (1.9)$$

1.5.2 Modele systemów zależnych od stanu

W ramach modeli systemów zależnych od stanu można wyróżnić dwa podstawowe modele: model systemu o ograniczonej dostępności (ang. *Limited Available System* – LAS) oraz model idealnego systemu niepełnodostępnego (ang. *Erlang Ideal System* – EIS). Model LAS został zaproponowany w pracy [48], natomiast model EIS w pracy [49]. W systemach tych pojawiające się na wejściu zgłoszenia mają dostęp nie do wszystkich zasobów systemu, a jedynie do ograniczonej ich części.

1.5.2.1 Model LAS

Model systemu o ograniczonej dostępności (LAS) jest modelem opisującym w inżynierii ruchu rzeczywisty system zbudowany z k podsystemów o pełnej dostępności [48]. Schemat takiego modelu został przedstawiony na rysunku 1.3. Pełna dostępność każdego podsystemu w tym przypadku oznacza, że pojawiające się na wejściu zgłoszenie posiada dostęp do wszystkich zasobów każdego z podsystemów. Pojemność każdego z podsystemów LAS jest równa $\frac{C_{\text{LAS}}}{k}$, gdzie C_{LAS} jest pojemnością całego systemu. System LAS obsługuje m klas żądań, wymagających c_i JA ($0 \leq i \leq m$) do obsługi.

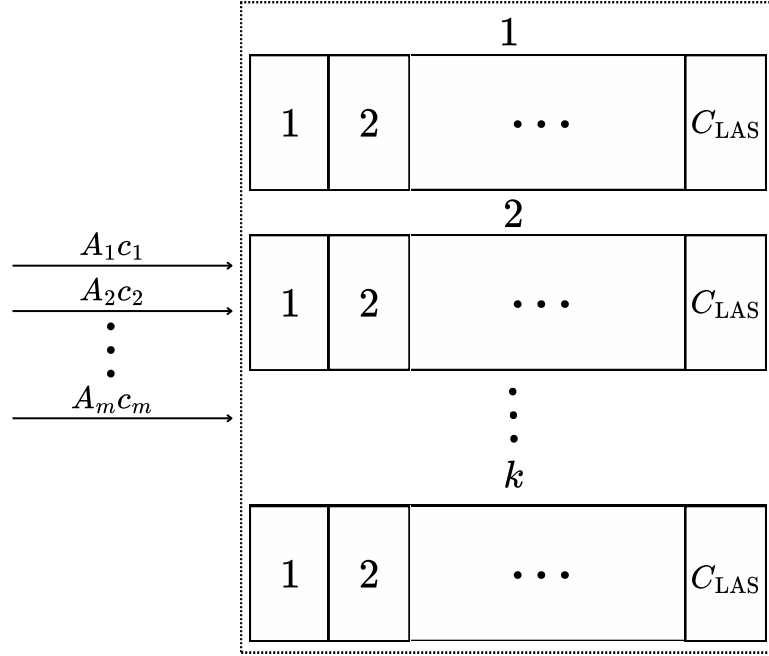
Zgłoszenia pojawiające się na wejściu do systemu LAS mogą zostać w nim obsłużone wtedy i tylko wtedy, gdy chociaż w jednym z podsystemów LAS jest wystarczająca liczba wolnych zasobów (JA) do ich obsługi. Co za tym idzie, nie jest możliwe podzielenie żądanych przez zgłoszenie c_i zasobów (JA) pomiędzy różne podsystemy.

Model analityczny dla systemów LAS został zaproponowany w [48], a idea systemu została przedstawiona na rysunku 1.3. Zgodnie z tym modelem, rozkład zajętości w LAS można określić w następujący sposób:

$$n[P_{\text{LAS}}(n)]_{kC_{\text{LAS}}} = \sum_{i=1}^m k A_{\text{LAS},i} c_i \sigma_{\text{LAS},i}(n - c_i) [P(n - c_i)]_{kC_{\text{LAS}}}, \quad (1.10)$$

gdzie:

- $A_{\text{LAS},i}$ to natężenie ruchu klasy ruchu i oferowane w systemie LAS;



Rysunek 1.3: Ilustracja konstrukcji modelu LAS [48].

- $[P(n)]_{kC_{\text{LAS}}}$ to prawdopodobieństwo zajęcia n JA w LAS o całkowitej pojemności C_{LAS} JA;
- $\sigma_{\text{LAS},i}(n)$ – to warunkowe prawdopodobieństwo przejścia pomiędzy sąsiednimi stanami zajęcia w LAS:

$$\sigma_{\text{LAS},i}(n) = 1 - \frac{F(kC_{\text{LAS}} - n, k, c_i - 1)}{F(kC_{\text{LAS}} - n, k, C_{\text{LAS}})}, \quad (1.11)$$

gdzie $F(x, k, c)$ to liczba możliwych rozmieszczeń x wolnych JA w k oddzielnych zasobach, gdzie każdy z zasobów ma pojemność równą $\frac{C_{\text{LAS}}}{k}$ JA:

$$F(x, k, f) = \sum_{i=0}^{\lfloor \frac{x}{f+1} \rfloor} (-1)^i \binom{k}{i} \binom{x+k-1-i(f+1)}{k-1}. \quad (1.12)$$

Na podstawie wyznaczonego rozkładu zajęcia możliwym jest określenie prawdopodobieństwa blokady dla każdej klasy zgłoszeń obsługiwanych w systemie na podstawie wzoru:

$$E_{\text{LAS},i} = \sum_{n=0}^{kC_{\text{LAS}}} [P(n)]_{\text{LAS}} (1 - \sigma_{\text{LAS},i}(n)). \quad (1.13)$$

W uproszczony sposób wyniki modelowania LAS, można przedstawić w następujący sposób:

$$\{\mathbf{P}_{\text{LAS}}, \mathbf{E}_{\text{LAS}}\} = \text{LAS}(\mathbf{A}_{\text{LAS}}, \mathbf{c}, kC_{\text{LAS}}), \quad (1.14)$$

gdzie \mathbf{A}_{LAS} zbiorem natężeń ruchu zgłoszeń oferowanych w systemie:

$$\mathbf{A}_{\text{LAS}} = \{A_{\text{LAS},1}, A_{\text{LAS},2}, \dots, A_{\text{LAS},m}\}; \quad (1.15)$$

\mathbf{c} jest zbiorem żądań poszczególnych klas zgłoszeń występujących w systemie:

$$\mathbf{c} = \{c_1, c_2, \dots, c_m\}; \quad (1.16)$$

\mathbf{E} jest zbiorem prawdopodobieństw blokady dla poszczególnych klas zgłoszeń w systemie LAS:

$$\mathbf{E}_{\text{LAS}} = \{E_{\text{LAS},1}, E_{\text{LAS},2}, \dots, E_{\text{LAS},m}\}. \quad (1.17)$$

natomiast \mathbf{P} reprezentuje rozkład zajętości w LAS:

$$\mathbf{P}_{\text{LAS}} = \{P_{\text{LAS},1}, P_{\text{LAS},2}, \dots, P_{\text{LAS},m}\}. \quad (1.18)$$

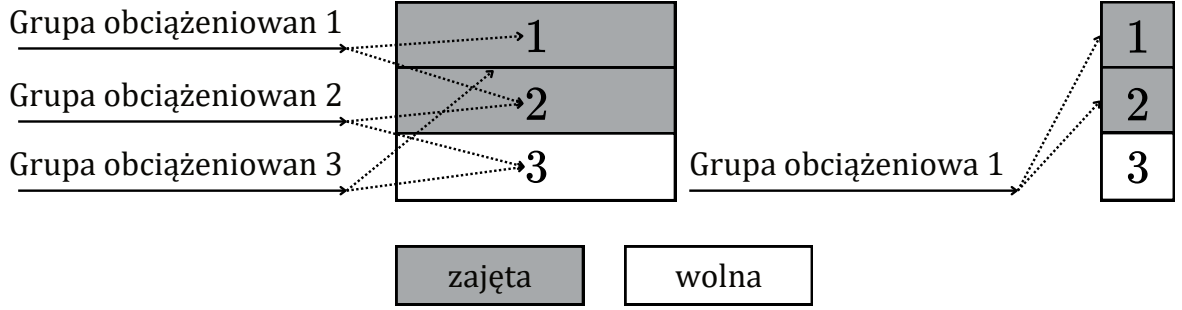
W dalszych badaniach doktorant założył, że model systemu LAS będzie aproksymować strukturę zewnętrzną systemu chmurowego, składającego się k maszyn fizycznych o jednakowych pojemnościach.

1.5.2.2 Model EIS

Idealny system niepełnodostępny (EIS) jest to abstrakcyjny system, w którym zgłoszenia pojawiające się na jego wejściu nie mają dostępu do wszystkich jego C_{EIS} , a jedynie ograniczonej ich części, nazywanej dostępnością d_i (rysunek 1.4). Konceptję oraz model analityczny dla ruchu jednosłużbowego w takim systemie został zaproponowany w pracy [49]. Zgłoszenia, które mają dostęp do tych samych zasobów systemu EIS, nazywane są grupą obciążeniową. Liczbę grup obciążeniowych w systemie (g_i), dla klasy i , można obliczyć na podstawie następującego wzoru:

$$g_i = \binom{C_{\text{EIS}}}{d_i} = \frac{C_{\text{EIS}}!}{(C_{\text{EIS}} - d_i)!d_i!}. \quad (1.19)$$

Parametr dostępności jest zdefiniowany dla każdej klasy zgłoszeń w systemie EIS, a jego wartość zależy od struktury systemu, mechanizmu kontroli dostępu do zasobów lub określonych wymagań co do samych zasobów. Warto zaznaczyć, że w obrębie jednej klasy zgłoszeń, rozumianej jako zgłoszenia o jednakowym rozmiarze i parametrze dostępności, niekoniecznie wszystkie zgłoszenia otrzymują dostęp do tych samych JA systemu EIS, a co za tym idzie mogą należeć do różnych grup obciążeniowych. Rysunek 1.4 ilustruje bardzo prosty przykład takiej sytuacji. Widoczny jest jednosłużbowy system EIS (występuje w systemie 1 klasa zgłoszeń i wymaga pojedynczego zasobu do obsługi) o pojemności 3 JA i dostępności $d = 2$ JA. W przedstawionej sytuacji, zgodnie ze wzorem (1.19), wyróżnić można 3 grupy obciążeniowe, posiadające dostęp kolejno do 1 i 2 zasobu w systemie, 2 i 3 zasobu w systemie oraz 1 i 3 zasobu w systemie. Na rysunku przedstawiony został stan blokady dla trzeciej grupy obciążeniowej, ponieważ choć w systemie występuje jeden zasób wolny (zasób nr 2 w systemie), zgłoszenie to nie posiada do niego dostępu i musi zostać odrzucone.



Rysunek 1.4: Ilustracja konstrukcji jednousługowego modelu EIS o pojemności 3 JA i dostępności 2 JA [J5].

Analogicznie jak w systemie LAS, wystarczająca liczba zasobów w całym systemie EIS nie jest gwarancją przyjęcia zgłoszenia do obsługi. Zgłoszenie pojawiające się na wejściu zostanie przyjęte do obsługi wtedy i tylko wtedy, gdy wystarczająca liczba wolnych zasobów będzie znajdowała się w ramach grupy obciążeniowej określonej dla tego zgłoszenia. Uwzględniając to założenie, warunkowy współczynnik przejścia w modelu EIS można określić w następujący sposób:

$$\sigma_{\text{EIS},i}(n) = 1 - \sum_{d_i - c_i + 1}^z \frac{\binom{d_i}{x} \binom{C_{\text{EIS}} - d_i}{n - x}}{\binom{C_{\text{EIS}}}{n}}, \quad (1.20)$$

gdzie:

$$z = n - c_i, \text{ if } (d_i - c_i + 1) \leq (n - c_i) \leq d_i,$$

$$z = d_i, \text{ if } (n - c_i) \geq d_i.$$

Po wyznaczeniu warunkowego współczynnika przejścia pomiędzy stanami, rozkład zajętości można wyznaczyć korzystając ze wzoru (1.10), podstawiając w miejsce warunkowego współczynnika przejścia, określonego dla modelu LAS, warunkowy współczynnik przejścia wyznaczony dla systemu EIS: $\sigma_i(n)_{\text{EIS}}$. Na tej podstawie możliwym jest wyznaczenie prawdopodobieństwa blokady w systemie EIS, zgodnie ze wzorem:

$$E_{\text{EIS},i} = \sum_{n=0}^{C_{\text{EIS}}} [1 - \sigma_{\text{EIS},i}(n)] [P_{\text{EIS}}(n)]_{C_{\text{EIS}}}. \quad (1.21)$$

Pomimo swojego teoretycznego charakteru, model EIS można wykorzystać do aproksymacji struktury wewnętrznej systemu chmurowego, w którym to żądania utworzenia maszyny wirtualnej, pojawiające się na wejściu pojedynczej maszyny fizycznej, posiadają dostęp do jednoznacznie określonej części zasobów tej maszyny.

W uproszczeniu, proces modelowania EIS można przedstawić symbolicznie w następujący sposób:

$$\{\mathbf{P}_{\text{EIS}}, \mathbf{E}_{\text{EIS}}\} = \text{EIS}(\mathbf{A}_{\text{EIS}}, \mathbf{c}, C_{\text{EIS}}), \quad (1.22)$$

gdzie \mathbf{P} reprezentuje rozkład zajętości w EIS, uzyskany na podstawie równania (1.10), w którym w miejsce warunkowego współczynnika przejścia, określonego dla modelu LAS, został wprowadzony warunkowy współczynnik przejścia wyznaczony dla systemu EIS, natomiast \mathbf{E} to zbiór prawdopodobieństw blokady uzyskanych za pomocą wzoru (1.21).

$$\mathbf{E}_{\text{EIS}} = \{E_{\text{EIS},1}, E_{\text{EIS},2}, \dots, E_{\text{EIS},m}\}. \quad (1.23)$$

$$\mathbf{P}_{\text{EIS}} = \{P_{\text{EIS},a}, P_{\text{EIS},2}, \dots, P_{\text{EIS},m}\}. \quad (1.24)$$

\mathbf{A}_{EIS} , \mathbf{c} i \mathbf{d} są zbiorami oferowanego ruchu i jego żądań oraz dostępności:

$$\mathbf{A}_{\text{EIS}} = \{A_{\text{EIS},1}, A_{\text{EIS},2}, \dots, A_{\text{EIS},m}\}, \quad (1.25)$$

$$\mathbf{c} = \{c_1, c_2, \dots, c_m\}, \quad (1.26)$$

$$\mathbf{d} = \{d_1, d_2, \dots, d_m\}. \quad (1.27)$$

1.5.2.3 Metoda ustalonego punktu

W metodzie ustalonego punktu (ang. *Fixed Point Method* - FP), problem określenia prawdopodobieństwa blokady dla zgłoszeń wymagających dostępu do kilku zasobów (podsystemów) w tym samym czasie jest rozwiązywany w oparciu o następujący schemat. Zakłada się, że danemu zasobowi jest oferowana tylko ta część całkowitego ruchu, która nie jest tracona w innych zasobach żądanych przez dane zgłoszenie. To założenie determinuje sposób określania prawdopodobieństwa blokady. Załóżmy, że klasa połączeń i ($0 \leq i \leq m$) żąda dostępu do s zasobów w tym samym czasie. Wartość ruchu oferowanego przez klasę i zasobowi j ($0 \leq j \leq s$) jest zdefiniowana następująco:

$$A_{\text{FP},i}(j) = A_{\text{FP},i} \prod_{l=1, l \neq j}^s [1 - E_{\text{FP},i}(l)], \quad (1.28)$$

gdzie $E_{\text{FP},i}(l)$ jest prawdopodobieństwem blokady dla zgłoszeń klasy i w zasobie (podsystemie) l . Należy zauważyć, że do określenia oferowanego ruchu $A_{\text{FP},i}(j)$ konieczna jest znajomość wartości prawdopodobieństwa blokady $E_{\text{FP},i}(l)$ w innych zasobach, tj. wszystkich $l \neq j$. Dlatego metoda FP jest metodą iteracyjną, której algorytm można zapisać następująco:

Krok 1: Inicjalizacja kroku iteracji: $z = 0$.

Krok 2: Określenie początkowych przybliżeń dla prawdopodobieństwa blokady wszystkich klas zgłoszeń we wszystkich zasobach:

$$\bigwedge_i \bigwedge_l E_i^{(0)}(l) = 0. \quad (1.29)$$

Krok 3: Zwiększanie kroku iteracji:

$$z = z + 1. \quad (1.30)$$

Krok 4: Określenie wartości ruchu oferowanego $A_i^k(l)$ przez zgłoszenia klasy j zasobowi l :

$$\bigwedge_i \bigwedge_l A_i^{(z)}(j) = A_i \prod_{l=1, l \neq j}^s [1 - E_i^{(z-1)}(l)]. \quad (1.31)$$

Krok 5: Określenie prawdopodobieństwa blokady dla wszystkich klas zgłoszeń we wszystkich zasobach:

$$\bigwedge_i \bigwedge_l E_i^{(z)}(j) = \text{FUN}(A^{(z)}, \mathbf{c}_j, C_j), \quad (1.32)$$

gdzie FUN jest modelem określającym prawdopodobieństwo blokady poszczególnych klas zgłoszeń w poszczególnych zasobach, np. FAS, a \mathbf{c}_j jest zbiorem żądań poszczególnych klas w zasobie j .

Krok 6: Określenie całkowitego prawdopodobieństwa blokady dla poszczególnych klas zgłoszeń:

$$\bigwedge_i E_i^{(z)} = \prod_{l=1, l \neq j}^s [1 - E_i^{(z)}(l)]. \quad (1.33)$$

Krok 7: Sprawdzanie warunku zakończenia procesu iteracyjnego:

$$\bigwedge_i \left| \frac{E_i^{(z)} - E_i^{(z-1)}}{E_i^{(z)}} \right| \leq \epsilon. \quad (1.34)$$

Jeśli warunek nie jest spełniony przynajmniej dla jednej klasy zgłoszeń, algorytm wraca do kroku 3; w przeciwnym razie $E_i = E_i^{(z)}$ i obliczenia kończą się.

W przedstawionym algorytmie zakłada się, że $X^{(z)}$ jest wartością parametru X w z -tym kroku iteracji. Parametr ϵ jest względnym błędem obliczeń, który określa dokładność procesu iteracji.

W celu skrócenia zapisu metody FP, umożliwiającej wyznaczenie prawdopodobieństwa blokady \mathbf{E} dla wszystkich klas ruchu, doktorant wprowadził następującą notację:

$$\mathbf{E} = \text{FP}(\mathbf{A}, \mathbf{c}_{\text{FP}}, \mathbf{C}_{\text{FP}}), \quad (1.35)$$

gdzie

- \mathbf{A} jest zbiorem wartości oferowanego ruchu (należy podkreślić, że ten zbiór \mathbf{A} jest taki sam dla wszystkich żądanych zasobów):

$$\mathbf{A} = \{A_1, A_2, \dots, A_m\}; \quad (1.36)$$

- \mathbf{c}_{FP} jest zbiorem żądań poszczególnych klas w poszczególnych zasobach:

$$\mathbf{c}_{\text{FP}} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_s\}, \quad (1.37)$$

gdzie

$$\mathbf{c}_l = \{c_1, c_2, \dots, c_m\}, 1 \leq l \leq s; \quad (1.38)$$

- \mathbf{C}_{FP} to zbiór pojemności poszczególnych zasobów:

$$\mathbf{C}_{\text{FP}} = \{\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_s\}. \quad (1.39)$$

1.6 Opis głównych osiągnięć doktoranta

1.6.1 Złożona struktura systemu

Idea obsługi nowych żądań maszyn wirtualnych w ramach systemu chmurowego działającego w modelu IaaS została schematycznie przedstawiona na rys. 1.5 [J4]. W rozważanym systemie chmurowym, nowe żądania utworzenia maszyn wirtualnych są odbierane przez urządzenia sterujące MRM (ang. *Main Resources Manager*). Aby uprościć zarządzanie zasobami, maszyny fizyczne (MF), np. serwery, są grupowane w zestawy o określonej funkcjonalności w wymaganej liczbie k . Każdy zestaw jest kontrolowany przez dedykowane urządzenie GM (ang. *Group Manager*), które jest odpowiedzialne za aktywację maszyn wirtualnych. Maszyny fizyczne mogą być grupowane w zależności od fizycznej lokalizacji urządzeń lub ich konkretnych specyfikacji (grupa składa się z urządzeń o identycznych lub podobnych specyfikacjach). MRM wysyła nowe żądanie do wybranego GM w celu aktywacji nowej maszyny wirtualnej. GM przekazuje informacje o bieżącym obciążaniu MF do MRM. Każda MF jest opisana przez następujące zasoby, które są używane do aktywacji nowych maszyn wirtualnych [J4, C3]:

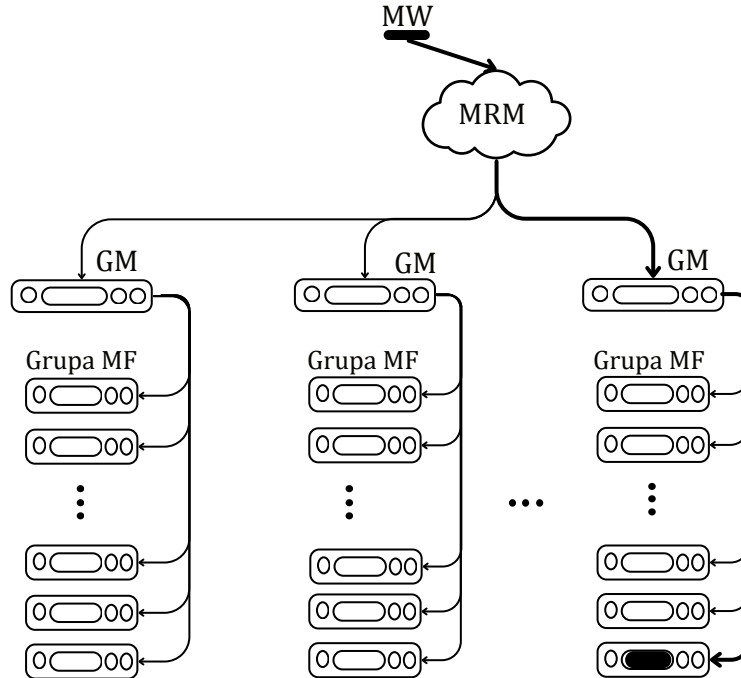
- C_P - liczba rdzeni procesora,
- C_R - całkowita pojemność pamięci RAM,
- C_D - całkowita pojemność dysku twardego,
- C_B - całkowita przepływność łącza sieciowego.

Żądanie utworzenia nowej maszyny wirtualnej klasy i można opisać czteroelementowym zbiorem $\mathbf{MW}_i = \{c_{i,P}, c_{i,R}, c_{i,D}, c_{i,B}\}$, gdzie:

- $c_{i,P}$ - liczba żądanych procesorów (rdzeni),
- $c_{i,R}$ - zapotrzebowanie na pojemność pamięci RAM,
- $c_{i,D}$ - wymagana pojemność dysku twardego,
- $c_{i,B}$ - wymagana prędkość łącza sieciowego,

gdzie i oznacza klasę maszyny wirtualnej, rozumianą jako grupę maszyn wirtualnych wymagających identycznych wartości parametrów zbioru \mathbf{MW}_i . Zazwyczaj przyjmuje się, że liczba klas maszyn wirtualnych jest równa m .

Po otrzymaniu nowego żądania, MRM próbuje zlokalizować maszynę fizyczną (MF), której zasoby można wykorzystać do aktywowania nowej maszyny wirtualnej (MW). Proces ten wykonywany jest zgodnie z zaimplementowanym w systemie chmurowym algorytmem alokacji zasobów fizycznych. Po wybraniu MF, informacja o nowej maszynie wirtualnej jest przekazywana do GM, który jest odpowiedzialny za aktywację maszyny. Liczba maszyn, które można aktywować na jednej MF, zależy od fizycznych zasobów MF i zestawu parametrów opisujących MW. Koniecznym jest, aby żądana maszyna wirtualna mogła być uruchomiona z wykorzystaniem zasobów tylko jednej maszyny fizycznej. Przykładowe parametry maszyn wirtualnych, oferowanych przez firmę Microsoft w ramach usługi Azure, zostały zaprezentowane w tabeli 1.1.



Rysunek 1.5: Idea działania systemu [J2].

Tabela 1.1: Parametry maszyn wirtualnych typu D2ds — D64ds v4 oferowanych przez Microsoft [10]

Instance	vCPU	RAM [GB]	Storage [GB]
D2ds v4	2	8	75
D4ds v4	4	16	150
D8ds v4	8	32	300
D16ds v4	16	64	600
D32ds v4	32	128	1200
D48ds v4	48	192	1800
D64ds v4	64	256	2400

Do zebranych w tabeli 1.1 maszyn wirtualnych w ramach oferty firmy Microsoft istnieje możliwość dodatkowego rozszerzenia przestrzeni dyskowej o dyski HDD i/lub SSD za dodatkową opłatą [10]. Taka rozbudowa wiąże się z ograniczeniem dostępności do konkretnych typów dysków zamontowanych w serwerze.

1.6.2 Modele jednoparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego.

Zgodnie z opisem umieszczonym w punkcie 1.6.1, system chmurowy może być rozpatrywany jako złożona struktura składająca się z maszyn fizycznych (połączonych ze sobą w grupę niezależnych urządzeń), którym oferowane są zgłoszenia opisane kilkoma

parametrami (żądana liczba rdzeni procesora, przestrzeni dyskowej, pamięci RAM oraz szybkości medium transmisyjnego). Jednocześnie, modele analityczne znane w literaturze przedmiotu reprezentują system chmurowy bardzo często przy pomocy jednego parametru, np. wyrażając żądanie jedynie w vCPU [34, 35, 36, 37, 38]. Takie podejście na potrzeby tworzenia modeli analitycznych, pomimo charakterystyki przyjętej przez dostawców usług, posiada swoje uzasadnienie. W wyniku dokładnej analizy parametrów typowych maszyn wirtualnych, umożliwiających realizację usług o różnych charakterystykach, np. przez Amazon i Microsoft, można zauważyć, że wartości pozostałych parametrów maszyn wirtualnych zwiększają się proporcjonalnie do wzrostu parametru vCPU. Dla przykładu, parametry maszyn wirtualnych oferowanych przez Microsoft Azure typu D2ds – D64ads v4 [10] zostały zebrane w Tabeli 1.1.

Na potrzeby opracowania modelu analitycznego rozważanego systemu, wymagane jest przeprowadzenie procesu dyskretyzacji i wyrażenia żądań w bezwymiarowych jednostkach alokacji zasobów [JA]. Są one wyznaczone jako najmniejszy wspólny dzielnik poszczególnych żądań. W przypadku maszyn wirtualnych oferowanych przez Microsoft typu D2ds – D64ads v4 (Tabela 1.1), pojedyncza jednostka alokacji może odpowiadać 2 rdzeniom procesora, 8 GB RAMu oraz 75 Gb przestrzeni dyskowej.

Tabela 1.2: Parametry maszyn wirtualnych typu D2ds — D64ds v4 wyrażone w JA

Instancja	vCPU [JA]	RAM [JA]	Dyks [JA]
D2ds v4	1	1	1
D4ds v4	2	2	2
D8ds v4	4	4	4
D16ds v4	8	8	8
D32ds v4	16	16	16
D48ds v4	24	24	24
D64ds v4	32	32	32

Jak łatwo zauważyć, po dokonaniu procesu dyskretyzacji poszczególne żądania dla maszyn wirtualnych są jednakowe. W związku z tym można sprowadzić je do pojedynczego parametru i na potrzeby obliczeń przyjąć reprezentację maszyny wirtualnej przez pojedynczą wartość. Gdyby jednak w systemie zastosowane skalowanie nie było możliwe, należałoby rozważyć system wieloparametryczny, reprezentując zarówno MF, jak i MW, przy pomocy więcej niż jednego parametru.

Dobór odpowiedniego modelu analitycznego zależy od przyjętej dostępności maszyn fizycznych. Gdy przyjmiemy, że wszystkie zasoby maszyny fizycznej są dostępne do utworzenia maszyny wirtualnej, to system ten można aproksymować przy pomocy modelu LAS (1.5.2.1). Nie zawsze jednak założenie to jest prawdziwe. Na przykład, w przypadku żądania dostępu do konkretnych zasobów systemu (dysk SSD/HDD, procesor o wybranej częstotliwości taktowania itp.). Wówczas rozważyć należy system o niepełnodostępnej strukturze wewnętrznej. Charakterystyka takiego systemu została opisana w [J1, J5, C1, C2], a na potrzeby rozprawy doktorskiej zebrana w podrozdziale 1.6.2.1.

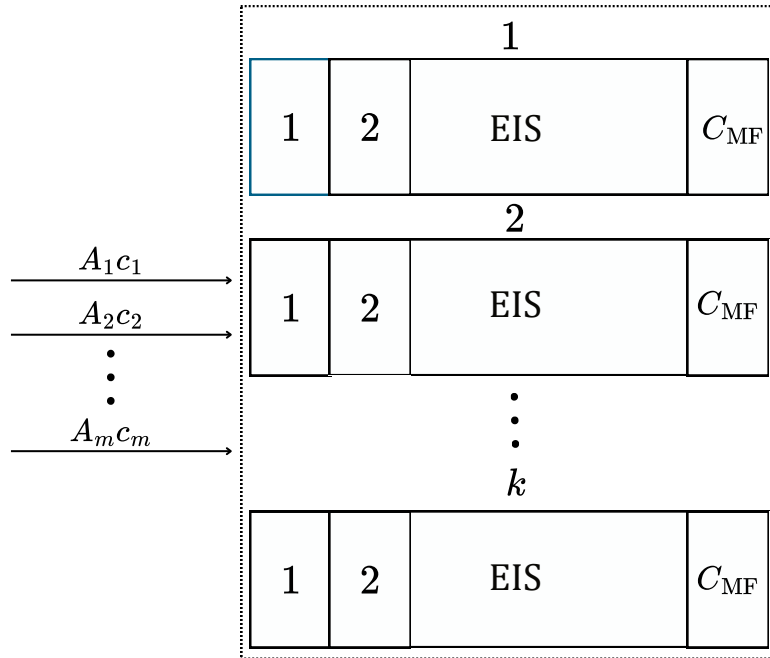
1.6.2.1 Opis i struktura jednoparametrycznego złożonego systemu chmurowego

Jak zostało opisane w punkcie 1.6.1, model systemu chmurowego opiera się na idei przedstawionej na rysunku 1.5. W modelu tym pojawienie się żądania utworzenia nowej maszyny wirtualnej (MW) trafia do urządzenia nazywanego Group Manager (GM). Urządzenie to odpowiedzialne jest za kontrolowanie zestawu maszyn fizycznych i aktywację maszyny wirtualnej na jednej z nich, zgodnie z przyjętym algorytmem dystrybucji. W swoich badaniach doktorant przyjął – jako algorytm dystrybucji – równomierny rozkład maszyn wirtualnych w zestawie maszyn fizycznych, co skutkuje minimalizacją strat. Rozważany przypadek ma zatem na celu optymalizację z punktu widzenia wydajności ruchu w systemie chmurowym.

W związku z założeniem, że zarówno żądanie utworzenia maszyny wirtualnej, jak i pojemność maszyn fizycznych (MF), może być opisane jednym parametrem, pojemność pojedynczej maszyny fizycznej może zostać określana jako C_{MF} . Jednocześnie, zapotrzebowanie na utworzenia nowej maszyny wirtualnej klasy i można opisać jako $c_{MW,i}$. Liczba różnych instancji maszyn wirtualnych (klas zgłoszeń) jest równa m .

Po otrzymaniu przez GM nowego żądania utworzenia MW, podejmuje on próbę wybrania MF w celu jej aktywacji. Liczba MW, które można aktywować na jednej MF, zależy od fizycznych zasobów (pojemności) MF i żądania opisującego MW. Przyjęto, że nowe żądanie utworzenia MW klasy i może zostać przyjęte do obsługi wtedy tylko wtedy, gdy co najmniej jedna MF w danym momencie ma wolne zasoby, które spełniają wymagania $c_{MW,i}$ MW. Zasoby te jednak nie mogą być dowolnymi zasobami maszyny, a jedynie ich szczegółowo określoną częścią. Liczba zasobów, do których dostęp w MF posiada instancja MW określona jest przez parametr dostępności d_i . Wymaganie konkretnych zasobów od maszyny fizycznej, jak zostało wspomniane wcześniej, może mieć związek z budową wewnętrzną systemu, wymaganiem konkretnego obszaru w przestrzeni dyskowej lub oferowania usługi tylko w części systemu (szybkość wybranych dysków, różna szybkość poszczególnych procesorów itp.). W opracowywanych modelach doktorant przyjął, że napływ zgłoszeń do systemu ma charakter Poissonowski (ruch Erlanga). Jednocześnie pojemność całego systemu chmurowego reprezentowanego przez zestaw MF została wyrażona jako iloczyn liczby maszyn fizycznych k i pojemności pojedynczej maszyny C_{MF} . W przypadku, w którym na żadnej z k maszyn fizycznych w systemie nie ma wolnych zasobów w wymaganej grupie zasobów (grupie obciążeniowej), zgłoszenie to jest odrzucane i uznawane za stracone.

W swoich badaniach prowadzących do opracowania jednoparametrycznych modeli analitycznych doktorant przyjął, że system chmurowy ma łączną pojemność C_{CH} , podzieloną na k jednakowych maszyn fizycznych. W takim systemie pojedynczy MF jest traktowana jako zasób o pojemności $C_{MF} = \frac{C_{CH}}{k}$. Ponadto, każda indywidualna MF otrzymuje żądania alokacji maszyn wirtualnych klasy i ($1 \leq i \leq m$) w każdym z zasobów należących do C_{MF} . Żądanie liczby JA pojedynczej klasy MW oznaczone zostało jako c_i . Jednocześnie, w każdej z MF pojawiające się zgłoszenie posiada dostęp jedynie do części zasobów określonych przez parametr dostępności d_i , a ruch oferowany w całym systemie oznaczony jest jako wektor \mathbf{A} , składający się z ruchu oferowanego poszczególnych klas zgłoszeń (a_i).



Rysunek 1.6: Ilustracja konstrukcji złożonego systemu niepełnodostępnego [J1].

Doktorant na potrzeby opracowania jednoparametrycznych modeli złożonych systemów chmurowych przyjął w swoich badaniach, że struktura zewnętrzna systemu, reprezentowana przez grupę k niezależnych MF o jednakowych parametrach (rozmiarach) może być aproksymowana przy pomocy modelu LAS. Jednocześnie, doktorant założył, że struktura wewnętrzna takiego systemu, reprezentowana przez pojedynczą maszynę fizyczną, w której zgłoszenia utworzenia MW pojawiające się na wejściu posiadają dostęp do ograniczonych zasobów MF, może być aproksymowana przy pomocy modelu EIS. Ilustracja budowy takiego systemu została przedstawiona na rysunku 1.6. Na podstawie tych założeń doktorant zaproponował 3 modele jednoparametrycznych złożonych systemów niepełnodostępnych, opisujących system chmurowy w modelu IaaS.

1.6.2.2 Model L-E-CBP I

Pierwszy z zaproponowanych przez doktoranta modeli L-E-CBP I (LAS-EIS-Cloud Blocking Probability I) [C2] bazował na własnościach systemów zależnych od stanu opisanych w podrozdziale 1.5 rozprawy doktorskiej. W celu wyznaczenia rozkładu zajętości w złożonym systemie niepełnodostępnym, doktorant zaproponował sposób wyznaczenia warunkowego współczynnika przejścia pomiędzy stanami w całym systemie dla każdej klasy zgłoszeń, opisanego jako σ_i^{LE} . Zgodnie z założeniami, system składał się ze struktury zewnętrznej opisanej jako k niezależnych maszyn fizycznych (MF), aproksymowanej przy pomocy modelu LAS (1.5.2.1), oraz struktury wewnętrznej reprezentującej pojedynczą maszynę fizyczną, aproksymowanej przy pomocy modelu EIS (1.5.2.2). Co za tym idzie, w tak scharakteryzowanym systemie, zależność strumienia zgłoszeń od

stanu, a w konsekwencji brak pełnej dostępności, wynika z dwóch czynników:

- wszystkie JA w systemie są podzielone na k identycznych oddzielnych zasobów (maszyn fizycznych), podczas gdy zgłoszenie musi być obsłużone przez pojedynczą maszynę fizyczną,
- każda wydzielona maszyna fizyczna ma ograniczony dostęp do własnych JA, co w konsekwencji prowadzi do sytuacji, w której – pomimo wystarczającej liczby wolnych JA – maszyna wirtualna nie może zostać utworzona z powodu braku dostępu do wolnych JA w wydzielonej dla niej części.

Na podstawie ogólnych własności systemów zależnych od stanu wskazanych w 1.5, a w szczególności wzoru (1.1), warunkowy współczynnik przejścia w całym systemie może być wyrażony jako:

$$\sigma_i(n)^{\text{LE}} = \sigma_i(n)^{\text{SZ}} \sigma_i(n)^{\text{SW}}, \quad (1.40)$$

gdzie $\sigma_i(n)^{\text{SZ}}$ jest warunkowym współczynnikiem przejścia, który wynika ze struktury zewnętrznej (podziału na k maszyn fizycznych), podczas $\sigma_i(n)^{\text{SW}}$ wynika z niepełnej dostępności w obrębie pojedynczej maszyny fizycznej (struktury wewnętrznej).

Doktorant w swoich badaniach zaproponował wyznaczenie warunkowego współczynnika przejścia dla struktury zewnętrznej $\sigma_i(n)^{\text{SZ}}$ zgodnie z opisem zamieszczonym w punkcie 1.5.2.1, dla systemu składającego się z k maszyn fizycznych o pojemnościach C_{MF} (wzór (1.11)):

$$[\sigma_i(n)]_{kC_{\text{MF}}}^{\text{SZ}} = 1 - \frac{F(kC_{\text{MF}} - n, k, c_i - 1)}{F(kC_{\text{MF}} - n, k, C_{\text{MF}})}. \quad (1.41)$$

Następnie wartość warunkowego współczynnika przejść dla struktury wewnętrznej (dla zgłoszeń o dostępności d_i i pojemności C_{MF}) doktorant wyznaczył na podstawie zmodyfikowanego wzoru (1.20), zgodnie z opisem modelu EIS zawartym w punkcie 1.5.2.2:

$$[\sigma_i(n)]_{C_{\text{MF}}}^{\text{SW}} = 1 - \sum_{d_i - c_i + 1}^z \frac{\binom{d_i}{x} \binom{C_{\text{MF}} - d_i}{n - x}}{\binom{C_{\text{MF}}}{n}}, \quad (1.42)$$

gdzie:

$$z = n - c_i, \text{ if } (d_i - c_i + 1) \leq (n - c_i) \leq d_i,$$

$$z = d_i, \text{ if } (n - c_i) \geq d_i.$$

Aby móc zastosować wyznaczone współczynniki $\sigma_i(n)^{\text{SZ}}$ i $\sigma_i(n)^{\text{SW}}$ we wzorze (1.40), doktorant musiał rozwiązać pojawiający się problem, związany z tym, że rozkład zajętości pojedynczego wydzielonego zasobu w postaci maszyny fizycznej zmienia się w zakresie od 0 do C_{MF} , natomiast całej grupy składającej się z k maszyn fizycznych w zakresie od 0 do kC_{MF} . Doktorant zauważył, że w związku z zastosowaniem mechanizmu sterowania tworzeniem i przydzielaniem maszyn wirtualnych w grupie maszyn fizycznych o rozkładzie równomiernym uzasadnione jest przyjęcie założenia, że w stanie zajętości n średnie obciążenie każdej z maszyn jest takie samo i równe $(\frac{n}{k})$.

W związku z tym doktorant zaproponował – w ramach modelu L-E-CBP I – wyznaczenie warunkowego współczynnika przejść pomiędzy stanami w całym jednoparametrycznym złożonym systemie niepełnodostępnym na podstawie wzoru:

$$\begin{aligned} [\sigma_i(n)]_{kC_{MF}}^{LE} &= [\sigma_i(n)]_{kC_{MF}}^{SZ} [\sigma_i(\frac{n}{k})]_{kC_{MF}}^{SW} = \\ &= \left[1 - \frac{F(kC_{MF} - n, k, c_i - 1)}{F(kC_{MF} - n, k, C_{MF})} \right] \left[1 - \sum_{d_i=c_i+1}^z \frac{\binom{d_i}{x} \binom{C_{MF}-d_i}{\lfloor \frac{n}{k} \rfloor - x}}{\binom{C_{MF}}{\lfloor \frac{n}{k} \rfloor}} \right]. \end{aligned} \quad (1.43)$$

Wyznaczony współczynnik przejścia posłużył do wyznaczenia rozkładu zajętości w tym systemie zgodnie ze wzorem:

$$n[P(n)]_{kC_{MF}} = \sum_{i=1}^m = A_i c_i [\sigma_i(n)]^{LE} [P(n - c_i)]_{kC_{MF}}. \quad (1.44)$$

Ostatecznie prawdopodobieństwo blokady w jednoparametrycznym złożonym systemie niepełnodostępnym dla zgłoszeń klasy i zgodnie z modelem L-E-CBP I może być wyznaczone przy pomocy wzoru:

$$E_i = \sum_{n=0}^{kC_{MF}-c_i} [P(n)]_{kC_{MF}} [1 - [\sigma_i(n)]_{kC_{MF}}^{LE}] + \sum_{n=kC_{MF}-c_i+1}^{kC_{MF}} [P(n)]_{kC_{MF}}. \quad (1.45)$$

Podsumowując powyższe rozważania, opracowaną metodę wyznaczania prawdopodobieństwa blokady można przedstawić w następujący sposób:

- Krok 1:** Wyznaczenie dla każdej klasy zgłoszeń warunkowego współczynnika przejść pomiędzy poszczególnymi stanami w systemie LAS o pojemności kC_{MF} i ruchu oferowanym **A** zgodnie ze wzorem (1.41).
- Krok 2:** Wyznaczenie dla każdej klasy zgłoszeń warunkowego współczynnika przejść pomiędzy poszczególnymi stanami w systemie EIS o pojemności C_{MF} i ruchu oferowanym **A** zgodnie ze wzorem (1.42).
- Krok 3:** Wyznaczenie dla każdej klasy zgłoszeń warunkowego współczynnika przejść pomiędzy poszczególnymi stanami w złożonym systemie niepełnodostępnym zgodnie ze wzorem (1.43).
- Krok 4:** Wyznaczenie rozkładu zajętości w złożonym systemie niepełnodostępnym zgodnie ze wzorem (1.44).
- Krok 5:** Wyznaczenie prawdopodobieństwa blokady dla poszczególnych klas zgłoszeń w systemie na podstawie rozkładu zajętości wyznaczonym w kroku 4, zgodnie ze wzorem (1.45).

Aby zweryfikować działanie zaproponowanego modelu L-E-CBP I, doktorant porównał wyniki uzyskane za jego pomocą z wynikami uzyskanymi w procesie symulacji cyfrowej. W tym celu wykorzystany został model symulacyjny szczegółowo opisany

w podrozdziale 1.6.5 rozprawy doktorskiej. Do określenia wartości pojedynczego pomiaru, posłużyło wykonanie 10 serii symulacji, każda z 1 000 000 zgłoszeń klasy, która wymagała największej liczby JA do obsługi. Uzyskane wyniki zostały przedstawione w funkcji ruchu oferowanego pojedynczej JA w systemie:

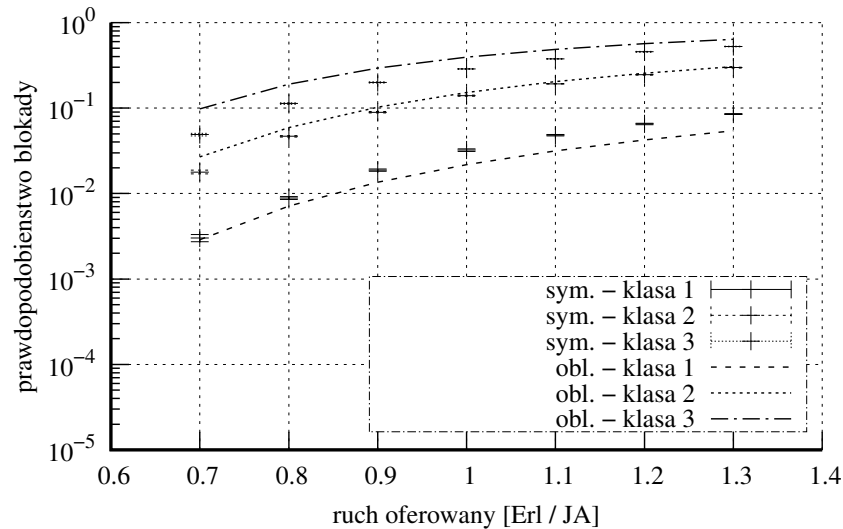
$$\mathbf{A} = \sum_{i=1}^m \frac{a_i c_i}{k C_{MF}}. \quad (1.46)$$

W badaniach przyjęto, że całkowity oferowany ruch jest dzielony między poszczególne klasy żądań utworzenia maszyny wirtualnej w następujących proporcjach: $a_1 c_1 : a_2 c_2 : \dots : a_m c_m = 1 : 1 : \dots : 1$.

Na rysunku 1.7 przedstawiono rezultaty modelowania uzyskane dla systemu scharakteryzowanego w tabeli 1.3.

Tabela 1.3: Parametry systemu – Model L-E-CBP I [C2]

System 1		
Liczba MF	Pojemność MF	
$k = 3$	$C_{MF} = 20$ JA	
Rozmiar MW		
MW klasy 1	$c_1 = 1$ JA	$d_1 = 75\% = 15$ JA
MW klasy 2	$c_2 = 2$ JA	$d_2 = 75\% = 15$ JA
MW klasy 3	$c_3 = 3$ JA	$d_3 = 75\% = 15$ JA



Rysunek 1.7: Prawdopodobieństwo blokady w systemie 1 [C2].

1.6.2.3 Model L-E-CBP II

Przeprowadzone badania wykazały, że zaproponowany w [C2] model L-E-CBP I wykazywał się niewystarczającą dokładnością dla klas zgłoszeń żądających najmniejszej liczby JA oraz w przypadku występowania dużej różnicy pomiędzy dostępnościami dla poszczególnych klas. Doktorant postanowił więc zaproponować nowy model, tj. Model L-E-CBP II, umożliwiający zwiększenie dokładności obliczeń prawdopodobieństwa blokady. Model ten został opublikowany w [J5], a dodatkowe szczegółowe jego porównanie z modelem L-E-CBP I [C2] zostało opisane w artykule [C1]. W modelu L-E-CBP II, w porównaniu z modelem L-E-CBP I, rozwiązany został dodatkowy problem związany wyznaczaniem warunkowego współczynnika przejścia między stanami w złożonym systemie niepełnodostępnym. Doktorant w swoich badaniach zauważył, że zgłoszenie utworzenia maszyny wirtualnej (MW) zostanie odrzucone w grupie maszyn fizycznych (MF) wtedy i tylko wtedy, gdy zostanie ono odrzucone na każdej maszynie fizycznej. W związku z tym, w ramach modelu L-E-CBP II, zaproponował modyfikacje wyznaczenia warunkowych współczynników przejść pomiędzy stanami w systemie.

W modelu L-E-CBP II założono, że algorytm przyjmowania zgłoszeń pozwala na sprawdzenie wszystkich wydzielonych zasobów (maszyn fizycznych), które mają co najmniej $c_{MW,i}$ wolnych JA. W celu uwzględnienia wpływu stanu każdego z wydzielonych zasobów na przyjęcie do obsługi nowego zgłoszenia utworzenia maszyny wirtualnej doktorant przyjął, że całkowity współczynnik przejścia, określający prawdopodobieństwo przejścia między stanami w jednoparametrycznym złożonym systemie niepełnodostępnym $\sigma_i(n)^{LE_2}$, będzie zależał iloczynowo od warunkowego współczynnika przejścia w strukturze zewnętrznej $[\sigma_i(n)]_{kC_{MF}}^{SZ}$ oraz warunkowego współczynnika przejść wynikającego ze struktury wewnętrznej $[\sigma_i(\frac{n}{k})]_{kC_{MF}}^{SW_2}$. Przyjęte założenie oznacza, że w stosunku do Modelu L-E-CBP I [C2], Model L-E-CBP II został rozszerzony o możliwość sprawdzenia dostępności wszystkich k maszyn fizycznych od momentu pojawienia się danego zgłoszenia. Zgłoszenie utworzenia maszyny wirtualnej zostanie przekazane do obsługi, gdy w co najmniej jednej z maszyn fizycznych w ramach danej grupy obciążeniowej będzie wystarczająca liczba dostępnych JA. W związku z tym, w modelu L-E-CBP II warunkowy współczynnik przejścia dla struktury wewnętrznej został wyznaczony zgodnie ze wzorem:

$$[\sigma_i(n)]_{C_{MF}}^{SW_2} = 1 - \left[1 - \left(\sum_{d_i=c_i+1}^z \frac{\binom{d_i}{x} \binom{C_{MF}-d_i}{\lfloor \frac{n}{k} \rfloor - x}}{\binom{C_{MF}}{\lfloor \frac{n}{k} \rfloor}} \right)^k \right]. \quad (1.47)$$

W Modelu L-E-CBP II nie zmieniono przyjętych założeń odnośnie struktury zewnętrznej, której wpływ na możliwość przyjmowania zgłoszeń określa współczynnik $[\sigma_i(n)]_{kC_{MF}}^{SZ}$. Całkowity warunkowy współczynnik przejścia w systemie $\sigma(n)_i^{LE_2}$ może być zatem określony wzorem:

$$[\sigma_i(n)]_{kC_{MF}}^{LE_2} = [\sigma_i(n)]_{kC_{MF}}^{SZ} \left[\sigma_i\left(\frac{n}{k}\right) \right]_{kC_{MF}}^{SW_2}. \quad (1.48)$$

Ostatecznie, uwzględniając wzór (1.48), rozkład zajętości w całym jednoparametrycznym złożonym systemie niepełnodostępnym może być wyznaczony na podstawie nastę-

pującego wzoru:

$$n[P(n)]_{kC_{MF}} = \sum_{i=1}^m A_i c_i [\sigma(n)_i]^{\text{LE}_2} [P(n - c_i)]_{kC_{MF}}. \quad (1.49)$$

Metodę obliczeń prawdopodobieństwa blokady zgodnie z Modelem II można zapisać w postaci następujących kroków:

- Krok 1:** Wyznaczenie dla każdej klasy zgłoszeń warunkowego współczynnika przejść pomiędzy poszczególnymi stanami w systemie LAS o pojemności kC_{MF} i ruchu oferowanym \mathbf{A} zgodnie ze wzorem (1.41).
- Krok 2:** Wyznaczenie dla każdej klasy zgłoszeń warunkowego współczynnika przejść pomiędzy poszczególnymi stanami w systemie EIS o pojemności C_{MF} i ruchu oferowanym \mathbf{A} zgodnie ze wzorem (1.47).
- Krok 3:** Wyznaczenie dla każdej klasy zgłoszeń warunkowego współczynnika przejść pomiędzy poszczególnymi stanami w złożonym systemie niepełnodostępnym zgodnie ze wzorem (1.48).
- Krok 4:** Wyznaczenie rozkładu zajętości poszczególnych stanów w złożonym systemie niepełnodostępnym zgodnie ze wzorem (1.49).
- Krok 5:** Wyznaczenie prawdopodobieństwa blokady dla poszczególnych klas zgłoszeń w systemie na podstawie rozkładu zajętości wyznaczonym w kroku 4, zgodnie ze wzorem:

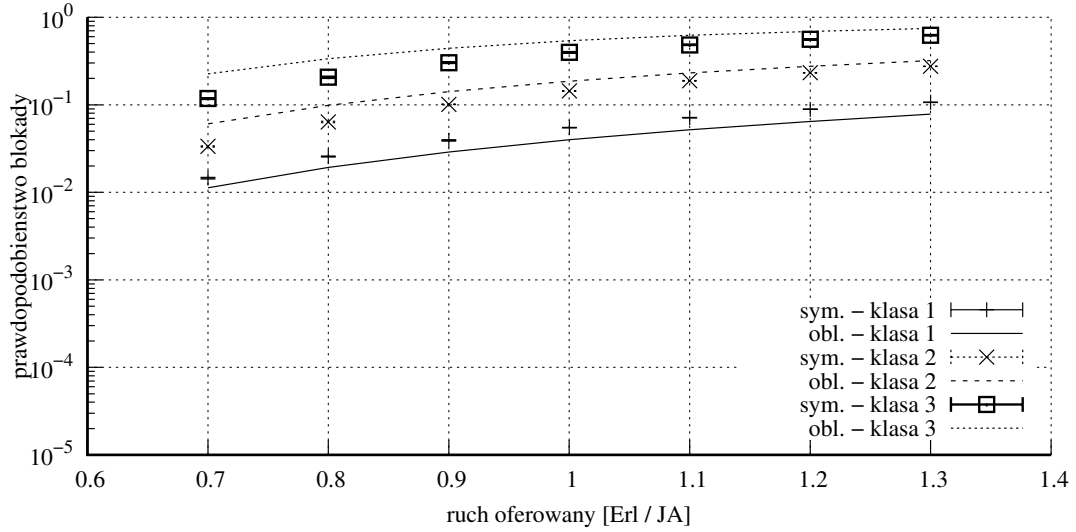
$$E_i = \sum_{n=0}^{kC_{MF}-c_i} [P(n)]_{kC_{MF}} [1 - [\sigma_i(n)]_{kC_{MF}}^{\text{LE}_2}] + \sum_{n=kC_{MF}-c_i+1}^{kC_{MF}} [P(n)]_{kC_{MF}}. \quad (1.50)$$

Podobnie jak w przypadku Modelu L-E-CBP I, opisanego w [C2], wyniki uzyskane za pomocą Modelu L-E-CBP II zostały porównane z wynikami uzyskanymi w wyniku przeprowadzenia symulacji cyfrowej. Założenia związane z liczbą serii symulacji oraz rozkładem ruchu zostały zachowane.

Rysunek 1.8 przedstawia wyniki dla systemu 2 scharakteryzowanego w tabeli 1.4. Analiza przedstawionych wyników wskazuje, że wprowadzona zmiana w Modelu L-E-CBP II przyczyniła się do znaczącej poprawy dokładności uzyskanych wyników, w przypadku niewielkich systemów oraz zmiennej dostępności dla każdej klasy zgłoszeń.

Tabela 1.4: Parametry systemu - Model L-E-CBP II [J5]

System 2		
Liczba MF	Pojemność MF	
$k = 3$	$C_{MF} = 18 \text{ JA}$	
Rozmiar MW		
MW klasy 1	$c_1 = 1 \text{ JA}$	$d_1 = 5 \text{ JA}$
MW klasy 2	$c_2 = 2 \text{ JA}$	$d_2 = 8 \text{ JA}$
MW klasy 3	$c_3 = 3 \text{ JA}$	$d_3 = 12 \text{ JA}$

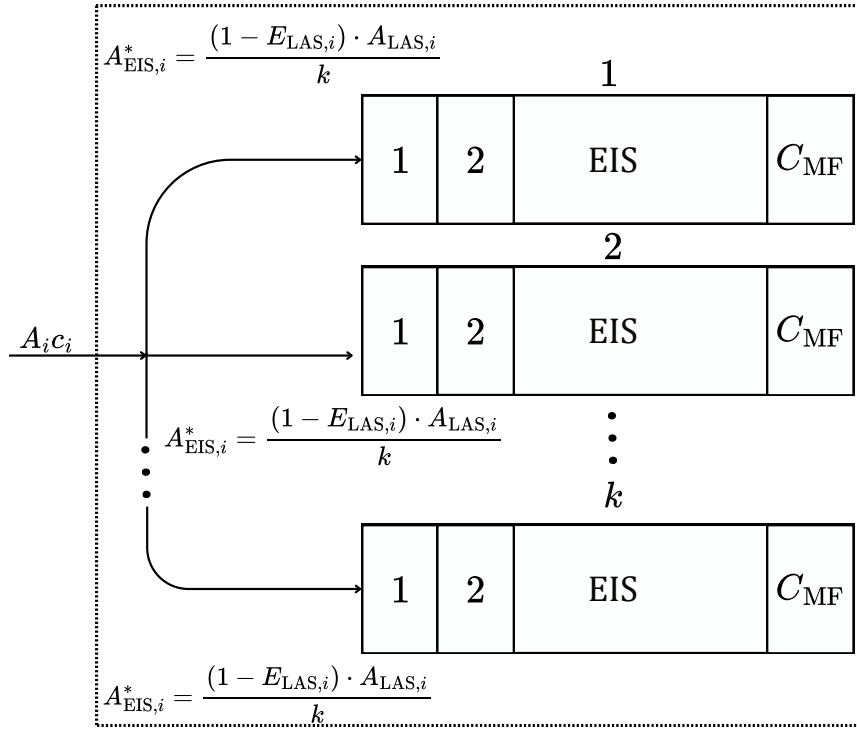


Rysunek 1.8: Prawdopodobieństwo blokady w systemie 2 [J5].

1.6.2.4 Model L-E-CBP III

W pracach [C3], [J5], doktorant zaproponował dwa modele analityczne systemu chmurowego, oparte na iloczynowym podejściu do określenia warunkowych prawdopodobieństw przejść w systemach, w których na zależność od stanu ma wpływ kilka niezależnych czynników [46]. W modelach tych, system chmurowy został potraktowany jako jednolity system z iloczynowo określonym, warunkowym prawdopodobieństwem przejścia pomiędzy stanami. Wadą takiego podejścia jest potrzeba przeskalowania stanu zajętości w pojedynczej maszynie fizycznej do stanu zajętości w całym systemie składającym się z k maszyn fizycznych, co w konsekwencji negatywnie wpływa na dokładność przyjętego rozwiązania.

W celu zwiększenia dokładności analitycznego modelowania systemów chmurowych doktorant w [J1] zaproponował nowy model L-E-CBP III (ang. LAS-EIS-Cloud Blocking Probability III). Zgodnie z wcześniej przedstawionymi rozważaniami, złożony charakter systemu chmurowego pozwala na zastosowanie podziału na strukturę zewnętrzną (grupa jednakowych maszyn fizycznych), aproksymowaną przez model LAS, oraz strukturę wewnętrzną (pojedyncza maszyna fizyczna), aproksymowaną modelem EIS. W Modelu L-E-CBP III, w pierwszym kroku określane jest prawdopodobieństwo strat wynikające ze struktury zewnętrznej systemu chmurowego, tj. systemu złożonego z k niez-



Rysunek 1.9: Ilustracja koncepcji opracowanego modelu [J1].

leżnych maszyn fizycznych o jednakowych parametrach, w których żądanie utworzenia maszyny wirtualnej może być obsłużone wyłącznie wtedy, gdy w choć jednej maszynie fizycznej znajdują się wolne zasoby, umożliwiające całkowitą obsługę maszyny wirtualnej, tj. bez uczestnictwa w obsłudze innych maszyn fizycznych. Następnie przyjmując, że algorytmy przydziału zasobów maszynom wirtualnym dążą do równomiernego obciążenia wszystkich maszyn fizycznych doktorant założył, że ruch oferowany poszczególnym maszynom fizycznym jest taki sam i wynika z podziału ruchu spływającego ze struktury zewnętrznej (tj. ruchu, który nie został stracony w strukturze zewnętrznej) na k strumieni ruchu o identycznym natężeniu. Taki ruch przez doktoranta został nazwany ruchem ekwiwalentnym i jest podstawą modelowania struktury wewnętrznej systemu chmurowego (pojedynczej maszyny fizycznej) na podstawie modelu EIS. W kolejnym kroku określone zostało rzeczywiste prawdopodobieństwo blokady w systemie chmurowym, uwzględniające zarówno dostępność w pojedynczej maszynie fizycznej jak i liczbę maszyn fizycznych w systemie. Na rysunku 1.9 przedstawiono koncepcję podziału ruchu w systemie chmurowym, zgodną z proponowanym modelem L-E-CBP III.

W zaproponowanym przez doktoranta modelu L-E-CBP III prawdopodobieństwo blokady klasy i , wynikające ze struktury zewnętrznej systemu chmurowego aproksymowane jest modelem LAS składającym się z k podsystemów, które z kolei są systemami EIS o pojemności $C_{MF} = C_{EIS} = \frac{C_{LAS}}{k}$ JA. Tak więc prawdopodobieństwo blokady w strukturze zewnętrznej systemu, zgodnie z opisem modelu LAS zawartym w punkcie

1.5.2.1, wyrażone zostało równaniem (1.13), które można w przyjętej notacji przepisać następująco:

$$E_{LAS,i} = \sum_{n=0}^{kC_{MF}} [1 - \sigma_{LAS,i}(n)] [P_{LAS}(n)]_{kC_{MF}}. \quad (1.51)$$

Doktorant zauważył, że prawdopodobieństwo blokady struktury zewnętrznej $\mathbf{E}_{LAS} = LAS(\mathbf{A}_{LAS}, \mathbf{c}, k, C_{MF})$ określane jest na podstawie ruchu oferowanego całemu systemowi chmurowemu, tj. ruchu \mathbf{A}_{LAS} . Natomiast prawdopodobieństwo strat struktury wewnętrznej $\mathbf{E}_{EIS} = EIS(\mathbf{A}_{EIS}^*, \mathbf{c}, \mathbf{d}, C_{MF})$ jest określane na podstawie ruchu ekwiwalentnego \mathbf{A}_{EIS}^* , gdzie każdy element wektora określający ruch oferowany każdej z klas przypadający na pojedynczą maszynę fizyczną $A_{EIS,i}^*$ określony jest wzorem:

$$A_{EIS,i}^* = \frac{(1 - E_{LAS,i}) A_{LAS,i}}{k}. \quad (1.52)$$

Tak określony ruch ekwiwalentny – oferowany pojedynczej maszynie fizycznej – jest ruchem, który nie jest tracony w strukturze zewnętrznej. Dzięki jego wyznaczeniu, możliwym jest określenie rzeczywistego prawdopodobieństwo blokady w pojedynczej maszynie fizycznej z uwzględnieniem zarówno dostępności, jak i struktury systemu (tj. zestawu maszyn fizycznych). Przyjmując założenie o aproksymacji struktury wewnętrznej przy pomocy modelu EIS (p. 1.5.2.2), prawdopodobieństwo blokady może być wyznaczone na podstawie zmodyfikowanego wzoru (1.21), do którego podstawione są parametry złożonego systemu uwzględniające zarówno ruch oferowany (ruch ekwiwalentny), jak i rozmiar pojedynczej maszyny fizycznej. W celu otrzymania prawdopodobieństwa blokady dla poszczególnych klas zgłoszeń $E_{EIS,i}$ doktorant zaproponował wykorzystanie następującego równania:

$$E_{EIS,i} = \sum_{n=0}^{C_{MF}} [1 - \sigma_{EIS,i}(n)] [P_{EIS}(n)]_{C_{MF}}. \quad (1.53)$$

Całkowite prawdopodobieństwo blokady w systemie chmurowym \mathbf{E}_{CBP} , uwzględniające blokadę w strukturze zewnętrznej (grupie maszyn fizycznych) i k strukturach wewnętrznych (pojedynczych maszynach fizycznych) w modelu L-E-CBP III jest aproksymowane prawdopodobieństwem strat ruchu. Prawdopodobieństwo strat ruchu jest definiowane jako stosunek ruchu traconego do ruchu oferowanego danemu systemowi. Ruch tracony klasy i w strukturze zewnętrznej ma wartość $A_{LAS,i} E_{LAS,i}$. Ruch tracony w pojedynczej strukturze wewnętrznej (maszynie fizycznej) jest równy $A_{EIS,i}^* E_{EIS,i}$. Zatem całkowity ruch klasy i tracony w strukturze zewnętrznej i k strukturach wewnętrznych jest równy $A_{LAS,i} E_{LAS,i} + k A_{EIS,i}^* (E_{EIS,i})^k$.

Na podstawie definicji prawdopodobieństwa strat ruchu można aproksymować całkowite prawdopodobieństwo blokady dla żądań klasy i w systemie chmurowym. Z racji tego, że zgłoszenie, aby zostało odrzucone w całym systemie chmurowym, musi zostać odrzucone jednocześnie przez wszystkie maszyny fizyczne, doktorant zaproponował wyznaczenie prawdopodobieństwa blokady w całej grupie składającej się z k maszyn

fizycznych na podstawie następującego wzoru:

$$E_{\text{CBP},i} = \frac{[A_i E_{\text{LAS},i} + k A_i^* (E_{\text{EIS},i})^k]}{A_i}. \quad (1.54)$$

Wartości $E_{\text{CBP},i}$ są elementami zbioru \mathbf{E}_{CBP} obejmującego prawdopodobieństwa blokady dla wszystkich klas ruchu oferowanego systemowi chmurowemu.

Podsumowując poprzednie rozważania, metodę L-E-CBP III można przedstawić w następujący sposób:

- Krok 1:** Określenie na podstawie modelu LAS (p. 1.5.2.1) rozkładu zajętości oraz prawdopodobieństwa blokady w strukturze zewnętrznej systemu zgodnie ze wzorem (1.51).
- Krok 2:** Określenie ruchu ekwiwalentnego pojawiającego się na wejściu pojedynczej maszyny fizycznej, jako ruchu obsługowanego na poziomie struktury zewnętrznej podzielonego przez liczbę maszyn fizycznych, zgodnie ze wzorem (1.52).
- Krok 3:** Wyznaczenie rozkładu zajętości i prawdopodobieństwa blokady w pojedynczej maszynie fizycznej o pojemności $C_{\text{MF}} = \frac{C_{\text{LAS}}}{k}$, na wejściu której pojawia się oferowany ruch ekwiwalentny, określony w kroku 2. Na podstawie modelu EIS (p. 1.5.2.2) może on zostać określony na podstawie wzoru (1.53).
- Krok 4:** Określenie na podstawie wzoru (1.54) prawdopodobieństwa blokady w całym złożonym systemie, uwzględniając, że żądanie utworzenia maszyny wirtualnej musi zostać odrzucone we wszystkich k maszynach fizycznych, aby zostało odrzucone w całym systemie:

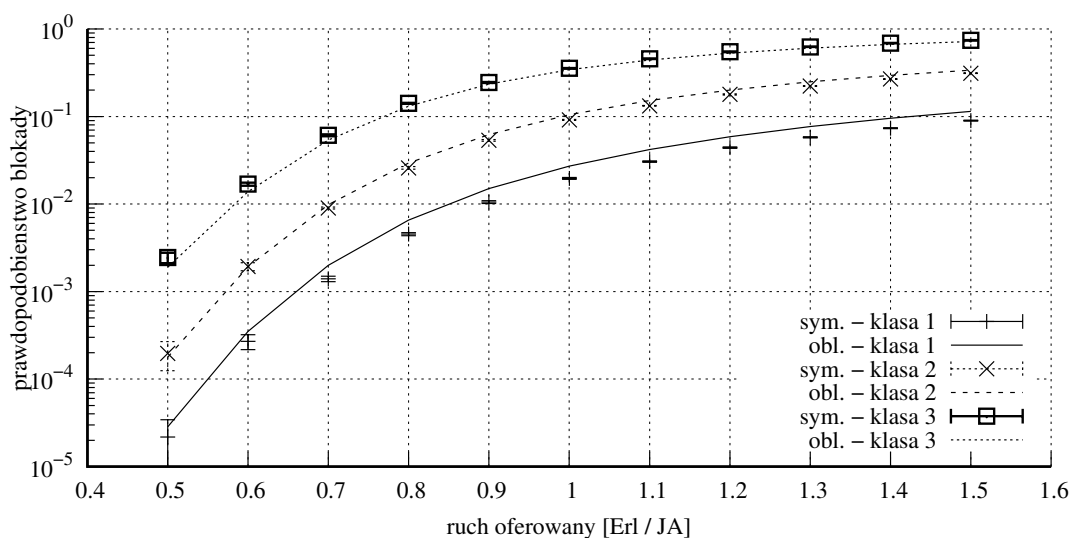
$$\mathbf{E}_{\text{CBP}} = \{E_{\text{CBP},i} : E_{\text{CBP},i} = \frac{A_{\text{LAS},i} E_{\text{LAS},i} + k A_{\text{EIS},i}^* (E_{\text{EIS},i})^k}{A_{\text{LAS},i}} \wedge 1 \leq i \leq m\}.$$

Analogicznie do poprzednio opisanych modeli, wyniki uzyskane przy pomocy modelu L-E-CBP III doktorant porównał z wynikami uzyskanymi przy pomocy modelu symulacyjnego (opisanego w p. 1.6.5). Wszystkie założenia związane z procesem symulacji są zgodne z przyjętymi przy opisie poprzednich modeli.

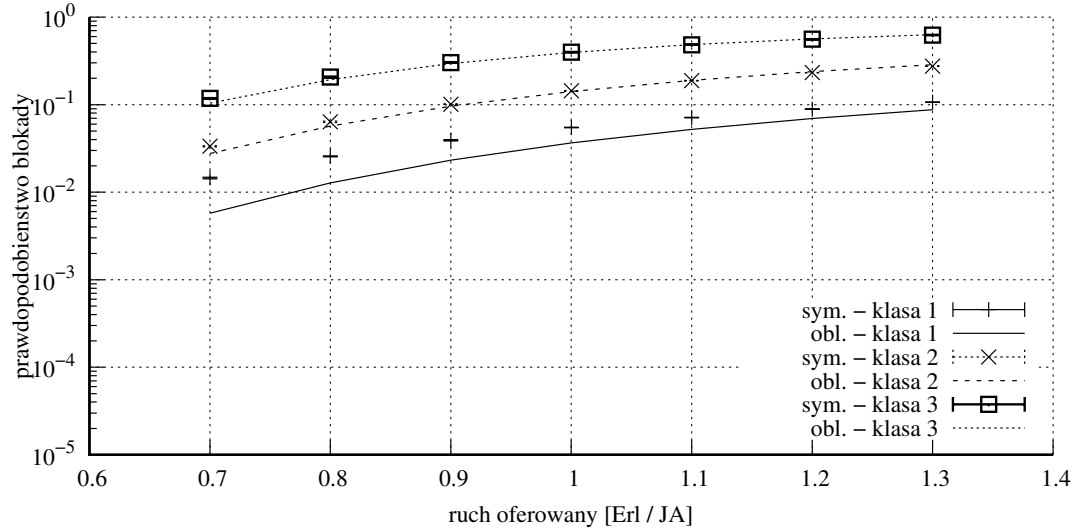
Rysunki 1.10 i 1.11 przedstawiają wyniki kolejno dla systemów 3 i 4, scharakteryzowanych w tabeli 1.5.

Tabela 1.5: Parametry systemów – Model L-E-CBP III [J1]

System 3		
Liczba MF	Pojemność MF	
$k = 3$	$C_{MF} = 25 \text{ JA}$	
Rozmiar MW		
MW klasy 1	$c_1 = 1 \text{ JA}$	$d_1 = 15 \text{ JA}$
MW klasy 2	$c_2 = 2 \text{ JA}$	$d_2 = 15 \text{ JA}$
MW klasy 3	$c_3 = 4 \text{ JA}$	$d_3 = 15 \text{ JA}$
System 4		
Liczba MF	Pojemność MF	
$k = 3$	$C_{MF} = 18 \text{ JA}$	
Rozmiar MW		
MW klasy 1	$c_1 = 1 \text{ JA}$	$d_1 = 5 \text{ JA}$
MW klasy 2	$c_2 = 2 \text{ JA}$	$d_2 = 8 \text{ JA}$
MW klasy 3	$c_3 = 3 \text{ JA}$	$d_3 = 12 \text{ JA}$



Rysunek 1.10: Prawdopodobieństwo blokady w systemie 3 [J1]



Rysunek 1.11: Prawdopodobieństwo blokady w systemie 4 [J1]

1.6.3 Modele wieloparametrycznego złożonego systemu teleinformatycznego z ruchem Erlanga typu strumieniowego

Analiza złożonych usług chmurowych wskazuje, że nie zawsze jest możliwe przeskalowanie żądań utworzenia maszyn wirtualnych w sposób umożliwiający opis żądanych zasobów za pomocą pojedynczego parametru, tak jak zostało to opisane w p. 1.6.2 rozprawy doktorskiej. Przykładem takiego rozwiązania mogą być maszyny wirtualne firmy Microsoft oferowane w ramach serii Av2 [10]. Parametry takich maszyn zostały przedstawione w tabeli 1.6.

Tabela 1.6: Parametry maszyn wirtualnych typu D2ds — D64ds v4 oferowanych przez Microsoft [10]

Instancja	vCPU	RAM [GB]	Dysk [GB]
D2ds v4	1	2	10
D4ds v4	2	4	20
D8ds v4	2	16	20
D16ds v4	4	8	40
D32ds v4	4	32	40
D48ds v4	8	16	80
D64ds v4	8	64	80

Wyznaczając żądania ekwiwalentne wyrażone w JA, maszyny te mogłyby zostać opisane zgodnie z parametrami wskazanymi w tabeli 1.7.

Analiza wartości żądań w tabeli 1.7 wskazuje, że nie jest możliwe określenie żądań z wykorzystaniem jednego parametru. Niezbędnym jest zatem uwzględnienie wszystkich parametrów zarówno maszyn fizycznych (MF), jak i maszyn wirtualnych (MW) niezależnie. Warto zwrócić uwagę, że nie zawsze pojemności maszyn fizycznych muszą

Tabela 1.7: Parametry maszyn wirtualnych typu D2ds — D64ds v4 oferowanych przez Microsoft [10] – wyrażone w AU

Instancja	vCPU [JA]	RAM [JA]	Dysk [JA]
D2ds v4	1	1	1
D4ds v4	2	2	2
D8ds v4	2	8	2
D16ds v4	4	4	4
D32ds v4	4	16	4
D48ds v4	8	8	8
D64ds v4	8	32	8

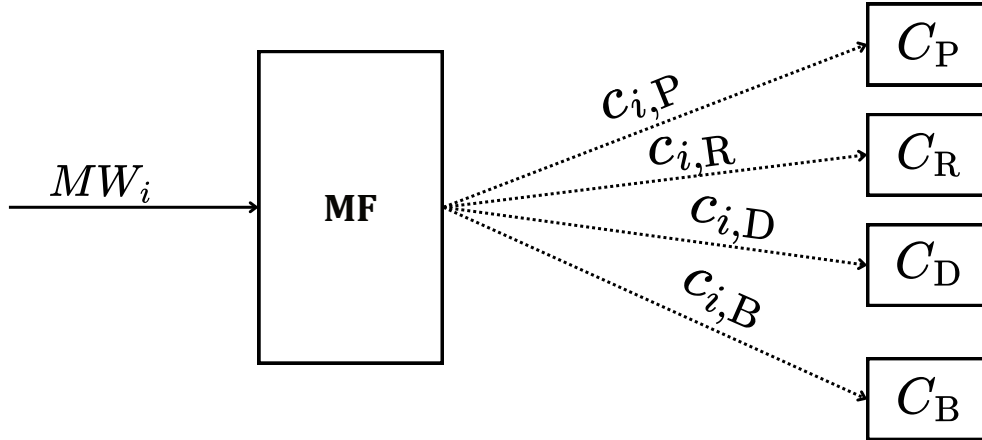
skalować się jednakowo dla wszystkich parametrów, co znacząco utrudnia zastosowanie modeli opisanych w p. 1.6.2 rozprawy doktorskiej. W związku z tym doktorant opracował model dla systemu składającego się z k jednakowych maszyn fizycznych opisanych parametrami wskazanymi w p. 1.6.1 rozprawy doktorskiej (procesor (P), RAM (R), dysk (D) i przepływność łącza (B)), którym oferowane są żądania utworzenia maszyn wirtualnych opisane za pomocą analogicznych parametrów (p. 1.6.1). Model wyznaczania prawdopodobieństwa blokady dla takiego systemu został opracowany i opisany przez doktoranta w [J4].

1.6.3.1 Opis i struktura wieloparametrycznego złożonego systemu chmurowego

W badaniach doktorant przyjął założenie, że nowe zgłoszenie klasy i w celu aktywacji maszyny wirtualnej zostanie obsłużone tylko wtedy, gdy co najmniej jedna maszyna fizyczna w danym momencie ma wolne zasoby, które spełniają wymagania MW_i tego zgłoszenia. Kolejnym przyjętym założeniem było to, że algorytm wyboru maszyny fizycznej zapewnia wyrównanie średnich obciążeń wszystkich maszyn fizycznych w systemie chmurowym (grupie k jednakowych maszyn fizycznych). Ponieważ czas życia maszyny wirtualnej znacząco przekracza czas oczekiwania na jej utworzenie, doktorant postanowił w badaniach nie uwzględniać czasu oczekiwania na aktywację maszyny wirtualnej. Jednocześnie, jeżeli w momencie pojawienia się nowego żądania utworzenia maszyny wirtualnej w systemie nie ma wolnych zasobów niezbędnych do jej utworzenia, w żadnej z dostępnych maszyn fizycznych, zgłoszenie utworzenia maszyny wirtualnej zostaje odrzucone.

Zaproponowane przez doktoranta modele analityczne dla systemów chmurowych oferujących usługi IaaS [J2], [J4] uwzględniają następujące czynniki: fizyczną architekturę systemu, parametry opisujące maszyny wirtualne oraz fakt, że maszyna wirtualna musi zostać uruchomiona w całości tylko na jednej maszynie fizycznej (tj. wszystkie zasoby wymagane przez nową maszynę wirtualną znajdują się na jednym serwerze).

Aktywacja nowej maszyny wirtualnej z wykorzystaniem zasobów maszyny fizycznej może nastąpić tylko wtedy, gdy ta MF ma wystarczającą liczbę wolnych zasobów zdefiniowanych przez zgłoszenie MW_i . W modelach Cloud II i Cloud II przeprowadzana została dekompozycja MF na cztery niezależne komponenty systemowe: procesory, pamięć



Rysunek 1.12: Ilustracja koncepcji dekompozycji zasobów żądania utworzenia maszyny wirtualnej [J4].

RAM, dysk twardy i dostępna przepływność (rys. 1.1). Założono, że każdy komponent systemu obsługuje żądania, które są dla niego odpowiednie; tj. P będzie obsługiwać żądania związane z liczbą procesorów, R będzie obsługiwać żądania związane z pamięcią RAM i tak dalej.

Zgłoszenie klasy i może zostać przyjęte do obsługi w danej maszynie fizycznej tylko wtedy, gdy maszyna ta ma wolne JA niezbędne do obsłużenia tego zgłoszenia w każdym z jego podsystemów. Doktorant założył również, że w celu utworzenia maszyny wirtualnej w obrębie maszyny fizycznej dostępne są wszystkie jej wolne zasoby, bez dodatkowego ograniczenia w postaci dostępności, jak miało to miejsce w przypadku jednoparametrycznych modeli opisanych w punkcie 1.6.2.

1.6.3.2 Model Cloud I

Podstawą modelu Cloud I jest dekompozycja grupy serwerów opisanych 4 czterema parametrami (procesor (P), pamięć RAM (R), dysk (D), przepływność łącza (B)), na cztery grupy serwerów, z których każda opisana jest jednym z wyżej wymienionych parametrów. Ilustracja idei dekompozycji zasobów żądania utworzenia maszyny wirtualnej została przedstawiona na rysunku 1.12. Dzięki tej dekompozycji, możliwym jest zastosowanie modelu LAS niezależnie dla każdej z czterech grup serwerów, odpowiednio dla k serwerów z zasobem P, k serwerów z zasobem R, k serwerów z zasobem D i k serwerów z zasobem B, do modelowania systemu chmurowego.

Jednocześnie, doktorant zauważył, że zastosowanie jedynie modelu LAS do zdekomponowanej grupy serwerów prowadzi do mylących i błędnych wyników, gdyż otrzymane rozwiązanie może uwzględniać również takie stany (traktując je jako stany nieblokujące), w których dostępne zasoby jednego typu, np. P, znajdują się na jednej maszynie fizycznej, podczas gdy dostępne zasoby innego typu, np. R, znajdują się w zupełnie innej maszynie fizycznej.

W związku z tym, doktorant zaproponował wykorzystanie metody ustalonego punktu (opisanej w punkcie 1.5.2.3) do wyznaczenia prawdopodobieństwa blokady w zbiorze k

MF, wynikającej z niedostępności (blokady) przynajmniej jednego z żądanych typów zasobów w ramach tej samej MF. Zauważmy, że sama metoda ustalonego punktu pozwala na wyznaczenie prawdopodobieństwa blokady jedynie dla jednej maszyny fizycznej. W związku z tym, w celu zastosowania metody ustalonego punktu dla k MF, doktorant zaproponował metodę odwzorowania prawdopodobieństwa blokady w pojedynczej maszynie fizycznej (przy wykorzystaniu modelu FAS do modelowania pojedynczego zasobu) na system składających się z k maszyn fizycznych (przy wykorzystaniu modelu LAS).

W pracy [J4] doktorant zaproponował rozwiązanie, które może być opisane w postaci następującej metody obliczeniowej:

Krok 1: Określenie prawdopodobieństwa blokady w każdym z typów zasobów w pojedynczej MF:

$$\mathbf{e}_x = \text{FAS}(\mathbf{a}, \mathbf{c}_X, C_X), \quad (1.55)$$

gdzie zbiór $\mathbf{c}_x = \{c_{1,x}, c_{2,x}, \dots, c_{m,x}\}$ jest zbiorem żądań, wyrażonych w JA, poszczególnych klas zgłoszeń w danym typie zasobu $\mathbf{X} = \{P, R, D, B\}$.

Zbiór C_X określa rozmiar każdego z typów zasobów X w danej MF.

W celu wyznaczenia prawdopodobieństwa blokady z wykorzystaniem modelu FAS niezbędnym jest dokonanie dekompozycji zgłoszenia na poszczególne niezależne żądania.

Krok 2: Określenie prawdopodobieństwa blokady w grupie k zdekomponowanych typów zasobów X , gdzie $\mathbf{X} = \{P, R, D, B\}$:

$$\mathbf{E}_X = \text{LAS}(k\mathbf{a}, \mathbf{c}_X, kC_X). \quad (1.56)$$

Krok 3: Wyznaczenie zbioru $\rho_{\mathbf{X}}$ relacji między prawdopodobieństwem blokady w każdym z pojedynczych zdekomponowanych typów zasobów \mathbf{e}_X i grupie k zdekomponowanych typów zasobów \mathbf{E}_X :

$$\rho_{\mathbf{X}} = \{\rho_{1,X}, \rho_{2,X}, \dots, \rho_{m,X}\}, \quad (1.57)$$

gdzie:

$$\bigwedge_{i=\{1,2,\dots,m\}} \rho_{i,X} = \frac{E_{i,X}}{e_{i,X}}. \quad (1.58)$$

Krok 4: Wyznaczenie prawdopodobieństwa blokady w pojedynczym MF (składającej się jednocześnie z X zasobów) na podstawie metodologii ustalonego punktu 1.5.2.3:

$$\mathbf{e}^* = \text{FP}(\mathbf{a}, \mathbf{c}_{CC}, \mathbf{C}_{CC}), \quad (1.59)$$

$$\bigwedge_{X=\{P,R,D,B\}} \mathbf{e}^* = \{e_{1,X}^*, e_{2,X}^*, \dots, e_{m,X}^*\}, \quad (1.60)$$

gdzie

- \mathbf{C}_{CC} to zbiór pojemności wszystkich typów zasobów w danej MF:

$$\mathbf{C}_{CC} = \{C_P, C_R, C_D, C_B\}; \quad (1.61)$$

- \mathbf{c}_{CC} jest zbiorem żądań poszczególnych klas utworzenia maszyn wirtualnych:

$$\mathbf{c}_{CC} = \{\mathbf{MW}_1, \mathbf{MW}_2, \dots, \mathbf{MW}_m\}. \quad (1.62)$$

Krok 5: Wyznaczenie prawdopodobieństwa blokady \mathbf{E} w całym systemie chmurowym:

$$\mathbf{E} = \{E_1, E_2, \dots, E_m\}, \quad (1.63)$$

$$\bigwedge_{i=\{1,2,\dots,m\}} E_i = 1 - [(1 - E_{i,P}^*)(1 - E_{i,R}^*)(1 - E_{i,D}^*)(1 - E_{i,B}^*)], \quad (1.64)$$

$$\bigwedge_{i=\{1,2,\dots,m\}} \bigwedge_{X=\{P,R,D,B\}} E_{i,X}^* = \rho_{i,X} e_{i,X}^*. \quad (1.65)$$

Zaproponowana przez doktoranta metoda obliczeniowa rozwiązuje problem obsługi zgłoszenia, które wymaga jednoczesności wolnych zasobów w każdym typie zasobów pojedynczej maszyny fizycznej. W metodzie tej doktorant – w wyróżnionych krokach – zaproponował wykonanie następujących operacji.

W kroku 1 na podstawie modelu FAS wyznaczane jest prawdopodobieństwo blokady w każdym z typów zasobów (P,R,D,B) pojedynczej MF. Towarzyszy temu założenie, że w wyniku działania algorytmów wyrównywania obciążenia, ruch oferowany pojedynczej MF \mathbf{a} jest k razy mniejszy niż ruch oferowany całemu systemowi.

W kroku 2 prawdopodobieństwo blokady E_X w każdej grupie zasobów (P,R,D,B) jest określane na podstawie modelu LAS. Zakłada się, że LAS składa się z k oddzielnych zasobów o pojemnościach C_X , które odpowiadają pojemnościom podsystemów pojedynczej MF.

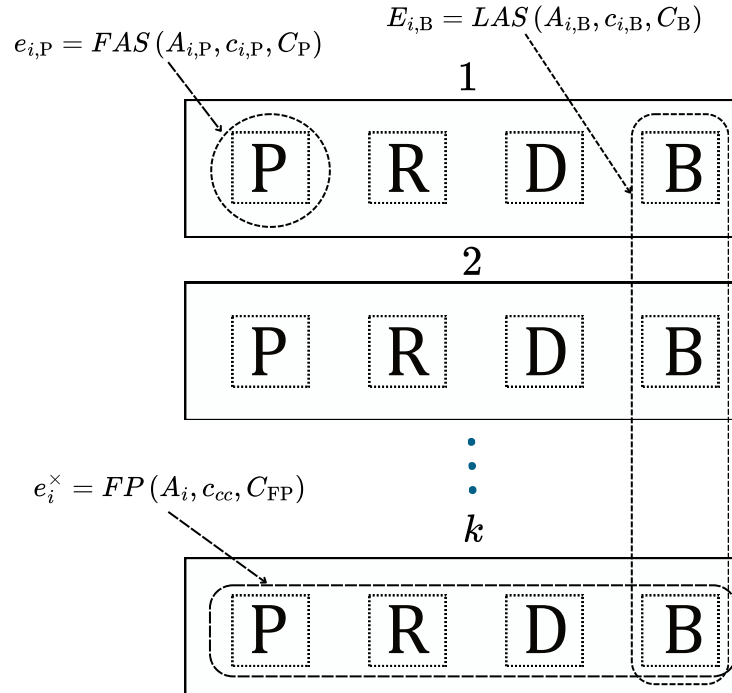
W kroku 3 określane są relacje ρ_X między prawdopodobieństwem blokady e_X w pojedynczych zasobach (zdekomponowanej MF) a prawdopodobieństwem blokady E_X w grupie zdekomponowanych MF. Bezpośrednie zastosowanie modelu LAS do wyznaczenia prawdopodobieństw blokady w całym wieloparametrycznym systemie chmurowym nie jest możliwe, gdyż – zgodnie z przyjętą definicją maszyny wirtualnej – jest ona opisana czterema parametrami, z których wszystkie muszą być dostępne w jednej MF jednocześnie. W związku z tym należy określić relację ρ dla każdego z komponentów systemu (typów zasobów). Relacja ta umożliwi uwzględnienie struktury składającej się z wielu maszyn fizycznych w obliczeniach dla pojedynczej maszyny fizycznej.

Relacja ρ jest kluczowym elementem zaproponowanego przez doktoranta modelu Cloud I, ponieważ zakłada się, że taka sama relacja będzie zachodziła między całkowitym prawdopodobieństwem blokady pojedynczej maszyny fizycznej a całkowitym prawdopodobieństwem blokady grupy k maszyn fizycznych, w których wymagane są wszystkie cztery parametry jednocześnie (P,R,D,B).

W kroku 4 wyznaczaniu podlega prawdopodobieństwo blokady e_X^* w pojedynczej maszynie fizycznej z uwzględnieniem wszystkich typów zasobów jednocześnie, które odpowiada całkowitemu prawdopodobieństwu blokady e_i^* MF wyznaczonemu na podstawie metodologii ustalonego punktu. Metoda ta pozwala określić prawdopodobieństwo blokady w pojedynczej maszynie fizycznej przy założeniu, że zgłoszenie w niej odrzucone nie jest oferowane innym maszynom fizycznym.

W kroku 5 całkowite prawdopodobieństwo blokady w całym systemie chmurowym składającym się z k maszyn fizycznych opisanych jednocześnie czterema różnymi parametrami jest określone na podstawie wzoru (1.64). Możliwość zastosowania wzoru (1.64) wynika z przyjętego założenia, że rzeczywiste prawdopodobieństwa blokady E_X^* w każdej grupie zasobów (P, R, D, B), które spełniają warunek jednoczesnej dostępności dla danego zgłoszenia w pojedynczej MF, podlegają relacji ρ do rzeczywistego prawdopodobieństwa e_X^* w pojedynczym zasobie, który spełnia warunek jednoczesnej dostępności w tej MF wyznaczonego przy pomocy metody ustalonego punktu.

Na rysunku 1.13 przedstawiono ideę całego modelu.



Rysunek 1.13: Idea metody Cloud 1 [J4].

W celu walidacji oraz weryfikacji działania zaproponowanego wieloparametrowego modelu analitycznego, doktorant wykorzystał autorski symulator opisany w p. 1.6.5. W modelu tym, wyróżnione zostało k serwerów o pełnej dostępności, gdzie każdy serwerów składał się z 4 zasobów (RAM, Procesor, Dysk, Łącze). Założono również konieczność utworzenia maszyny wirtualnej na pojedynczej maszynie fizycznej oraz algorytm dystrybucji zgłoszeń do maszyn fizycznych dążący do równomiernego ich obciążenia. Uzyskane wyniki zostały przedstawione w funkcji ruchu oferowanego pojedynczej JA

dostępnej w podsystemie odpowiadającym pamięci RAM

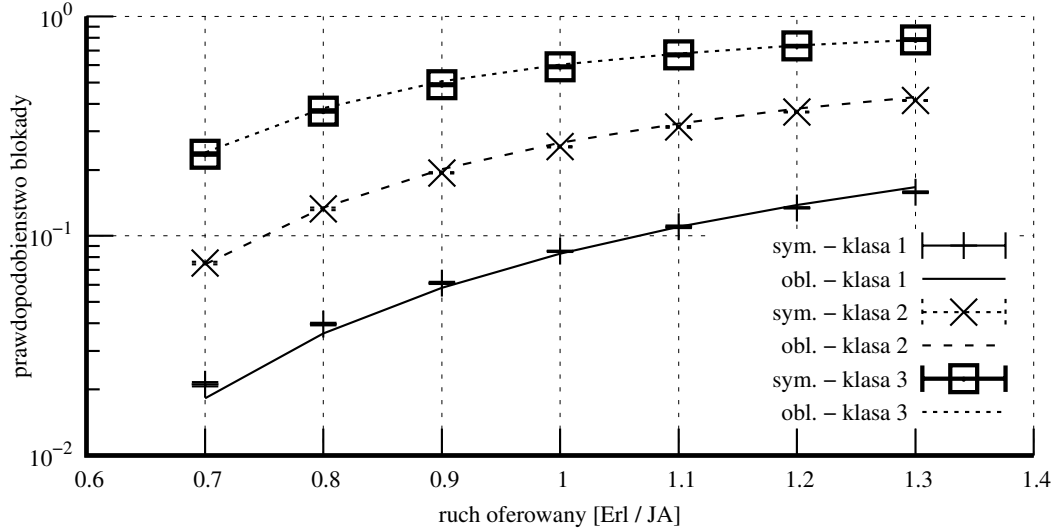
$$a = \frac{\sum_{i=1}^m a_i c_{i,R}}{C_R}. \quad (1.66)$$

Kolejnym założeniem poczynionym przez doktoranta było to, że całkowity oferowany ruch został podzielony między poszczególne klasy zgłoszeń w następujących proporcjach: $a_1 c_{1,R} : a_2 c_{2,R} : \dots : a_m c_{m,R} = 1 : 1 : \dots : 1$.

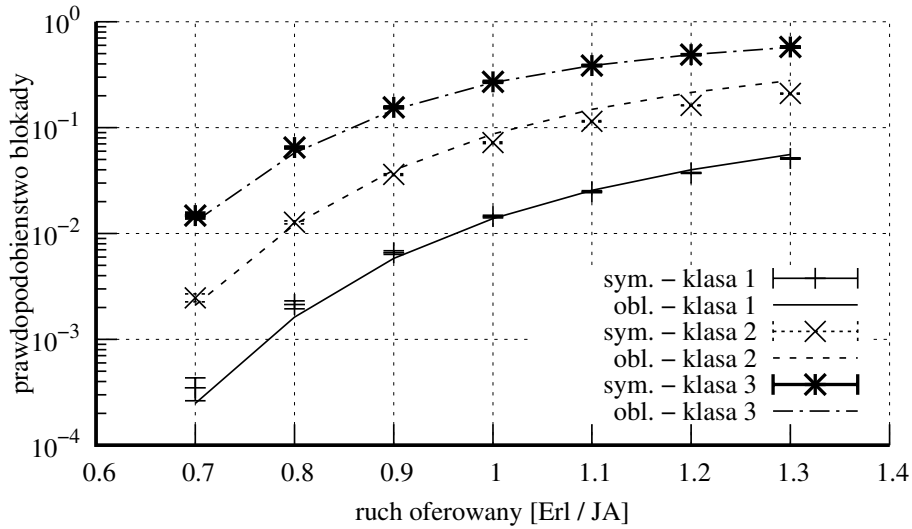
Parametry poszczególnych przeanalizowanych systemów zostały przedstawione w Tabeli 1.8, natomiast uzyskane wartości prawdopodobieństwa blokady na rysunkach 1.14–1.16.

Tabela 1.8: Parametry rozważanych systemów [J4]

System 5				
Liczba MF	Pojemność MF [JA]			
$k = 3$	$C_R = 32$	$C_P = 24$	$C_D = 28$	$C_B = 40$
Rozmiar MW [JA]				
MW klasy 1	$c_{1,R} = 1$	$c_{1,P} = 1$	$c_{1,D} = 1$	$c_{1,B} = 2$
MW klasy 2	$c_{2,R} = 2$	$c_{2,P} = 2$	$c_{2,D} = 2$	$c_{2,B} = 2$
MW klasy 3	$c_{3,R} = 4$	$c_{3,P} = 4$	$c_{3,D} = 4$	$c_{3,B} = 2$
System 6				
Liczba MF	Pojemność MF [JA]			
$k = 5$	$C_R = 24$	$C_P = 24$	$C_D = 32$	$C_B = 28$
Rozmiary MW [JA]				
MW klasy 1	$c_{1,R} = 1$	$c_{1,P} = 1$	$c_{1,D} = 1$	$c_{1,B} = 1$
MW klasy 2	$c_{2,R} = 2$	$c_{2,P} = 2$	$c_{2,D} = 3$	$c_{2,B} = 1$
MW klasy 3	$c_{3,R} = 4$	$c_{3,P} = 3$	$c_{3,D} = 3$	$c_{3,B} = 2$
System 7				
Ruch oferowany traffic	Pojemność MF [JA]			
$a = 1$	$C_R = 16$	$C_P = 16$	$C_D = 14$	$C_B = 12$
Rozmiary MW [JA]				
MW klasy 1	$c_{1,R} = 1$	$c_{1,P} = 1$	$c_{1,D} = 1$	$c_{1,B} = 2$
MW klasy 2	$c_{2,R} = 2$	$c_{2,P} = 1$	$c_{2,D} = 1$	$c_{2,B} = 1$
MW klasy 3	$c_{3,R} = 3$	$c_{3,P} = 2$	$c_{3,D} = 1$	$c_{3,B} = 2$



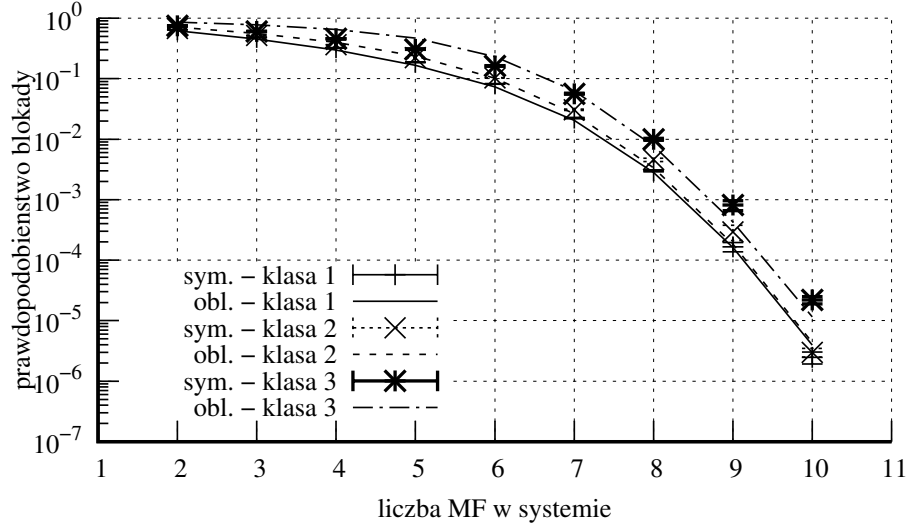
Rysunek 1.14: Prawdopodobieństwo blokady w systemie 5 [J4]



Rysunek 1.15: Prawdopodobieństwo blokady w systemie 6 [J4]

1.6.3.3 Model Cloud II

Analogiczne podejście do opisanego w [C3] doktorant zastosował w modelu Cloud II opublikowanym w [J2], umożliwiającym wyznaczenie średniego wykorzystania poszczególnych zasobów (P,R,D,B) wieloparametrycznego systemu chmurowego. W modelu tym wykorzystano modele LAS i FAS oraz metodę ustalonego punktu, tym razem w celu określenia wykorzystania zasobów w grupie k maszyn fizycznych kontrolowanych przez GM. Model Cloud II w początkowej fazie pozwala na wyznaczenie rozkładu zajętości w indywidualnych (zdekomponowanych) zasobach maszyny fizycznej (przy pomocy modelu FAS) i grupy maszyn fizycznych (przy pomocy modelu LAS). Uzyskane w ten sposób rozkłady w kolejnych krokach posłużyły doktorantowi do określenia średniego wykorzystania każdego zasobu niezależnie oraz do określenia współczynni-



Rysunek 1.16: Prawdopodobieństwo blokady w systemie 7 [J4]

ków zależności dla każdego z typów zasobów (P,R,D,B). Jednocześnie metoda ustalonego punktu umożliwiła doktorantowi określenie rozkładu zajętości dla pojedynczej maszyny fizycznej, uwzględniając jednocześnie zapotrzebowanie na wszystkie rodzaje zasobów (P,R,D,B). Rozkład ten został wyznaczony przy założeniu o równomiernym obciążeniu MF. W związku z tym ruch pojawiający się na wejściu MF jest k -tą częścią całego ruchu pojawiającego się na wejściu do grupy k maszyn fizycznych.

Dalsze obliczenia umożliwiają uwzględnienie wcześniej wyznaczonych współczynników korelacji średniego wykorzystania poszczególnych (zdekomponowanych) zasobów z wykorzystaniem modeli LAS i FAS do określenia wykorzystania zasobów w MF znajdującej się w systemie chmurowym. Wyznaczone współczynniki korelacji uwzględniają występowanie w systemie k MF i pozwalają na odwzorowanie wartości średniego wykorzystania poszczególnych typów zasobów, uzyskanych metodą ustalonego punktu dla pojedynczej maszyny wirtualnej, na wartości średniego wykorzystania tych zasobów w systemie składającym się z k MF.

Model ten można opisać w postaci następujących kroków obliczeniowych:

Krok 1: Określenie rozkładów $\mathbf{P}_{MF,X}$ w pojedynczej MF dla każdego typu zasobu ($X \in \{P, R, D, B\}$) przy założeniu, że zasoby C_X obsługują maszyny wirtualne z żadaniami $c_{i,X}$ ($1 \leq i \leq m$), niezależnie od obsługi tych maszyn wirtualnych w innych zasobach \mathbf{C}_{MF} :

$$\mathbf{P}_{MF,X} = FAS(\mathbf{A}_{MF,X}, \mathbf{c}_{MF,X}, C_X) = \{[P_{FAS}(n)]_{C_X}, 0 \leq n \leq C_X\}. \quad (1.67)$$

Krok 2: Określenie średniego wykorzystania poszczególnych zasobów $L_{MF,X}$ dla każdego typu zasobu MF ($X \in \{P, R, D, B\}$) niezależnie:

$$L_{MF,X} = \sum_{n=0}^{C_X} n [P_{FAS}(n)]_{C_X}. \quad (1.68)$$

Krok 3: Określenie rozkładów zajętości $P_{kMF,X}$ w grupie k MF tworzących systemy LAS dla każdego typu zasobu ($X \in \{P, R, D, B\}$) niezależnie (po dekompozycji), zakładając, że systemowi oferowany jest ruch o natężeniu k razy większym niż ruch oferowany pojedynczej MF:

$$\mathbf{A}_{kMF,X} = k\mathbf{A}_{MF,X} = \{kA_{i,X}, 1 \leq i \leq m\}, \quad (1.69)$$

$$\mathbf{P}_{kMF,X} = LAS(\mathbf{A}_{kMF,X}, \mathbf{c}_{MF,x}, kC_X) = \{[P_{LAS}(n)]_{kC_X}, 0 \leq n \leq kC_X\}. \quad (1.70)$$

Krok 4: Określenie średniego wykorzystania zasobów $L_{kMF,X}$ dla każdego z zdekomponowanych zasobów ($X \in \{P, R, D, B\}$) w grupie k MF:

$$\mathbf{L}_{kMF,X} = \sum_{n=0}^{kC_X} n[P_{LAS}(n)]_{kC_X}. \quad (1.71)$$

Krok 5: Obliczenie, dla każdego typu zasobu X , współczynnika wykorzystania zasobu ϑ_X stanowiącego relację pomiędzy wykorzystaniem zdekomponowanych zasobów wyznaczonych przy pomocy modelu FAS dla pojedynczej MF (krok 2) i modelu LAS dla grupy k MF (krok 4):

$$\vartheta_X = \frac{L_{kMF,X}}{kL_{MF,X}}. \quad (1.72)$$

Krok 6: Określenie z wykorzystaniem metody ustalonego punktu rozkładów $\mathbf{P}_{MF,X}^*$ w pojedynczej MF dla każdego typu zasobu ($X \in \{P, R, D, B\}$), przy założeniu, że obsługa maszyn wirtualnych wymaga jednoczesnej alokacji żądań \mathbf{c}_i w każdym zasobie należącym do zbioru zasobów \mathbf{C}_{MF} :

$$\mathbf{P}_{MF,X}^* = FP(\mathbf{A}_{MF,X}, \mathbf{c}_{MF,X}, C_X). \quad (1.73)$$

Krok 7: Wyznaczenie średniego wykorzystania zasobów $L_{MF,X}^*$ w pojedynczej MF dla każdego typu zasobu X ($X \in \{P, R, D, B\}$) wykorzystując w obliczeniach ostatni uzyskany przy pomocy metody ustalonego punktu rozkład zajętości wyznaczony w kroku 6:

$$L_{MF,X}^* = \sum_{n=0}^{C_X} n[P_{FP}^*(n)]C_X, \quad (1.74)$$

gdzie:

$$[P_{FP}^*(n)]C_X \in \mathbf{P}_{MF,X}^*. \quad (1.75)$$

Krok 8: Wyznaczanie średniego wykorzystania zasobów $L_{kMF,X}^*$ w grupie k MF dla każdego typu zasobu X ($X \in \{P, R, D, B\}$), uwzględniając występowanie w systemie grupy k maszyn fizycznych poprzez uwzględnienie wyznaczonej w kroku 5 relacji:

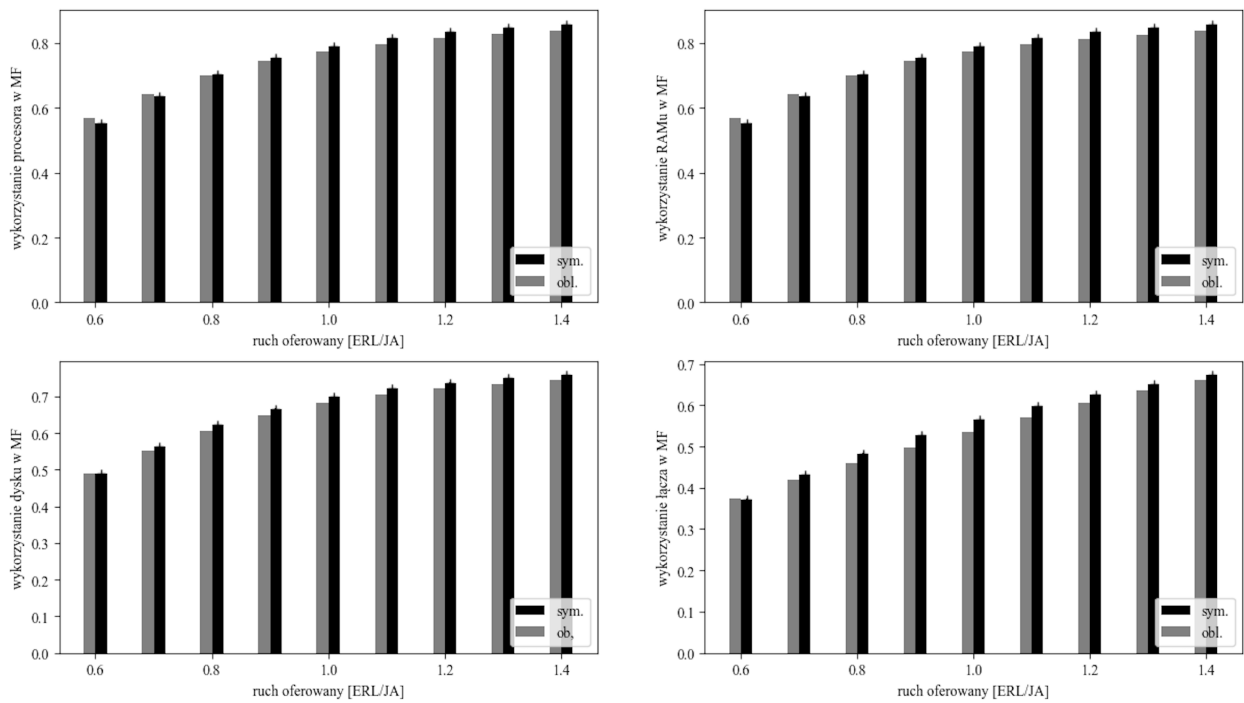
$$L_{kMF,X}^* = k\vartheta_X L_{MF,X}^*. \quad (1.76)$$

Aby określić rzeczywiste średnie wykorzystanie zasobów $L_{kMF,X}^*$ w grupie k MF, doktorant zastosował współczynnik wykorzystania zasobów ϑ_X , obliczony przy założeniu, że obsługa ruchu w określonych zasobach C_X danej MF jest niezależna od obsługi ruchu w pozostałych zasobach należących do zbioru \mathbf{C}_{MF} . Oznacza to, że obliczenia parametru ϑ_X bazują na ruchu $\mathbf{A}_{MF,X}$ ($\mathbf{A}_{MF,X} = \{[A_{i,X}, 1 \leq i \leq m]\}$), bezpośrednio oferowanym pojedynczej MF. Jednocześnie, obliczenia rzeczywistego średniego wykorzystania zasobów $L_{MF,X}^*$ w pojedynczej MF wynikają z zastosowania metody ustalonego punktu, a zatem opierają się na efektywnym ruchu $\mathbf{A}_{MF,X}^{(z)}$, wyznaczonym na podstawie równania (1.31) w ostatniej iteracji z metody ustalonego punktu (punkt 1.5.2.3). W związku z tym, aby określić rzeczywiste średnie wykorzystanie zasobów $L_{kMF,X}^*$ w grupie k MF, w metodzie Cloud II zastosowany został współczynnik wykorzystania zasobów ϑ_X (równanie (1.76)), wyznaczony wcześniej dla ruchu $\mathbf{A}_{MF,X}$.

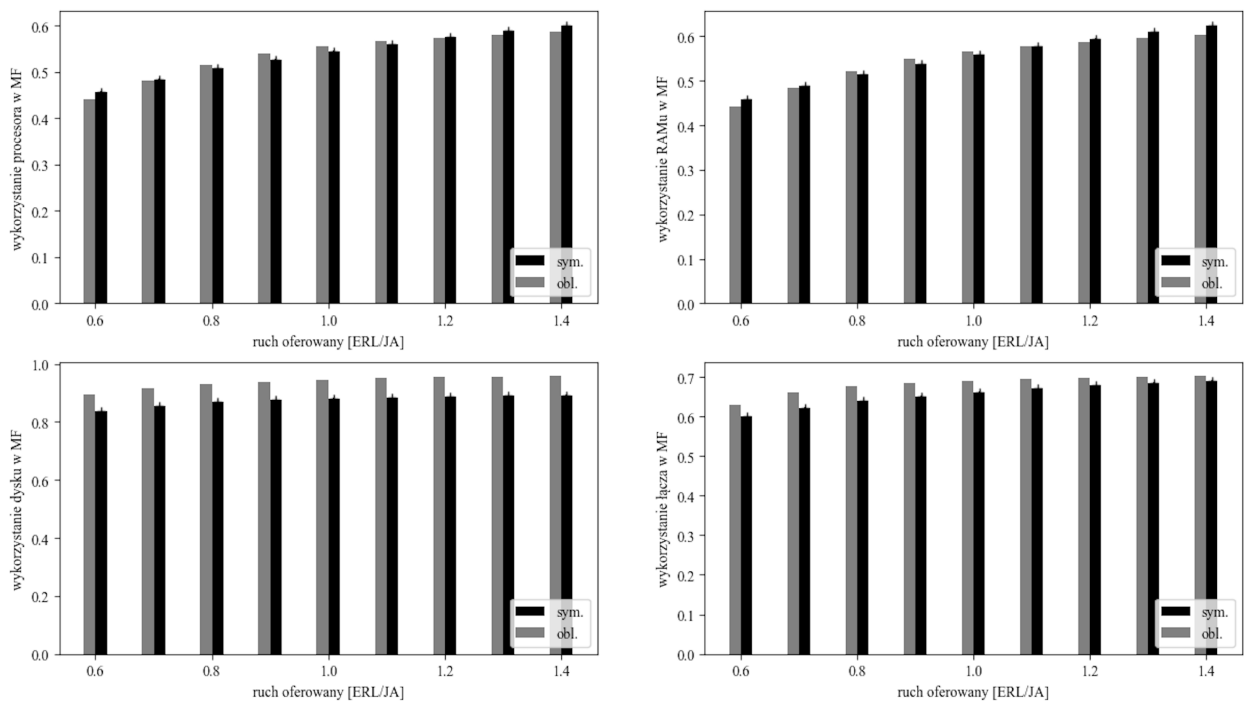
Wyniki uzyskane przy pomocy opracowanego modelu analitycznego (Cloud II) zostały porównane z wynikami uzyskanymi przy pomocy opracowanego przez doktoranta symulatora. Założenia związane z procesem symulacji były analogiczne do założeń opisanych w p. 1.6.3.2 rozprawy doktorskiej. W tabeli 1.9 zebrane zostały parametry analizowanych systemów, natomiast uzyskane wyniki zostały przedstawione na wykresach 1.17 i 1.18.

Tabela 1.9: Parametry analizowanych systemów [J2]

System 8				
Liczba MF	Pojemność MF [JA]			
$k = 3$	$C_R = 23$	$C_P = 23$	$C_D = 26$	$C_B = 20$
Rozmiary MW [JA]				
MW klasy 1	$c_{1,R} = 1$	$c_{1,P} = 1$	$c_{1,D} = 1$	$c_{1,B} = 1$
MW klasy 2	$c_{2,R} = 2$	$c_{2,P} = 2$	$c_{2,D} = 2$	$c_{2,B} = 1$
MW klasy 3	$c_{3,R} = 4$	$c_{3,P} = 4$	$c_{3,D} = 4$	$c_{3,B} = 1$
System 9				
Liczba MF	Pojemność MF [JA]			
$k = 5$	$C_R = 38$	$C_P = 35$	$C_D = 27$	$C_B = 25$
Rozmiary MW [JA]				
MW klasy 1	$c_{1,R} = 1$	$c_{1,P} = 1$	$c_{1,D} = 3$	$c_{1,B} = 1$
MW klasy 2	$c_{2,R} = 3$	$c_{2,P} = 3$	$c_{2,D} = 3$	$c_{2,B} = 4$
MW klasy 3	$c_{3,R} = 5$	$c_{3,P} = 4$	$c_{3,D} = 1$	$c_{3,B} = 2$



Rysunek 1.17: Średnie wykorzystanie zasobów (P, R, D, B) w pojedynczej maszynie fizycznej w Systemie 8 [J2]



Rysunek 1.18: Średnie wykorzystanie zasobów (P, R, D, B) w pojedynczej maszynie fizycznej w Systemie 9 [J2]

1.6.4 Model jednoparametrycznego kolejkowego systemu teleinformatycznego z ruchem Erlanga, Engseta i Pascala (BPP) typu elastycznego i adaptacyjnego.

W celu rozbudowy opracowywanych i opublikowanych w [J1, J2, J4, J5, C1, C2, C3] modeli, doktorant podjął pracę nad modelowaniem systemów z uwzględnieniem nie tylko ruchu Erlanga, ale również ruchu Engseta i Pascala. Ponadto w rozważaniach uwzględniony został typ ruchu elastycznego i adaptacyjnego.

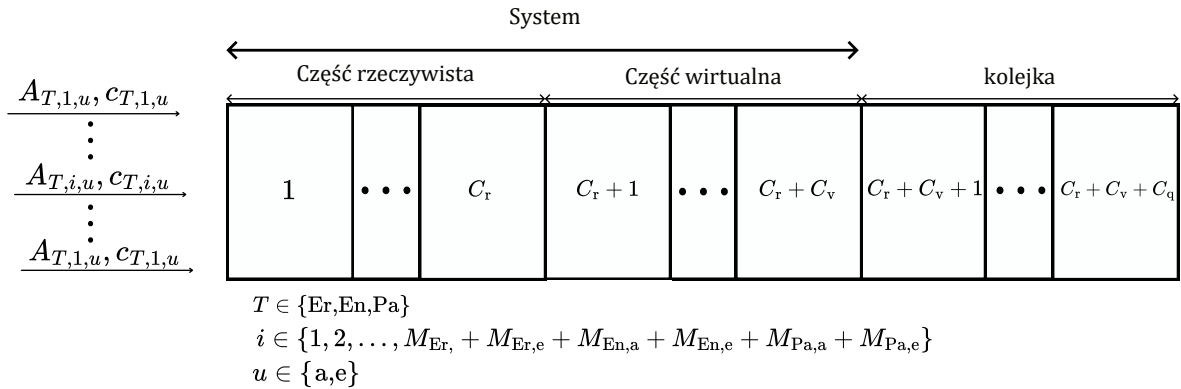
1.6.4.1 Opis i struktura systemu

W analizowanym przez doktoranta jednoparametrycznym kolejkowym systemie teleinformatycznym (nazywanym dalej systemem), wyróżniony został ruch Erlanga, Engseta oraz Pascala. Strumień zgłoszeń każdej klasy ruchu Erlanga jest strumieniem Poissona, którego intensywność nie zależy od stanu zajętości systemu (tzn. od liczby obsługiwanych zgłoszeń). Strumień zgłoszeń każdej klasy ruchu Engseta to strumień Bernoulliego, którego intensywność maleje wraz ze wzrostem stanu zajętości systemu. Z kolei w przypadku klas ruchu Pascala strumień zgłoszeń jest strumieniem Pascala, którego intensywność rośnie wraz ze wzrostem stanu zajętości systemu. Taka mieszanina strumieni ruchu w literaturze przedmiotu nazywana jest ruchem BPP [50, 51]. Ponadto, we wcześniejszych rozważaniach doktorant uwzględniał ruch strumieniowy, który jest typem ruchu niezmiennym swoich parametrów bez względu na stan zajętości systemu. W badaniach opublikowanych w artykule [J3], doktorant uwzględnił zamiast ruchu strumieniowego, dwa inne typy ruchu, a mianowicie ruch elastyczny i adaptacyjny. Co za tym idzie, każdy z oferowanych systemowi strumieni ruchu mógł być ruchem typu elastycznego lub adaptacyjnego, co oznacza, że proces obsługi zgłoszeń stał się również zależny od stanu. Zależność tą można opisać w następujący sposób: brak wolnych zasobów do obsługi nowego zgłoszenia implikuje kompresję (zmniejszenie rozmiaru) wszystkich aktualnie obsługiwanych zgłoszeń. Jeżeli przyjęcie nowego zgłoszenia spowodowałoby przekroczenie pewnego poziomu kompresji (nazwanego granicą kompresji), to zgłoszenie takie przenoszone jest do kolejki w nieskompresowanej formie. Brak miejsca w kolejce, rozumiany jako brak wolnych zasobów w liczbie większej bądź równej rozmiarowi nieskompresowanego zgłoszenia, powoduje jego odrzucenie. Jeżeli dokonanie kompresji łączy się z wydłużeniem czasu obsługi potrzebnym do wykonania określonego zadania, to taki ruch określany jest jako ruch elastyczny. Przykładami aplikacji wiążących się z tym typem ruchu mogą być wszystkie operacje polegające na pobieraniu danych, przesyłaniu e-maili, wykonywaniu obliczeń z wykorzystaniem procesora, itp. Natomiast, jeżeli wraz z dokonaniem kompresji zmianie nie podlega czas obsługi zgłoszenia, to ruch określany jest jako ruch adaptacyjny. Ruch ten jest typowym ruchem dla usług czasu rzeczywistego, w którym w celu transmisji danych zmieniana jest jakość, na przykład zmiana kodowania sygnału wideo z formatu 4K na HD. Systemy z ruchem elastycznym i adaptacyjnym, określane są jako systemy z kompresją bezprogową. W takich systemach do kompresji dochodzi w momencie, w którym wszystkie zasoby systemu są zajęte.

W analizowanym systemie oprócz pojemności rzeczywistej systemu C_r , w celu określenia granicy kompresji, doktorant wprowadził pojęcie pojemności wirtualnej systemu,

oznaczonej parametrem C_v ($C_v > C_r$). Wprowadzenie parametru C_v pozwala na przyjęcie następującej interpretacji mechanizmu kompresji: zgłoszenia mogą podlegać mechanizmowi kompresji dopóki liczba zajętych JA w maszynie fizycznej, określona jako suma nieskompresowanych żądań zgłoszeń wszystkich klas, nie przekroczy pojemności wirtualnej systemu. Oznacza to, że wybór pojemności wirtualnej systemu C_v jest równoznaczny ze wskaźnikiem maksymalnego poziomu kompresji w serwerze, który definiuje się jako stosunek rzeczywistej pojemności systemu do pojemności wirtualnej $\frac{C_r}{C_v}$.

Na rysunku 1.19 przedstawiono poglądowy schemat wielosługowego systemu z ruchem BPP, kompresją i kolejką.



Rysunek 1.19: Wielosługowy system kolejkowy z ruchem BPP typu elastycznego i adaptacyjnego [J3]

1.6.4.1.1 Model I

W procesie opracowywania modelu analitycznego, doktorant przyjął następujące formalne oznaczenia strumieni zgłoszeń:

- ruch Erlanga:
 - $M_{\text{Er},u}$ – liczba klas zgłoszeń Erlanga typu kompresji u ,
 - $\lambda_{\text{Er},i,u}$ – intensywność strumienia Poissona klasy i typu kompresji u ,
 - $\mu_{\text{Er},i,u}$ – intensywność strumienia obsługi zgłoszeń klasy i typu kompresji u (podlegającego rozkładowi wykładniczemu),
 - $c_{\text{Er},i,u}$ – liczba zasobów niezbędnych do zestawienia połączenia zgłoszeń klasy i typu kompresji u (wyrażona w JA),
- ruch Engseta:
 - $M_{\text{En},u}$ – liczba klas zgłoszeń Engseta typu kompresji u oferowana systemowi,
 - $N_{\text{En},i,u}$ – liczba źródeł ruchu klasy i typu kompresji u ,
 - $\Lambda_{\text{En},i,u}$ – intensywność strumienia Bernoullego klasy i typu kompresji u generowanego przez pojedyncze wolne źródło ruchu,

- $\lambda_{\text{En},i,u}(x_{\text{En},i,u}^*)$ – intensywność napływu zgłoszeń klasy Engseta i typu kompresji u , parametr ten stanowi funkcję liczby obsługiwanych zgłoszeń i wyraża się wzorem:

$$\lambda_{\text{En},j,u}(x_{\text{En},j,u}^*) = [N_{\text{En},j,u} - x_{\text{En},j,u}^*] \Lambda_{\text{En},j,u}, \quad (1.77)$$

- $\mu_{\text{En},i,u}$ – intensywność strumienia obsługi zgłoszeń klasy i typu kompresji u (podlegającego rozkładowi wykładniczemu),
- $c_{\text{En},i,u}$ – liczba zasobów niezbędnych do zestawienia połączenia zgłoszeń klasy i typu kompresji u (wyrażona w JA),

• ruch Pascala:

- $M_{\text{Pa},u}$ – liczba klas zgłoszeń Pascala typu kompresji u oferowana systemowi,
- $S_{\text{Pa},i,u}$ – liczba źródeł ruchu klasy i typu kompresji u ,
- $\Lambda_{\text{Pa},i,u}$ – intensywność strumienia Pascala klasy i typu kompresji u generowanego przez pojedyncze wolne źródło ruchu,
- $\lambda_{\text{Pa},i,u}(x_{\text{Pa},i,u}^*)$ – intensywność napływu zgłoszeń Pascala klasy i typu kompresji u , parametr ten stanowi funkcję liczby obsługiwanych zgłoszeń i wyraża się wzorem:

$$\lambda_{\text{Pa},k,u}(x_{\text{Pa},k,u}^*) = [S_{\text{Pa},k,u} + x_{\text{Pa},k,u}^*] \Lambda_{\text{Pa},k,u}, \quad (1.78)$$

- $\mu_{\text{Pa},i,u}$ – intensywność strumienia obsługi zgłoszeń klasy i typu kompresji u (podlegającego rozkładowi wykładniczemu),
- $c_{\text{Pa},i,u}$ – liczba zasobów niezbędnych do zestawienia połączenia zgłoszeń klasy i typu kompresji u (wyrażona w JA).

Kompresja klas ruchu elastycznego powoduje jednoczesne wydłużenie czasu obsługi zgłoszeń i zmniejszenie liczby żądanych zasobów. Zatem intensywność obsługi $\mu_{d,i,u}$ oraz liczba żądanych zasobów $c_{d,i,u}$ mogą być zapisane następująco:

$$\mu_{d,i,e}(X) = \begin{cases} \mu_{d,i,e}, & \text{dla } 0 \leq n_X \leq C_r, \\ \frac{C_r}{n_X} \mu_{d,i,e}, & \text{dla } C_r < n_X \leq C_r + C_v, \\ \frac{C_r}{C_v} \mu_{d,i,e}, & \text{dla } C_r + C_v < n_X \leq C_r + C_v + C_q, \end{cases} \quad (1.79)$$

oraz

$$c_{d,i,e}(X) = \begin{cases} c_{d,i,e}, & \text{dla } 0 \leq n_X \leq C_r, \\ \frac{C_r}{n_X} c_{d,i,e}, & \text{dla } C_r < n_X \leq C_r + C_v, \\ \frac{C_r}{C_v} c_{d,i,e}, & \text{dla } C_r + C_v < n_X \leq C_r + C_v + C_q, \end{cases} \quad (1.80)$$

gdzie $d \in \{\text{Er}, \text{En}, \text{Pa}\}$ i $u \in \{e, a\}$.

W przypadku klas ruchu adaptacyjnego, kompresja zgłoszeń powoduje wyłącznie zmniejszenie liczby żądanych JA, możemy zatem zapisać:

$$\mu_{d,i,a}(X) = \mu_{d,i,a}, \quad \text{for } 0 \leq n_X \leq C_r + C_v + C_q, \quad (1.81)$$

oraz

$$c_{d,i,a}(X) = \begin{cases} c_{d,i,a}, & \text{dla } 0 \leq n_X \leq C_r, \\ \frac{C_r}{n_X} c_{d,i,a}, & \text{dla } C_r < n_X \leq C_r + C_v, \\ \frac{C_r}{C_v} c_{d,i,a}, & \text{dla } C_r + C_v < n_X \leq C_r + C_v + C_q, \end{cases} \quad (1.82)$$

gdzie $d \in \{\text{Er,En,Pa}\}$ i $u \in \{e,a\}$.

Współczynnik $\frac{C_r}{C_v}$ w obszarze kolejki (w mikrostanach dla których spełniona jest nierówność $C_v < n_X \leq C_q$) nie oznacza, że zgłoszenia przyjmowane do kolejki podlegają mechanizmowi kompresji. Oznacza jedynie, że przyjęcie zgłoszenia z obszaru kolejki do obszaru obsługi implikuje maksymalną kompresję wszystkich obsługiwanych zgłoszeń typu elastycznego i adaptacyjnego. Podobna interpretacja odnosi się do czasu obsługi zgłoszeń klas typu elastycznego. Czas obsługi takiego zgłoszenia wydłuża się maksymalnie, a zatem intensywność jego obsługi maleje $\frac{C_r}{C_v}$ razy.

Na początkowym etapie badań, doktorant rozważał system na poziomie mikrostanów. Założył jednocześnie, że w rozważanym systemie wielousługowym z ruchem BPP podlegającym kompresji, proces obsługi jest procesem odwracalnym. Założenie to implikuje możliwość zastosowania równań lokalnej równowagi pomiędzy sąsiadującymi mikrostanami. Na rysunku 1.20 przedstawiony został fragment wielowymiarowego procesu Markowa zachodzącego w rozważanym systemie.

W takim przypadku równanie lokalnej równowagi dla zgłoszeń kolejno Erlanga, Engseta oraz Pascala, można zapisać w następującej postaci:

$$A_{\text{Er},i,u}(X) c_{\text{Er},i,u}(X) P(X - 1_{\text{Er},i,u}) = c_{\text{Er},i,u}(X) x_{\text{Er},i,u}^*(X) P(X), \quad (1.83)$$

$$\begin{aligned} \alpha_{\text{En},j,u}(X) [N_{\text{En},j,u} - x_{\text{En},j,u}^*(X - 1_{\text{En},j,u})] c_{\text{En},j,u}(X) P(X - 1_{\text{En},j,u}) = \\ = c_{\text{En},j,u}(X) x_{\text{En},j,u}^*(X) P(X), \end{aligned} \quad (1.84)$$

$$\begin{aligned} \beta_{\text{Pa},k,u}(X) [S_{\text{Pa},k,u} + x_{\text{Pa},k,u}^*(X - 1_{\text{Pa},k,u})] c_{\text{Pa},k,u}(X) P(X - 1_{\text{Pa},k,u}) = \\ = c_{\text{Pa},k,u}(X) x_{\text{Pa},k,u}^*(X) P(X), \end{aligned} \quad (1.85)$$

gdzie:

$$\begin{aligned} A_{\text{Er},i,u}(X) &= \frac{\lambda_{\text{Er},i,u}}{\mu_{\text{Er},i,u}(X)} = \\ &= \begin{cases} A_{\text{Er},i,u}, & \text{dla } (u = a \wedge 0 \leq n_X \leq C_r + C_v + C_q) \wedge (u = e \wedge 0 \leq n_X \leq C_r), \\ \frac{n_X}{C_r} A_{\text{Er},i,u}, & \text{dla } u = e \wedge C_r < n_X \leq C_r + C_v, \\ \frac{C_v}{C_r} A_{\text{Er},i,u}, & \text{dla } u = e \wedge C_r + C_v < n_X \leq C_r + C_v + C_q \end{cases} \end{aligned} \quad (1.86)$$

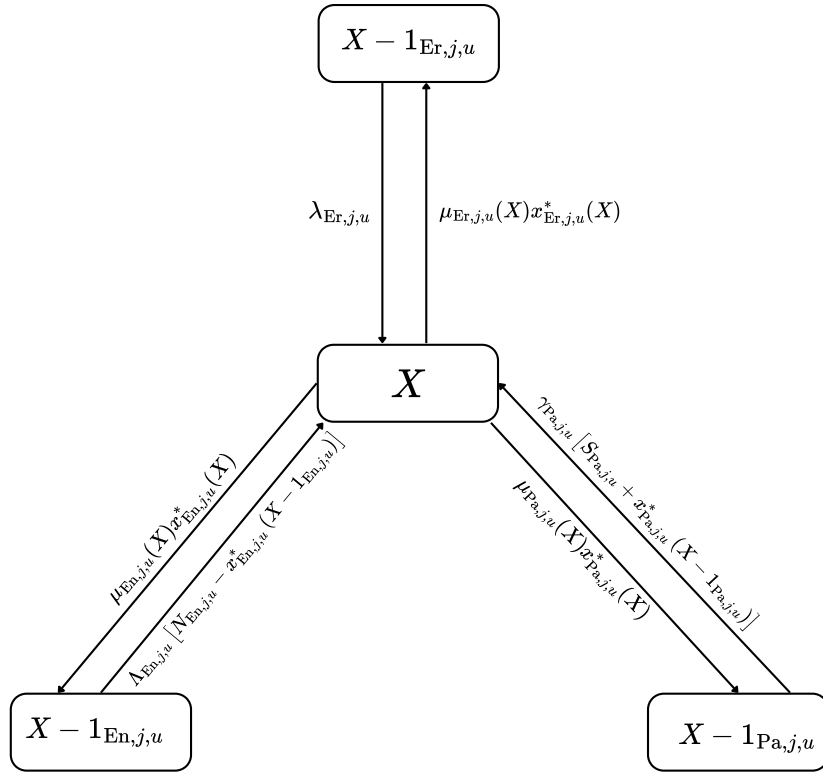
określa natężenie ruchu oferowanego Erlanga klasy i typu u , natomiast

$$\alpha_{\text{En},j,u}(X) = \frac{\Lambda_{\text{En},j,u}}{\mu_{\text{En},j,u}(X)} = \begin{cases} \alpha_{\text{En},j,u}, & \text{dla } (u = a \wedge 0 \leq n_X \leq C_r + C_v + C_q) \wedge (u = e \wedge 0 \leq n_X \leq C_r), \\ \frac{n_X}{C_r} \alpha_{\text{En},j,u}, & \text{dla } u = e \wedge C_r < n_X \leq C_r + C_v, \\ \frac{C_v}{C_r} \alpha_{\text{En},j,u}, & \text{dla } u = e \wedge C_r + C_v < n_X \leq C_r + C_v + C_q \end{cases} \quad (1.87)$$

i

$$\beta_{\text{Pa},k,u}(X) = \frac{\Lambda_{\text{Pa},k,u}}{\mu_{\text{Pa},k,u}(X)} = \begin{cases} \beta_{\text{Pa},k,u}, & \text{dla } (u = a \wedge 0 \leq n_X \leq C_q) \wedge (u = e \wedge 0 \leq n_X \leq C_r), \\ \frac{n_X}{C_r} \beta_{\text{Pa},k,u}, & \text{dla } u = e \wedge C_r < n_X \leq C_v, \\ \frac{C_v}{C_r} \beta_{\text{Pa},k,u}, & \text{dla } u = e \wedge C_v < n_X \leq C_q \end{cases} \quad (1.88)$$

określają natężenie ruchu oferowanego przez pojedyncze wolne źródło ruchu Engseta klasy j typu kompresji u oraz Pascala k typu kompresji u , odpowiednio.



Rysunek 1.20: Fragment wielowymiarowego procesu Markowa w systemie z ruchem elastycznym i adaptacyjnym BPP [J3].

Jednocześnie, w pracy [52] przeprowadzona została analiza procesu obsługi zachodzącego w systemach z ruchem adaptacyjnym i elastycznym, która wykazała, że w systemach tych nie jest spełniona właściwość odwracalności procesu obsługi. W związku z tym, wyrażenie $c_{d,j,u}(X) x_{d,j,u}^*(X)$ zostało zastąpione taką wartością $r_{d,j,u}(X)$ wyrażoną w JA, aby spełniony był warunek odwracalności procesu obsługi. Uwzględniając równania (1.83)-(1.85) można zapisać:

$$r_{Er,i,u}(X) = \frac{A_{Er,i,u}(X) c_{Er,i,u}(X) P(X - 1_{Er,i,u})}{P(X)}, \quad (1.89)$$

$$\begin{aligned} r_{En,j,u}(X) &= \\ &= \frac{\alpha_{En,j,u}(X) [N_{En,j,u} - x_{En,j,u}^*(X - 1_{En,j,u})] c_{En,j,u}(X) P(X - 1_{En,j,u})}{P(X)}, \end{aligned} \quad (1.90)$$

$$\begin{aligned} r_{Pa,k,u}(X) &= \\ &= \frac{\beta_{Pa,k,u}(X) [S_{Pa,k,u} + x_{Pa,k,u}^*(X - 1_{Pa,k,u})] c_{Pa,k,u}(X) P(X - 1_{Pa,k,u})}{P(X)}. \end{aligned} \quad (1.91)$$

Niezależność strumieni zgłoszeń oferowanych systemowi pozwala na zsumowanie wszystkich $M_{Er,a}$ i $M_{Er,e}$ równań typu (1.83) dla elastycznych i adaptacyjnych strumieni Erlanga, $M_{En,a}$ i $M_{En,e}$ równań typu (1.84) dla adaptacyjnych i elastycznych strumieni Engseta, oraz $M_{Pa,a}$ i $M_{Pa,e}$ równań typu (1.85) strumieni elastycznych i adaptacyjnych strumieni Pascala dla mikrostanu X . W związku z tym rozkład zajętości w rozważanym systemie, na poziomie mikrostanów, został zapisany przez doktoranta jako:

$$\begin{aligned} P(X) \min \{n_X, C_r\} &= \sum_{u \in \{e,a\}} \sum_{i=1}^{M_{Er,u}} A_{Er,i,u}(X) c_{Er,i,u}(X) P(X - 1_{Er,i,u}) + \\ &+ \sum_{u \in \{e,a\}} \sum_{j=1}^{M_{En,u}} \alpha_{En,j,u}(X) [N_{En,j,u} - x_{En,j,u}^*(X - 1_{En,j,u})] c_{En,j,u}(X) P(X - 1_{En,j,u}) + \\ &+ \sum_{u \in \{e,a\}} \sum_{k=1}^{M_{Pa,u}} \beta_{Pa,k,u}(X) [S_{Pa,k,u} + x_{Pa,k,u}^*(X - 1_{Pa,k,u})] c_{Pa,k,u}(X) P(X - 1_{Pa,k,u}), \end{aligned} \quad (1.92)$$

dla $0 \leq n_X \leq C_r + C_v + C_q$, w pozostałych przypadkach $P(X) = 0$, a wartość prawdopodobieństwa $P(0)$ wynika z warunku normującego.

Dzięki zdefiniowaniu przez doktoranta rozkładu zajętości systemu na poziomie mikrostanu, możliwym stało się określenie rozkładu zajętości na poziomie makrostanów. Prawdopodobieństwo makrostanu $P(n)$ określa prawdopodobieństwo zajętości n JA i może być wyrażone jako suma prawdopodobieństw odpowiednich mikrostanów:

$$P(n) = \sum_{\Omega(n)} P(X), \quad (1.93)$$

gdzie $\Omega(n)$ jest makrostanem, tj. zbiorem wszystkich mikrostanów X , które spełniają następujący warunek:

$$n_X = \sum_{u \in \{a,e\}} \sum_{i=1}^{M_{Er,u}} x_{Er,i,u}^* c_{Er,i,u} + \sum_{u \in \{a,e\}} \sum_{i=1}^{M_{En,u}} x_{En,i,u}^* c_{En,i,u} + \sum_{u \in \{a,e\}} \sum_{i=1}^{M_{Pa,u}} x_{Pa,i,u}^* c_{Pa,i,u}. \quad (1.94)$$

Przyjęte założenie o odwracalności procesu obsługi na poziomie mikrostanów pozwala na obustronne zsumowanie równania (1.92) po wszystkich mikrostanach należących do makrostanu $\Omega(n)$. Uwzględniając zatem definicję prawdopodobieństwa makrostanu (1.93) oraz wzór (1.92) prawdopodobieństwo zdarzenia, że system znajduje się w stanie zajętości n JA można zapisać następująco:

$$\begin{aligned} \min \{n, C_r\} \sum_{\Omega(n)} P(X) &= \sum_{\Omega(n)} \left(\sum_{u \in \{e,a\}} \sum_{i=1}^{M_{Er,u}} A_{Er,i,u}(X) c_{Er,i,u}(X) P(X - 1_{Er,i,u}) + \right. \\ &+ \sum_{u \in \{e,a\}} \sum_{j=1}^{M_{En,u}} \alpha_{En,j,u}(X) [N_{En,j,u} - x_{En,j,u}^*(X - 1_{En,j,u})] c_{En,j,u}(X) P(X - 1_{En,j,u}) + \\ &\left. + \sum_{u \in \{e,a\}} \sum_{k=1}^{M_{Pa,u}} \beta_{Pa,k,u}(X) [S_{Pa,k,u} + x_{Pa,k,u}^*(X - 1_{Pa,k,u})] c_{Pa,k,u}(X) P(X - 1_{Pa,k,u}) \right), \end{aligned} \quad (1.95)$$

Średnia liczba obsługiwanych zgłoszeń ruchu $d \in \{Er, En, Pa\}$, klasy j typu kompresji $u \in \{e, a\}$ w stanie zajętości n JA może być zdefiniowana w następujący sposób:

$$y_{d,j,u}(n) = \sum_{X \in \Omega(n)} \frac{x_{d,j,u}^*(X) P(X)}{P(n)}. \quad (1.96)$$

Dla uproszczenia zapisu doktorant wprowadził następującą notację dla natężenia ruchu oferowanego przez klasy zgłoszeń Engseta i Pascala typu elastycznego i adaptacyjnego:

$$A_{En,j,u}(n) = \alpha_{En,j,u}(n) [N_{En,j,u} - y_{En,j,u}(n - c_{En,j,u})], \quad (1.97)$$

$$A_{Pa,k,u}(n) = \beta_{Pa,k,u}(n) [S_{Pa,k,u} + y_{Pa,k,u}(n - c_{Pa,k,u})]. \quad (1.98)$$

Doktorant dokonał także aproksymacji, w której liczbę obsługiwanych zgłoszeń ruchu typu Engseta i Pascala, uwzględnionych w (1.95), zastąpiono ich średnimi wartościami określonymi wzorem (1.96). Średnią liczbę obsługiwanych zgłoszeń w poszczególnych stanach zajętości systemu można wyznaczyć na podstawie procesu obsługi zgłoszeń na poziomie makrostanów. W tym celu konieczne jest określenie średniej liczby obsługiwanych zgłoszeń w systemie. Aby wyznaczyć wartości liczbowe prawdopodobieństwa stanów doktorant zastosował modyfikację iteracyjnej metody MIM-BPP, zaprezentowanej w pracy [53]. Zmiana ta polegała na uwzględnieniu mechanizmu kompresji zgłoszeń. Wartości prawdopodobieństw $P(n)$ oraz średnia liczba obsługiwanych zgłoszeń

ruchu typu d klasy w typu kompresji u w makrostanie n stanowią wartości wyznaczone w ostatnim kroku iteracji metody MIM-BPP [53]. Schemat modyfikacji został szczegółowo opisany w [J3].

Dzięki uwzględnieniu przyjętych aproksymacji w zmodyfikowanej metodzie MIM-BPP, doktorant określił rozkład zajętości systemu. Na jego podstawie możliwym jest określenie charakterystyk systemu, takich jak:

- Średnia liczba zajętych JA w systemie:

$$E_C = \sum_{n=0}^{C_r+C_v} nP(n) + \sum_{n=C_r+C_v+1}^{C_r+C_v+C_q} (C_r + C_v)P(n). \quad (1.99)$$

- Średnia długość kolejki:

$$Q = \sum_{n=C_r+C_v+1}^{C_r+C_v+C_q} (n - C_r - C_v)P(n). \quad (1.100)$$

- Prawdopodobieństwo blokady:

$$E_{d,i,u} = \sum_{n=C_r+C_v+C_q-c_{d,i,u}+1}^{C_r+C_v+C_q} P(n). \quad (1.101)$$

- Prawdopodobieństwo strat:

- Dla zgłoszeń ruchu Erlanga:
w przypadku zgłoszeń typu Erlanga prawdopodobieństwo strat jest równe prawdopodobieństwu blokady:

$$B_{Er,d,u} = E_{Er,d,u}; \quad (1.102)$$

- Dla zgłoszeń ruchu Engseta:

$$B_{En,i,u} = \frac{\sum_{n=C_r+C_v+C_q-c_{En,i,u}+1}^{C_r+C_v+C_q} \Lambda_{En,i,u} [N_{En,i,u} - y_{En,i,u}(n)] P(n)}{\sum_{n=0}^{C_r+C_v+C_q} \Lambda_{En,i,u} [N_{En,i,u} - y_{En,i,u}(n)] P(n)}; \quad (1.103)$$

- Dla zgłoszeń ruchu Pascala:

$$B_{Pa,i,u} = \frac{\sum_{n=C_r+C_v+C_q-c_{Pa,i,u}+1}^{C_r+C_v+C_q} \Lambda_{Pa,w,u} [N_{Pa,i,u} - y_{Pa,i,u}(n)] P(n)}{\sum_{n=0}^{C_r+C_v+C_q} \Lambda_{Pa,i,u} [N_{Pa,i,u} - y_{Pa,i,u}(n)] P(n)}. \quad (1.104)$$

W celu weryfikacji poprawności oraz dokładności zaproponowanego modelu systemu kolejkowego z ruchem Erlanga, Engseta oraz Pascala typu elastycznego i adaptacyjnego, doktorant porównał wyniki modelowania analitycznego z wynikami eksperymentów symulacyjnych. W tym celu wykorzystał autorski symulator opisany w punkcie 1.6.5.

Tabela 1.10: Parametry analizowanego systemu [J3]

System 10				
Pojemność	Rzeczywista	Wirtualna	Kolejka	Całkowita
-	10	10	5	25
Klasy zgłoszeń				
Numer klasy (i)	Żądanie (c)	Typ Ruchu	Rodzaj ruchu	Liczba źródeł (N)
1	1	elastyczny	Erlang	-
2	2	adaptacyjny	Erlang	-
3	2	elastyczny	Engset	40
4	3	adaptacyjny	Engset	40
5	2	elastyczny	Pascal	70
6	1	adaptacyjny	Pascal	70

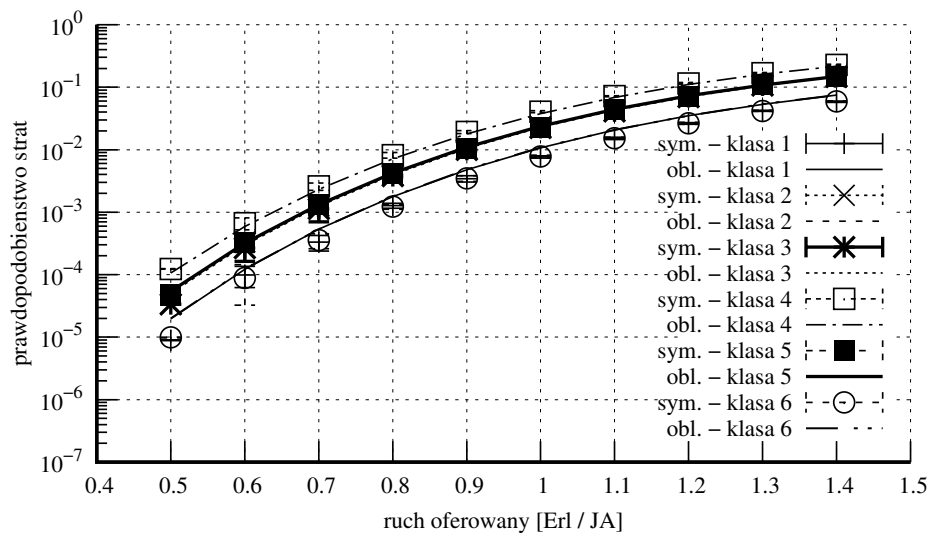
Każdy eksperyment składa się z 8 serii symulacji, każda z minimum 500 000 zgłoszeń klasy, która wymagała największej liczby JA. Wyniki symulacji na wykresach zostały zaprezentowane w postaci punktów z zaznaczonym przedziałem ufności 95%, obliczonym zgodnie z rozkładem t-Studenta. Uzyskane wyniki zaprezentowane zostały w funkcji oferowanego ruchu na pojedynczą JA systemu:

$$a = \frac{\sum_{i=1}^M A_{i,u}}{C_r}. \quad (1.105)$$

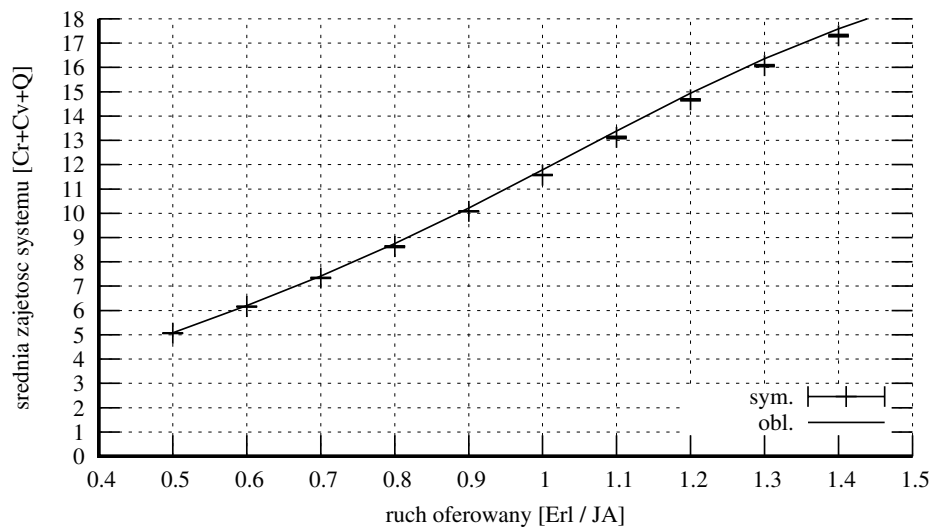
Przyjęto też, że mieszanina różnych klas ruchu jest oferowana systemowi w następującej proporcji:

$$a_1 c_1 : a_2 c_2 : \dots : a_M c_M = 1 : 1 : \dots : 1, \quad (1.106)$$

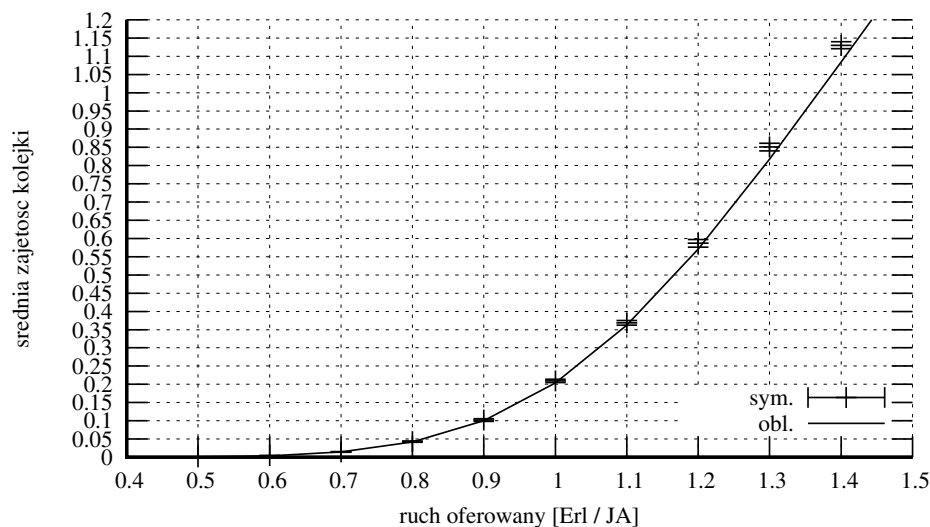
gdzie M oznacza liczbę wszystkich klas ruchu oferowanych systemowi.



Rysunek 1.21: Prawdopodobieństwo strat w systemie 10 [J3].



Rysunek 1.22: Średnia zajętość Systemu 10 [J3].



Rysunek 1.23: Średnia zajętość kolejki w Systemie 10 [J3].

1.6.5 Modułowy symulator umożliwiający analizę złożonych systemów teleinformatycznych z ruchem Erlanga, Engseta, Pacala typu strumieniowego, elastycznego oraz adaptacyjnego.

Doktorant w celu weryfikacji i oceny dokładności opracowanych przez siebie modeli analitycznych systemów teleinformatycznych opracował oraz zaimplementował autorski model symulacyjny. Symulator umożliwia wyznaczenie parametrów jakościowych takich jak: prawdopodobieństwo blokady, prawdopodobieństwo strat, wykorzystanie zasobów systemu oraz kolejki, a także analizę wpływu poszczególnych parametrów systemu na jego działanie. W modelu symulacyjnym wyróżnione zostały moduły odpowiadające za jego działanie, takie jak:

- Dane wejściowe – wprowadzenie do systemu podstawowych parametrów, takich jak: rodzaj i liczba klas zgłoszeń wraz z ich rozmiarem i dostępnością, parametry systemu (struktura zewnętrzna i wewnętrzna) wraz z pojemnościami, określenie kolejki w systemie wraz z algorytmem zarządzania;
- Zgłoszenia – inicjowanie poszczególnych klas zgłoszeń, wraz z ich podstawową obsługą zaimplementowaną w wyróżnionych metodach, zależnie od przyjętych parametrów (rodzaj ruchu (BPP), typ ruchu elastyczny, adaptacyjny, strumieniowy);
- System – inicjowanie struktury systemu (liczby maszyn fizycznych, dostępności w ich obszarze) oraz metod służących obsłudze serwerów;
- Symulacja – inicjowanie i kontrola nad procesem przeprowadzanej symulacji zgodnie z przyjętym scenariuszem;
- Zdarzenia – opis i realizacja występujących w systemie zdarzeń czasowych oraz warunkowych;
- Lista zdarzeń – przechowywanie zdarzeń występujących w systemie i ich obsługi zgodnie z kolejnością występowania
- Generatory – generowanie liczb losowych zgodnie z przyjętymi rozkładami;
- Statystyka – wyznaczanie danych statystycznych związanych z parametrami jakościowymi systemu;
- Dane wyjściowe – zapis w plikach wyjściowych danych statystycznych będących wynikiem przeprowadzonego procesu symulacji cyfrowej.

Opracowany przez doktoranta model symulacyjny dyskretny z czasem ciągłym został zaimplementowany w języku C++ przy pomocy metody planowania zdarzeń [54]. W modelu tym zarówno aktywność źródeł ruchu, jak i proces zajmowania zasobów systemu zmienia się w sposób skokowy. W modelu wyróżnione zostały dwa zdarzenia czasowe związane z pojawieniem się nowego zgłoszenia, a także zakończeniem jego obsługi. Dodatkowo wyróżniono jedno zdarzenie warunkowe, związane z przyjęciem zgłoszenia do obsługi. Warunkowość tego zdarzenia wynika z zależności od stanu systemu, ponieważ zgłoszenie może zostać przyjęte do obsługi tylko wtedy, gdy w systemie zgodnie z przyjętym algorytmem dystrybucji zgłoszeń oraz przyjętą architekturą, występuje wystarczająca liczba wolnych zasobów do obsługi jego obsługi.

Ze względu na wykorzystane metody planowania zdarzeń, zaimplementowana została lista uporządkowana względem czasu, na której umieszczane były występujące w systemie zdarzenia. W przypadku pojawienia się w systemie nowego zdarzenia, lista każdorazowo podlega sortowaniu, tak aby zdarzenie umieszczone zostało na niej w odpowiedniej dla siebie kolejności. Wszystkie parametry zgłoszeń przechowywane są w dedykowanym dla zgłoszeń module i weryfikowane w momencie wystąpienia zdarzenia w którym dane zgłoszenie jest obsługiwane.

W momencie wystąpienia zdarzenia weryfikowane są parametry zgłoszenia z nim skojarzonego, takie jak rodzaj ruchu, typ, żądania itp. Zależnie od rodzaju zdarzenia,

weryfikacji podlega albo aktualny stan zajętości systemu w celu przekazania zgłoszenia do obsługi lub następuje zwolnienie zajmowanych zasobów w przypadku zakończenia obsługi. Jeżeli zgłoszenie zostanie przyjęte do obsługi, zasoby są zajmowane oraz generowany jest czas zakończenia zgodnie z przyjętym rozkładem.

Zaimplementowany model symulacyjny każdorazowo obsługuje najbliższe względem czasu zdarzenie, a po obsłudze skojarzonych z nim zadań, następuje jego usunięcie z listy i przejście do kolejnego zdarzenia. Czas symulacji jest reprezentowany w sposób względny i wskazuje czas pozostały do wystąpienia kolejnego zdarzenia. Czas ten po wystąpieniu zdarzenia każdorazowo jest zwiększany o zapisaną wartość czasu wystąpienia pierwszego zdarzenia z listy.

Całość modelu podzielona jest na dwie fazy, fazę początkową oraz fazę stacjonarną. Celem takiego podziału jest doprowadzenie procesu symulacji do stanu, w którym napływ nowych zgłoszeń nie wpływa znacząco na wyznaczone parametry jakościowe. W związku z tym, przejście przez fazę początkową umożliwia doprowadzenie procesu symulacji do stanu stacjonarnego działania systemu. Proces symulacji kończy się w momencie wystąpienia określonej przez użytkownika liczby zdarzeń przyjęcia do obsługi zgłoszeń wybranej klasy.

Schemat działania modelu symulacyjnego można zapisać w sposób następujący:

- Krok 1:** Wstępna konfiguracja modelu symulacyjnego – utworzenie wszystkich źródeł generujących zgłoszenia różnych klas ruchu, ustawienie parametrów wejściowych;
- Krok 2:** Zainicjowanie modelu poprzez aktywację źródeł ruchu i umieszczenie zdarzeń (pojawienie się zgłoszenia) na liście zdarzeń;
- Krok 3:** Sprawdzenie warunku końca symulacji. Jeśli warunek końca jest spełniony, symulacja jest przerywana, a wyniki zapisywane do pliku;
- Krok 4:** Uaktualnienie czasu systemowego do czasu wystąpienia pierwszego zdarzenia z listy.
- Krok 5:** Wykonanie pierwszego zdarzenia z listy zdarzeń;
- Krok 6:** Usunięcie pierwszego zdarzenia z listy i powrót do Kroku 3.

Doktorant w modelu symulacyjnym zaimplementował generatory liczb pseudolosowych. Podstawowym zaimplementowanym generatorem był generator rozkładu równomiernego, stanowiący bazę dla pozostałych generatorów. W tym celu opracowany został generator liniowy multiplikatywny, zaproponowany w pracy [55], działający zgodnie ze wzorem:

$$X_{n+1} = (aX_n)\%(m), \quad (1.107)$$

gdzie:

- $a = 16807$,
- $m = 2^{31} - 1$.

Opracowany generator stał się podstawą generatora liczb wykładniczych, niezbędnego w procesie analizy systemów z ruchem Erlanga, Engseta oraz Pascala. Rozkład wykładniczy został otrzymany zgodnie z metodą odwracania dystrybuanty (metodę inwersyjną), której podstawą jest Twierdzenie 1.6.1:

Twierdzenie 1.6.1 *Jeżeli zmienna losowa U ma rozkład równomierny na odcinku $(0; 1)$, to $F^{-1}(U)$ ma dystrybuantę F . I na odwrót. Jeżeli X ma dystrybuantę F , to $F(X)$ ma rozkład równomierny.*

Z racji tego, że dla rozkładu wykładniczego dystrybuanta opisana jest wzorem:

$$\begin{aligned} F(X) &= 0 \quad \wedge \quad x < 0, \\ F(X) &= 1 - e^{-\lambda x} \quad \wedge \quad (\lambda > 0 \wedge x \geq 0), \end{aligned} \quad (1.108)$$

to możliwym jest wyznaczenie dystrybuanty odwrotnej rozkładu dla $x \geq 0$ stanowiącą podstawę rozkładu wykładniczego:

$$X = -\frac{1}{\lambda} \ln(1 - u), \quad (1.109)$$

gdzie:

- X – stanowi kolejną wyznaczoną liczbę losową o rozkładzie wykładniczym,
- u – stanowi kolejną wartość wyznaczoną przez generator wykładniczy.

W związku z tym, że U jest liczbą losową o rozkładzie równomiernym określoną na przedziale $(0; 1)$, to rozkład wykładniczy może być opisany jako:

$$X = -\frac{1}{\lambda} \ln(u). \quad (1.110)$$

W opracowanym modelu symulacyjnym zaimplementowana została możliwość obsługi o systemów złożonej budowie: z ruchem Erlanga, Engseta oraz Pascala, typu strumieniowego, adaptacyjnego oraz elastycznego.

1.6.5.1 Charakterystyka struktury systemów

Opracowany symulator umożliwia analizę systemów o złożonej strukturze. W momencie inicjalizacji systemu, określana jest struktura systemu w następujący sposób:

- Charakterystyka pojemności systemu oraz zgłoszeń – system jedno lub wieloparametryczny;
- Liczba występujących w systemie maszyn fizycznych;
- Pojemność systemu – wszystkich jego zasobów;
- Struktura maszyn fizycznych – pełna dostępność lub ograniczona dostępność dla wybranych zasobów.

W symulatorze zaimplementowany został szereg możliwych do zrealizowania scenariuszy. Scenariusze te bazują na strukturze systemu i algorytmach przyjmowania zgłoszenia do obsługi. W modelu symulacyjnym wyróżnione zostały elementy umożliwiające budowanie różnych scenariuszy badań:

- Weryfikacja liczby dostępnych podsystemów – w momencie pojawienia się zgłoszenia analizie mogą podlegać wszystkie występujące w systemie podsystemy (maszyny fizyczne) w sposób równomierny lub tylko wybrane, np. jedna losowa maszyna fizyczna, najmniej obciążona maszyna fizyczna itp.
- Struktura wewnętrzna podsystemu (maszyny fizycznej) może być pełnodostępna lub niepełnodostępna z ograniczeniem dostępu do wybranych zasobów maszyny;
- Analizie może podlegać więcej niż jeden dostępny w podsystemach parametr opisujący rozmiar maszyny (analiza systemów jedno lub wieloparametrycznych);
- Podsystemy (maszyny fizyczne) mogą różnić się między sobą zarówno rozmiarem (niezależnie dla każdego z parametrów), jak i strukturą (pełnodostępna, niepełnodostępna);
- Możliwym jest dodanie pojedynczej kolejki do systemu lub kolejek dedykowanych poszczególnym klasom zgłoszeń. Kolejki te mogą być obsługiwane zgodnie z różnymi algorytmami kolejkowania (FIFO, LIFO, C-FIFO, S-FIFO).

1.6.5.2 Obsługa strumieni Erlanga, Engseta, Pascala

Oprócz określenia struktury systemu, w momencie konfiguracji należy określić typy źródeł ruchu występujące w systemie. W symulatorze zaimplementowana została możliwość wyboru ruchu typu Erlanga, Engseta oraz Pascala. Dla każdej klasy zgłoszeń, wyróżnianego poprzez strumień zgłoszeń oraz obsługi przypisany został osobny generator zwracający wartości zgodne z rozkładem wykładniczym. Wyróżnione zostały funkcje realizujące zdarzenia związane z pojawieniem się zgłoszeń kolejno Erlanga, Engseta, Pascala.

Dla klas zgłoszeń Erlanga obsługa może być określona w następujący sposób:

Krok 1: Zaplanowanie pojawienia się nowego zgłoszenia klasy i zgodnie z rozkładem wykładniczym, z parametrem intensywności λ i umieszczenie zdarzenia na liście zdarzeń zgodnie z czasem wystąpienia.

Krok 2: Obsługa zgłoszenia w systemie poprzez weryfikację możliwości przyjęcia do obsługi zgodnie z przyjętym algorytmem dystrybucji zgłoszeń i struktury systemu:

- w przypadku przyjęcia zgłoszenia do obsługi następuje zaplanowanie zdarzenia zakończenia obsługi i umieszczenie go na liście zdarzeń zgodnie z czasem wystąpienia;
- w przypadku braku miejsca w systemie następuje odrzucenie zgłoszenia lub umieszczenie w kolejce, zależnie od przyjętego algorytmu.

Krok 3: Usunięcie zdarzenia z listy zdarzeń.

Krok 4: Sprawdzenie warunku końca symulacji. Jeśli warunek ten jest spełniony, zakończenie procesu symulacji. Jeśli warunek ten nie jest spełniony, uaktualnienie czasu systemowego do czasu pierwszego zdarzenia na liście zdarzeń.

Dla klas zgłoszeń Engseta obsługa może być określona w następujący sposób:

Krok 1: Obsługa zgłoszenia w systemie poprzez weryfikację możliwości przyjęcia do obsługi, zgodnie z przyjętym algorytmem dystrybucji zgłoszeń i struktury systemu:

- w przypadku przyjęcia zgłoszenia do obsługi następuje zaplanowanie zdarzenia zakończenia obsługi i umieszczenie go na liście zdarzeń zgodnie z czasem wystąpienia;
- w przypadku braku miejsca w systemie następuje odrzucenie zgłoszenia i zaplanowanie pojawienia się nowego zgłoszenia klasy i zgodnie z rozkładem wykładniczym, z parametrem intensywności skojarzonym z pojedynczym źródłem ruchu Engseta i umieszczenie zdarzenia na liście zdarzeń zgodnie z czasem wystąpienia lub umieszczenie w kolejce, zależnie od przyjętego algorytmu.

Krok 3: Usunięcie zdarzenia z listy zdarzeń.

Krok 4: Sprawdzenie warunku końca symulacji. Jeśli warunek ten jest spełniony, zakończenie procesu symulacji. Jeśli warunek ten nie jest spełniony, uaktualnienie czasu systemowego do czasu pierwszego zdarzenia na liście zdarzeń.

Jak można zauważyć, w przypadku źródeł ruchu Engseta, w odróżnieniu od źródeł ruchu Erlanga, generowanie nowego zgłoszenia następuje tylko w przypadku odrzucenia zgłoszenia w systemie lub w momencie zakończenia obsługi zgłoszenia w systemie. Wiąże się to z ograniczoną liczbą źródeł ruchu, w których to źródło może wygenerować zgłoszenie wtedy i tylko wtedy, gdy źródło to jest wolne.

Natomiast dla klas zgłoszeń Pascala obsługa może być określona w następujący sposób:

Krok 1: Zaplanowanie pojawienia się nowego zgłoszenia klasy i zgodnie z rozkładem wykładniczym, z parametrem intensywności skojarzonym ze źródłem klasy Pascala i umieszczenie zdarzenia na liście zdarzeń zgodnie z czasem wystąpienia.

Krok 2: Obsługa zgłoszenia w systemie poprzez weryfikację możliwości przyjęcia do obsługi zgodnie z przyjętym algorytmem dystrybucji zgłoszeń i struktury systemu:

- w przypadku przyjęcia zgłoszenia do obsługi, następuje zaplanowanie zdarzenia zakończenia obsługi i umieszczenie go na liście zdarzeń zgodnie z czasem wystąpienia. Ponadto zachodzi dodanie nowego źródła i zaplanowanie pojawienia się nowego zgłoszenia klasy i zgodnie z rozkładem wykładniczym, z parametrem intensywności skojarzonym ze źródłem klasy Pascala i umieszczenie zdarzenia na liście zdarzeń zgodnie z czasem wystąpienia;

- w przypadku braku miejsca w systemie odrzucenie zgłoszenia lub umieszczenie w kolejce, zależnie od przyjętego algorytmu.

Krok 3: Usunięcie zdarzenia z listy zdarzeń.

Krok 4: Sprawdzenie warunku końca symulacji. Jeśli warunek ten jest spełniony, zakończenie procesu symulacji. Jeśli warunek ten nie jest spełniony, uaktualnienie czasu systemowego do czasu pierwszego zdarzenia na liście zdarzeń.

W przypadku procesu zdarzeń klasy Pascala, kluczowym jest weryfikacja w momencie zakończenia obsługi, czy zgłoszenie powstałe w wyniku wygenerowania kopii źródła ruchu jest aktualnie obsługiwane. Jeżeli w momencie zakończenia obsługi zgłoszenia, zdarzenie związane z pojawieniem się zgłoszenia wygenerowane przez kopię nie zostało jeszcze obsłużone, zgłoszenie to jest usuwane z listy zdarzeń. Natomiast w przypadku, w którym w momencie zakończenia obsługi zgłoszenia zdarzenie związane z pojawieniem się zgłoszenia wygenerowanego przez kopię źródła ruchu zostało obsłużone, w systemie nic się nie dzieje.

1.6.5.3 Metody zarządzania ruchem – kompresja bezprogowa

Opracowany przez doktoranta model symulacyjny umożliwia także analizę systemów z kompresją bezprogową. W systemach tych oprócz ruchu strumieniowego, można wyróżnić także ruch elastyczny oraz adaptacyjny. Ruch elastyczny i adaptacyjny są typami ruchu, w których w momencie braku dostępnych zasobów w systemie, stosowana jest kompresja wszystkich zgłoszeń (typu elastycznego i adaptacyjnego) już obsługiwanych, tak aby zgłoszenie to mogło zostać przyjęte do obsługi. Jednocześnie czas obsługi zgłoszeń typu elastycznego jest wydłużany, proporcjonalnie do zastosowanego współczynnika kompresji.

W modelu symulacyjnym doktorant, w celu uniknięcia kompresji zgłoszeń i zmiany ich rozmiaru, zastosował dodatkowy obszar, nazywany pojemnością wirtualną. W związku z tym, cała pojemność systemu wyrażona jest jako suma pojemności rzeczywistej systemu C_r oraz pojemności wirtualnej C_v . Pojemność części wirtualnej określana jest na podstawie maksymalnego poziomu kompresji, który wynika ze stosunku całej pojemności systemu ($C_r + C_v$) do pojemności części rzeczywistej (C_r). Następnie zgłoszenia pojawiające się na wejściu systemu zależnie od swojego typu obsługiwane są w różny sposób. Dla ruchu typu strumieniowego, obsługa może być wyrażona w następujący sposób:

Krok 1: Weryfikacja czy w systemie zgodnie z przyjętym algorytmem dystrybucji zgłoszeń oraz charakterystyką dostępności występuje wystarczająca liczba wolnych zasobów w części rzeczywistej systemu:

- Jeżeli tak, zgłoszenie typu strumieniowego przyjmowane jest do obsługi;
- Jeżeli nie, weryfikowane jest czy w systemie występuje wystarczająca liczba zasobów w części wirtualnej, która w przypadku kompresji zgłoszeń typu elastycznego i adaptacyjnego umożliwiłaby przyjęcie zgłoszenia typu strumieniowego w części rzeczywistej systemu;

- Jeżeli tak, zgłoszenia typu adaptacyjnego lub elastycznego jest przenoszone do części wirtualnej, jednocześnie, wyznaczany jest nowy współczynnik kompresji. Czas trwania pozostały do obsługi zgłoszeń ruchu elastycznego jest zwiększany zgodnie z nowym współczynnikiem kompresji, a lista zdarzeń sortowana. Co za tym idzie, czas obsługi zgłoszeń elastycznych wydłużany jest proporcjonalnie do zajętości części wirtualnej (stosunek stanu zajętości n do pojemności rzeczywistej C_r);
- Jeżeli nie, zgłoszenie ruchu strumieniowego jest odrzucane lub umieszczane w kolejce.

Dla ruchu typu adaptacyjnego, obsługa może być wyrażona w następujący sposób:

Krok 1: Weryfikacja czy w systemie zgodnie z przyjętym algorytmem dystrybucji zgłoszeń oraz charakterystyką dostępności występuje wystarczająca liczba wolnych zasobów w części rzeczywistej systemu:

- Jeżeli tak, zgłoszenie typu adaptacyjnego przyjmowane jest do obsługi;
- Jeżeli nie, weryfikowane jest czy w systemie występuje wystarczająca liczba zasobów w części wirtualnej, do przyjęcia tego zgłoszenia;
 - Jeżeli tak, zgłoszenie to jest przyjmowane do obsługi, a jednocześnie, wyznaczany jest nowy współczynnik kompresji. Czas trwania pozostały do obsługi zgłoszeń ruchu elastycznego jest zwiększany zgodnie z nowym współczynnikiem kompresji, a lista zdarzeń sortowana. Co za tym idzie, czas obsługi zgłoszeń elastycznych, wydłużany jest proporcjonalnie do zajętości części wirtualnej (stosunek stanu zajętości n do pojemności rzeczywistej C_r);
 - Jeżeli nie, zgłoszenie ruchu adaptacyjnego jest odrzucane lub umieszczane w kolejce.

Jednocześnie dla ruchu typu elastycznego, obsługa może być wyrażona w następujący sposób:

Krok 1: Weryfikacja czy w systemie zgodnie z przyjętym algorytmem dystrybucji zgłoszeń oraz charakterystyką dostępności występuje wystarczająca liczba wolnych zasobów w części rzeczywistej systemu.

- Jeżeli tak, zgłoszenie typu elastycznego przyjmowane jest do obsługi.
- Jeżeli nie, weryfikowane jest czy w systemie występuje wystarczająca liczba zasobów w części wirtualnej, do przyjęcia tego zgłoszenia;
 - Jeżeli tak, zgłoszenie to jest przyjmowane do obsługi, a jednocześnie, wyznaczany jest nowy współczynnik kompresji. Czas trwania pozostały do obsługi zgłoszeń ruchu elastycznego (w tym nowo przyjętego zgłoszenia) jest zwiększany zgodnie z nowym współczynnikiem kompresji, a lista zdarzeń sortowana. Co za tym idzie, czas obsługi zgłoszeń elastycznych, wydłużany jest proporcjonalnie do zajętości części wirtualnej (stosunek stanu zajętości n do pojemności rzeczywistej C_r);

- Jeżeli nie, zgłoszenie ruchu elastycznego jest odrzucane lub umieszczane w kolejce.

1.6.5.4 Zakończenie symulacji i wyznaczenie mierzonych parametrów

W procesie symulacji, doktorant przyjął warunek jej zakończenia zdefiniowany przez liczbę pojawiających się zgłoszeń wybrane klasy. W procesie analizy systemów opisanych w rozprawie doktorskiej, każdorazowo wybierana była klasa zgłoszeń, której zgłoszenia pojawiały się najrzadziej, tak aby zapewnić minimalną wymaganą wartość liczby pojawiających się zgłoszeń. Każdy proces symulacji wykonywany był wielokrotnie, a wyniki zbierane dla każdej próby zapisywane do pliku. W większości przeprowadzonych badań wykonywanych było od 8 do 10 serii symulacji, w których w systemie pojawiała się od 800 tysięcy do miliona zgłoszeń klasy najmniej aktywnej. Pozwoliło to na uzyskanie 95% przedziału ufności zgodnie z rozkładem t-Studenta dla wszystkich wyznaczanych parametrów.

Za analizę statystyczną w symulatorze odpowiadał moduł Statystyka, w którym to, po przeprowadzeniu wszystkich serii symulacji, przedziały ufności były wyznaczone zgodnie ze wzorem:

$$\left(\bar{X} - t_{\alpha} \frac{\sigma}{\sqrt{r}}; \bar{X} + t_{\alpha} \frac{\sigma}{\sqrt{r}} \right), \quad (1.111)$$

gdzie:

- \bar{X} – średnia arytmetyczna obliczana ze wszystkich wyznaczonych wartości w określonej liczbie serii symulacji,
- t_{α} – wartość rozkładu t-Studenta dla $r - 1$ stopni swobody,
- σ – odchylenie standardowe.

Bibliografia

- [1] Cisco Annual Internet Report - Cisco annual Internet report (2018–2023) White Paper. Publication Title: Cisco.
- [2] Ericsson Mobility Report Q4 update February 2022. strona 4, 2022.
- [3] Timm Böttger, Ghida Ibrahim, Ben Vallis. How the Internet reacted to Covid-19: A perspective from Facebook’s edge network. *Proceedings of the ACM Internet measurement conference*, strony 34–41, Virtual Event USA, Październik 2020. ACM.
- [4] World Development Report 2021 team. World Development Report 2021: data for better lives \textbar How COVID-19 increased data consumption and highlighted the digital divide.
- [5] Michał P. Karpowicz. COVID-19 pandemic and Internet traffic in Poland: evidence from selected regional networks. *Journal of Telecommunications and Information Technology*, 3(2021):86–91, Wrzesień 2021.
- [6] Commission and European regulators calls on streaming services, operators and users to prevent network congestion | shaping Europe’s digital future, online: <https://digital-strategy.ec.europa.eu/en/news/commission-and-european-regulators-calls-streaming-services-operators-and-users-prevent-network>, Marzec 2020.
- [7] Cloud computing market to reach \$1,554.94Bn by 2030, online: <https://www.grandviewresearch.com/press-release/global-cloud-computing-market>.
- [8] Peter Mell, Timothy Grance. The NIST definition of Cloud Computing.
- [9] Infographic: Amazon maintains cloud lead as Microsoft edges closer | statista daily data, online: <https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers>, Luty 2024.
- [10] Seria maszyn wirtualnych | Microsoft Azure, online: <https://azure.microsoft.com/pl-pl/pricing/details/virtual-machines/series/>.
- [11] EC2 on-demand instance pricing – Amazon Web Services, online: <https://aws.amazon.com/ec2/pricing/on-demand/>.

- [12] Maciej Stasiak. Queuing systems for the Internet. *IEICE Transactions on Communications*, E99-B(6):1234–1242, Czerwiec 2016.
- [13] Sławomir Hanczewski, Maciej Stasiak, Joanna Weissenberg. The model of the queuing system with adaptive traffic. *2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, strony 303–308, Grudzień 2017.
- [14] Sławomir Hanczewski, Maciej Stasiak, Joanna Weissenberg. A queueing model of a multi-service system with state-dependent distribution of resources for Each Class of Calls. *IEICE Transactions on Communications*, E97-B(8):1592–1605, Sierpień 2014.
- [15] Sławomir Hanczewski, Adam Kaliszan, Maciej Stasiak. Convolution model of a queueing system with the cFIFO service discipline. *Mobile Information Systems*, 2016:e2185714, Listopad 2016.
- [16] Sławomir Hanczewski, Maciej Stasiak, Joanna Weissenberg. Non-full-available queueing model of an EON node. *Optical Switching and Networking*, 33:131–142, Lipiec 2019.
- [17] Carlos Oliveira, Jaime Bae Kim, Tatsuya Suda. An adaptive bandwidth reservation scheme for high-speed multimedia wireless networks. *IEEE Journal on Selected Areas in Communications*, 16(6):858–874, Sierpień 1998.
- [18] Maciej Stasiak, Sławomir Hanczewski. Approximation for multi-service systems with reservation by systems with limited-availability. Nigel Thomas, Carlos Juiz, redaktorzy, *Computer Performance Engineering*, Lecture Notes in Computer Science, strony 257–267, Berlin, Heidelberg, 2008. Springer.
- [19] Ioannis D. Moscholios, Michael D. Logothetis, Anthony C. Boucouvalas. Blocking probabilities of elastic and adaptive calls in the Erlang multirate loss model under the threshold policy. *Telecommunication Systems*, 62(1):245–262, Maj 2016.
- [20] K. Subramaniam, A.A. Nilsson. An analytical model for adaptive call admission control scheme in a heterogeneous UMTS-WCDMA system. *IEEE International Conference on Communications, 2005. ICC 2005. 2005*, wolumen 5, strony 3334–3338 Vol. 5, Maj 2005.
- [21] Sławomir Hanczewski, Maciej Stasiak, Piotr Zwierzykowski. Modelling of the access part of a multi-service mobile network with service priorities. *EURASIP Journal on Wireless Communications and Networking*, 2015(1):194, Lipiec 2015.
- [22] Joseph S. Kaufman. Blocking in a shared resource environment. *IEEE Transactions on Communications*, 29(10):1474–1481, Październik 1981.

- [23] Sándor Rácz, Balázs Péter Gerő, Gábor Fodor. Flow level performance analysis of a multi-service system supporting elastic and adaptive services. *Performance Evaluation*, 49(1):451–469, Wrzesień 2002.
- [24] Sławomir Hanczewski, Maciej Stasiak, Joanna Weissenberg. Model of a multiservice server with stream and elastic traffic. Michał Choraś, Ryszard S. Choraś, redaktorzy, *Image Processing and Communications, Advances in Intelligent Systems and Computing*, strony 287–294, Cham, 2020. Springer International Publishing.
- [25] Sławomir Hanczewski, Maciej Stasiak, Joanna Weissenberg. A model of a system with stream and elastic traffic. *IEEE Access*, 9:7789–7796, 2021.
- [26] Shun-Ping Chung, Jin-Chang Lee. Performance analysis and overflowed traffic characterization in multiservice hierarchical wireless networks. *IEEE Transactions on Wireless Communications*, 4(3):904–918, Maj 2005.
- [27] Mariusz Głabowski, Damian Kmiecik, Maciej Stasiak. Overflows in multiservice systems. *IEICE Transactions on Communications*, E102-B(5):958–969, Maj 2019.
- [28] Mariusz Głabowski, Damian Kmiecik, Maciej Stasiak. On increasing the accuracy of modeling multi-service overflow systems with Erlang-Engset-Pascal streams. *Electronics*, 10(4):508, Styczeń 2021.
- [29] Jorma Virtamo, Allan T. Andersen. Broadband network reletraffic: final report of Action COST 242. *Broadband Network Teletraffic*, 1996.
- [30] Joseph Y. Hui. Resource allocation for broadband networks. *IEEE Journal on Selected Areas in Communications*, 6(9):1598–1608, Grudzień 1988.
- [31] Thomas Bonald, James W. Roberts. Internet and the Erlang formula. *ACM SIGCOMM Computer Communication Review*, 42(1):23–30, Styczeń 2012.
- [32] Ehsan Ataie, Reza Entezari-Maleki, Leila Rashidi, Kishor Trivedi, Ali Movaghar. Hierarchical stochastic models for performance, availability, and power consumption analysis of IaaS clouds. *IEEE Transactions on Cloud Computing*, PP:1–1, Październik 2017.
- [33] Yonal Kirsal. Performance modelling and analysis of clustered servers in mobile edge computing. strony 1–4, Październik 2020.
- [34] Xiaolin Chang, Ruofan Xia, Jogesh K. Muppala, Kishor S. Trivedi, Jiqiang Liu. Effective modeling approach for IaaS data center performance analysis under heterogeneous workload. *IEEE Transactions on Cloud Computing*, 6(4):991–1003, Październik 2018. Conference Name: IEEE Transactions on Cloud Computing.
- [35] Hamzeh Khazaei, Jelena Mišić, Vojislav Misic, Nasim Beigi. Modeling the performance of heterogeneous IaaS cloud centers. strony 232–237, Lipiec 2013.

- [36] Xiaolin Chang, Bin Wang, Jogesh K. Muppala, Jiqiang Liu. Modeling active virtual machines on IaaS clouds using an M/G/m/m+K queue. *IEEE Transactions on Services Computing*, 9(3):408–420, Maj 2016. Conference Name: IEEE Transactions on Services Computing.
- [37] Dario Bruneo. A stochastic model to investigate data center performance and QoS in IaaS cloud Computing Systems. *IEEE Transactions on Parallel and Distributed Systems*, 25(3):560–569, Marzec 2014. Conference Name: IEEE Transactions on Parallel and Distributed Systems.
- [38] Rahul Ghosh, Francesco Longo, Vijay K. Naik, Kishor S. Trivedi. Modeling and performance analysis of large scale IaaS Clouds. *Future Generation Computer Systems*, 29(5):1216–1234, Lipiec 2013.
- [39] Tajwar Mehmood, Seemab Latif, Shehryaar Malik. Prediction of cloud computing resource utilization. *2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT)*, strony 38–42, Październik 2018. ISSN: 1949-4106.
- [40] Michael Borkowski, Stefan Schulte, Christoph Hochreiner. Predicting cloud resource utilization. *Proceedings of the 9th International Conference on Utility and Cloud Computing*, strony 37–42, Shanghai China, Grudzień 2016. ACM.
- [41] Qazi Zia Ullah, Shahzad Hassan, Gul Muhammad Khan. Adaptive resource utilization prediction system for Infrastructure as a Service cloud. *Computational Intelligence and Neuroscience*, 2017:e4873459, Lipiec 2017. Publisher: Hindawi.
- [42] Sadeka Islam, Jacky Keung, Kevin Lee, Anna Liu. Empirical prediction models for adaptive resource provisioning in the cloud. *Future Generation Comp. Syst.*, 28:155–162, Styczeń 2012.
- [43] K. Beghdad Bey, Farid Benhamadi, Aicha Mokhtari, Zahia Guessoum. CPU load prediction model for distributed computing. *2009 Eighth International Symposium on Parallel and Distributed Computing*, strony 39–45, Czerwiec 2009. ISSN: 2379-5352.
- [44] Rongting Li, Xiaogang Wang, Dun Xiao, Chao Huang. Cloud instance resources prediction based on hidden Markov model. *2023 IEEE 9th International Conference on Cloud Computing and Intelligent Systems (CCIS)*, strony 516–520, Sierpień 2023. ISSN: 2376-595X.
- [45] Fahimeh Farahnakian, Pasi Liljeberg, Juha Plosila. LiRCUP: linear regression based CPU usage prediction algorithm for live migration of virtual machines in data centers. strony 358–364, Wrzesień 2013.
- [46] Mariusz Głabowski, Adam Kaliszan, Maciej Stasiak. Modeling product-form state-dependent systems with BPP traffic. *Performance Evaluation*, 67(3):174–197, 2010.

- [47] John Roberts. A service system with heterogeneous user requirement. *Performance of Data Communications Systems and Their Applications*, strony 423–431, 1981.
- [48] Maciej Stasiak. Blocking probability in a limited-availability group carrying mixture of different multichannel traffic streams. *Annales des Télécommunications*, Styczeń 1993. Publisher: Springer Science and Business Media LLC.
- [49] Erlang A. K. Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Post Office Electrical Engineer's Journal*, 10:189–197, 1917.
- [50] Mariusz Głabowski, Adam Kaliszan, Maciej Stasiak. Modeling product-form state-dependent systems with bpp traffic. *Performance Evaluation*, 67(3):174–197, 2010.
- [51] Villy Iversen. Teletraffic engineering and network planning. Lyngby, Technical University of Denmark, 2009. Conference Name:IEICE General Conference.
- [52] Sławomir Hanczewski, Maciej Stasiak, Joanna Weissenberg. Queueing model of a multi-service system with elastic and adaptive traffic. *Computer Networks*, 147:146–161, Grudzień 2018.
- [53] Mariusz Głabowski, Maciej Stasiak, Joanna Weissenberg. Properties of recurrent equations for the full-availability group with BPP traffic. *Mathematical Problems in Engineering*, (2012):17, Sierpień 2012.
- [54] Jerzy Tyszer. *Object-oriented computer simulation of discrete-event systems: 10*. Boston, 1999.
- [55] Derrick H. Lehmer. Mathematical methods in large-scale computing units. *Proceedings of the Second Symposium on Large Scale Digital Computing Machinery*, strony 141–146, Cambridge, United Kingdom, 1951. Harvard University Press.

Rozdział 2

Wykaz pozostałych osiągnięć doktoranta

2.1 Spis pozostałych publikacji doktoranta

2.1.1 Publikacje w obszarze modelowania systemów teleinformatycznych

- [D1] Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg. Determining resource utilization in cloud systems: an analytical algorithm for IaaS architecture. *The 17th International Conference on Telecommunications – Contel 2023*. 11-13 Lipiec 2023, Graz, Austria. strony 1–8, 2023.
https:
doi.org/10.1109/Con58387.2023.10199064.[MEiN: 20]
- [D2] Sławomir Hanczewski, Michał Weissenberg. Simulation of queuing systems with non-full-availability server. *11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2018)*. Institute of Electrical and Electronics Engineers (IEEE), 2018. [MEiN: 20]
- [D3] Sławomir Hanczewski, Michał Weissenberg. Simulation studies of a complex non-full-availability systems. *Image Processing and Communications: Techniques, Algorithms and Applications. in Advances in Intelligent Systems and Computing, AISC. Springer*. strony. 303–310, 2020.
https:
doi.org/10.1007/978-3-030-31254-1_36. [MEiN: 20]
- [D4] Mariusz Głąbowski, Piotr Zwierzykowski, Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg, Vito Bai. Traffic modeling for industrial Internet of Things (IIoT) networks. *Image Processing and Communications: Techniques, Algorithms and Applications. in Advances in Intelligent Systems and Computing, AISC. Springer*. strony. 264–271, 2020. [MEiN: 20]
- [D5] Mariusz Głąbowski, Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg, Piotr Zwierzykowski. Modelowanie źródeł ruchu dla sieci przemysłowego internetu

rzeczy. *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*. 2019.
<https://doi.org/10.15199/59.2019.7.9> [MEiN: 5]

- [D6] Sławomir Hanczewski, Michał Weissenberg. Badania symulacyjne złożonych systemów niepełnodostępnych. *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*, 2019.
<https://doi.org/10.15199/59.2019.7.38> [MEiN: 5]
- [D7] Sławomir Hanczewski, Michał Weissenberg. Model złożonego systemu niepełnodostępnego. *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*. 2019.
<https://doi.org/10.15199/59.2019.7.36> [MEiN: 5]
- [D8] Sławomir Hanczewski, Michał Weissenberg. Badania symulacyjne systemów kolejkowych z niepełnodostępnym serwerem. *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*. 2018.
<https://doi.org/10.15199/59.2018.8-9.22> [MEiN: 9]
- [D9] Sławomir Hanczewski, Michał Weissenberg. Koncepcja modelu analitycznego chmury obliczeniowej. *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*. 2018.
<https://doi.org/10.15199/59.2018.8-9.62> [MEiN: 9]
- [D10] Mariusz Głąbowski, Piotr Zwierzykowski, Sławomir Hanczewski, Maciej Stasiak, Michał Weissenberg, Vito Bai. Traffic modeling in Industrial Ethernet Networks. *International Journal of Electronics and Telecommunications*. 2020. [MEiN: 40]

2.1.2 Pozostałe publikacje

- [D11] Bartosz Ptak, Przemysław Aszkowski, Joanna Weissenberg, Marek Kraft, Michał Weissenberg. ISO-Compatible Personal Temperature Measurement Using Visual and Thermal Images With Facial Region of Interest Detection. *IEEE Access*. wolumen 12, strony 44262–44277, 2024.
<https://10.1109/ACCESS.2024.3377448>. [IF: 3.9, MEiN: 100]
- [D12] Michał Weissenberg, Maja Maćkowiak. An overview of lie detection methods. *Proceedings of the 42nd International Business Information Management Association Computer Science Conference IBIMA*. 29-30 Listopad 2023, Sevilla, Hiszpania. Theory and Practice in Modern Computing: From Artificial Intelligence to Computer Security. IBIMA Publishing, strony. 319–325, 2023. [MEiN: 70]
- [D13] Kuba Gryglak, Piotr Kontowicz, Konrad Śniatała, Michał Weissenberg, S.S. Iyengar, Jerry Miller. Data collection in smartwatches detection of vhanages in heartbeat in medical applications. *Proceedings of the 40th International Business Information Management Association (IBIMA) [Tech Conference]*. 23-24 Listopad 2022, Sevilla, Hiszpania. IBIMA Publishing, strony 138–148, 2022. [MEiN: 70]

- [D14] Konrad Śniatała, Yashas Hariprasad, K.J. Latesh Kumar, Naveen Kumar Chaudhary, Michał Weissenberg. Fog forensics: a comprehensive review of forensic models for fog computing environment. *Lecture Notes in Electrical Engineering, LNEE*. Springer, Information Security, Privacy and Digital Forensics: Select Proceedings of the International Conference, ICISPD 2022, strony 31–42, 2023.
https://10.1007/978-981-99-5091-1_3 [MEiN: 20]
- [D15] Mariusz Głąbowski, Michał Weissenberg. Cybersecurity of virtual reality solutions. *NTES 2024, Seminar New Technologies in Education and Society. Abstract Book, Seminar New Technologies in Education and Society, NTES 2024* 4-6 kwiecień 2024. Viseu, Portugalia
- [D16] Mariusz Głąbowski, Jakub Grzelski, Konrad Śniatała, Paweł Śniatała, Michał Weissenberg. Cyber Security - training students and scholars for the challenges of information and communication technologies in eesearch and studies for internationalisation: handbook. Rozdział w książce. University of Maribor Press, online: <https://press.um.si/index.php/ump/catalog/book/753>

2.2 Wykaz udziału w komitetach organizacyjnych i naukowych konferencji krajowych lub międzynarodowych, z podaniem pełnionej funkcji.

2.2.1 Członek komitetu organizacyjnego:

- IEICE Information and Communication Technology Forum, 2017, Poznań, Polska;
- IEICE Information and Communication Technology Forum, 2019, Bydgoszcz, Polska;
- Krajowa Konferencja Radiokomunikacji i Teleinformatyki 2024, Poznań, Polska.

2.3 Wykaz uczestnictwa w pracach zespołów badawczych realizujących projekty, z podziałem na projekty zrealizowane i będące w toku realizacji oraz z uwzględnieniem informacji o pełnionej funkcji w ramach prac zespołów.

2.3.1 Projekty międzynarodowe

- 2021-2024: REUNICE, projekt badawczo-naukowy finansowany ze środków Unii Europejskiej w ramach programu Horyzont 2020 powiązany z projektem EUNICE (European University for Customized Education), lider Pakietu Roboczego

2 (Promoting cooperation between universities and other sectors) oraz od listopada 2023 oficer projektu,

Projekt realizowany przy współpracy z 6 Uniwersytetami: Brandenburgische Technische Universität Cottbus-Senftenbera (Niemcy), University of Cantabria (Hiszpania), University of Mons (Belgia), University of Catania (Włochy), Université Polytechnique Hauts-de-France (Francja), University of Vaasa (Finlandia);

- 2017-2023: EUNICE, projekt Uniwersytetu Europejskiego finansowany ze środków Komisji Europejskiej w ramach programu Erasmus+ (call for European Universities), lider Pakietu Roboczego 4 (Open Virtual System for Education, Research and Innovation),

Projekt realizowany przy współpracy z 6 Uniwersytetami: Brandenburgische Technische Universität Cottbus-Senftenbera (Niemcy), University of Cantabria (Hiszpania), University of Mons (Belgia), University of Catania (Włochy), Université Polytechnique Hauts-de-France (Francja), University of Vaasa (Finlandia);

- 2024 - obecnie: EUNICE4U, rojekt Uniwersytetu Europejskiego finansowany ze środków Komisji Europejskiej w ramach programu Erasmus+ (call for European Universities) będący kontynuacją projektu EUNICE, lider Pakietu Roboczego 5 (Communication, Impact and Dissemination) oraz członek zarządu IT Uniwersytetu Europejskiego EUNICE,

Projekt realizowany przy współpracy z 9 Uniwersytetami: Brandenburgische Technische Universität Cottbus-Senftenbera (Niemcy), University of Cantabria (Hiszpania), University of Mons (Belgia), University of Catania (Włochy), Université Polytechnique Hauts-de-France (Francja), University of Vaasa (Finlandia), University of the Peloponese (Grecja), University of Karlstad (Szwecja), Politecnico de Viseu (Portugalia);

- 2022-2023: Cyberbezpieczeństwo dla Uniwersytetu Europejskiego EUNICE, projekt finansowany ze środków Narodowej Agencji Wymiany Akademickiej (NAWA) w ramach programu Partnerstwa Strategiczne (BPI/PST/2021/1/00073), lider zadania II.10.1. (Przygotowanie pilotażowego laboratorium Wnioskodawcy do pracy zdalnej z lokalizacji Partnerów oraz adaptacja VirtualLab do realizacji zajęć w trybie hybrydowym) oraz wykonawca w pozostałych zadaniach;

Projekt realizowany przy współpracy z 3 Uniwersytetami: Brandenburgische Technische Universität Cottbus-Senftenbera (Niemcy), University of Mons (Belgia), University of Catania (Włochy);

- 2024 - obecnie: Enhancing IoT System Security, projekt finansowany ze środków NATO w ramach programu Science for Peace and Security Programme, stypendysta, uczestnik projektu oraz lider zadania 3 "Develop best practices for the design and construction of IoT systems for smart cities"

Projekt realizowany przy współpracy z 3 Uniwersytetami: University of Mons (Belgia), University of Nis (Serbia), Universiapolis (Maroko) oraz Miastem Poznania;

- 2024 - obecnie: Towards Networked Airborne Computing in Uncertain Airspace: a Control and Networking Facilitated Distributed Computing Framework, projekt finansowany ze środków NAWA w ramach programu IMPRESS_U, wykonawca Projekt realizowany przy współpracy z Partnerem University of San Diego (USA) poprzez dołączenie do projektu finansowanego ze środków National Science Foundation (NSF) w USA;
- 2020-2023: CyberF-IT: Cyber Security - Training Students and Scholars for the Challenges of Information and Communication Technologies Research and Studies for Internationalisation (Cyber F-IT, projekt finansowany ze środków Komisji Europejskiej w ramach programu ERASMUS+, wykonawca. Projekt realizowany przy współpracy z 3 Uniwersytetami: University of Maribor (Słowenia), Brandenburgische Technische Universität Cottbus-Senftenbera (Niemcy), Université Polytechnique Hauts-de-France (Francja). Zakończony publikacją [D16].

Oprócz projektów wymienionych powyżej, doktorant w styczniu 2024 złożył jako kierownik projektu wniosek do NAWA w ramach programu Partnerstwa Strategiczne. Projekt zatytułowany *Inkubatory współpracy międzynarodowej w zakresie badań i dydaktyki dla Uniwersytetu Europejskiego EUNICE*, zgodnie z założeniami ma być realizowany przy współpracy z 4 uniwersytetami (University of Mons (Belgia), University of Cantabria (Hiszpania), Université Polytechnique Hauts-de-France (Francja), Politecnico de Viseu (Portugalia), a całkowity budżet projektu został oszacowany na kwotę 1 777 690,35 złotych. Rozstrzygnięcie konkursu powinno zostać ogłoszone do końca czerwca 2024.

2.3.2 Projekty krajowe

- 2022-2023: Cyberbezpieczeństwo jako katalizator umiędzynarodowienia uczelni i rozwoju kadry, projekt finansowany ze środków NAWA w ramach programu Spinaker, lider zadania II.10.7. (Prowadzenie IMPK w formie hybrydowej w ramach szkoły letniej dla studentów EUNICE i uczelni partnerskich) oraz wykonawca w pozostałych zadaniach;
- 2017-2021: Modelowanie i ocena jakości usług sieci Internet, projekt finansowany ze środków Narodowego Centrum Nauki (NCN), wykonawca;
- 2024 - obecnie: Modele analityczne i struktury optycznych pól komutacyjnych, projekt finansowany z funduszy na działalność statutową dla młodej kadry, kierownik projektu;
- 2021: Optymalizacja oraz modelowanie optycznych pól komutacyjnych, projekt finansowany ze środków z funduszy na działalność statutową dla młodej kadry, wykonawca;
- 2020: Modele analityczne i struktury optycznych pól komutacyjnych, projekt finansowany ze środków z funduszy na działalność statutową dla młodej kadry, wykonawca;

- 2019: Modelowanie wielousługowych sieci i systemów komutacyjnych, projekt finansowany ze środków z funduszy na działalność statutową dla młodej kadry, wykonawca;
- 2018: Modelowanie optycznych, wielousługowych pól komutacyjnych, projekt finansowany ze środków z funduszy na działalność statutową dla młodej kadry, wykonawca.

2.4 Wykaz staży w instytucjach naukowych lub artystycznych, w tym zagranicznych, z podaniem miejsca, terminu, czasu trwania stażu i jego charakteru

Doktorant w trakcie pracy na Politechnice Poznańskiej odbył 5 wyjazdów realizowanych w ramach staży dydaktycznych i szkoleniowych:

- University of Mons, Mons, Belgia - 09-13 maja 2022, wyjazd z grupą studentów specjalności Cybersecurity, realizacja projektów z zakresu IoT Security;
- University of Mons, Mons, Belgia - 24-28 stycznia 2023 - realizacja szeregu aktywności, m.in. zapoznanie się z infrastrukturą badawczą University of Mons, aktualnymi badaniami uczelni oraz udział w warsztatach i seminariach;
- University of Cantabria, Santander, Hiszpania - 26-30 czerwca 2023 - realizacja szeregu aktywności, m.in. prezentacja możliwości wykorzystania wirtualnych laboratoriów Politechniki Poznańskiej;
- University of Mons, Mons, Belgia - 30 sierpnia - 2 września 2024 - realizacja szeregu aktywności, m.in. prezentacja możliwości wykorzystania wirtualnych laboratoriów Politechniki Poznańskiej;
- University of Catania, Katania, Włochy - 20-25 maja 2024, wyjazd z grupą studentów specjalności Cybersecurity, realizacja projektów i zajęć w ramach Malware Analysis oraz prowadzenie 8h zajęć dydaktycznych w obszarze Podstaw cyberbezpieczeństwa, kryminalistyki cyfrowej oraz analizy złośliwego oprogramowania.

2.5 Współpraca z otoczeniem społeczno-gospodarczym

2.5.1 Współpraca z sektorem gospodarczym.

Doktorant od roku 2017 brał udział w 4 projektach realizowanych realizowanych na Politechnice Poznańskiej przy współpracy z otoczeniem gospodarczym. Ponadto był odpowiedzialny za prowadzenie centrum egzaminacyjnego.

- INTEL (2021-2023) – projekt: Algorytmy AI dla bezpiecznego i wydajnego przetwarzania, przechowywania i transmisji danych, wykonawca
Organizacja i budowa Laboratorium Gaudiego na Politechnice Poznańskiej wyposażonego w Super Server do realizacji badań z wykorzystaniem sztucznej inteligencji;
- 2018-2019: Huawei (2018-2019) – projekt: Research on network and protocol architecture for industry4.0/IIoT, wykonawca
Udział w realizacji badań naukowych nad modelowaniem ruchu w sieciach przemysłowego Internetu Rzeczy (IIoT) zakończony 2 publikacjami naukowymi [D4, D10];
- META (2023-obecnie) – projekt: Metawersum przeciw wykluczeniu społecznemu, wykonawca,
Udział w realizacji badań nad bezpieczeństwem rozwiązań rzeczywistości wirtualnej i rozszerzonej w celu weryfikacji bezpieczeństwa środowiska Metaverse. Badania zakończone publikacją [D15];
- Squaretec Balicki, Kozłowski, Weissenberg, sp.j. - projekt: FaceCOV (podwykonawstwo w ramach projektu finansowanego ze środków NCBR (Program Operacyjny Inteligentny Rozwój 2014-2020) realizowanego przez firmę Milton Essex), wykonawca
Zarządzanie projektem badawczym prowadzącym do opracowania oprogramowania i implementacji algorytmu automatycznego wykrywania biomarkerów temperaturowych w powiązaniu z optymalnym mechanizmem umożliwiającym zautomatyzowane wyszukiwanie i identyfikację pól pomiarowych po stronie firmy Sqaretec oraz realizacja badań nad oprogramowaniem umożliwiającym pomiar temperatury ciała u osoby dorosłej na podstawie obrazu termowizyjnego obszarów ROI-1/ROI-2 opisanych w normie ISO/TR 13154 po stronie Politechniki Poznańskiej. Badania zakończone publikacją [D11] oraz raportem z dzieła "Opracowane i implementacja oprogramowania umożliwiającego pomiar temperatury ciała u osoby dorosłej na podstawie obrazu termowizyjnego obszarów ROI-1/ROI-2 opisanych w w normie ISO/TR 13154";
- Santander (2022-obecnie) – projekt: Szkolenie Santander Cyberbezpieczeństwo z Politechniką Poznańską, wykonawca.
Wsparcie przy przeprowadzeniu kursu w zakresie cyberbezpieczeństwa w ramach projektu;
- Pearson VUE (2016-2020)
Administrator centrum egzaminacyjnego Pearson VUE na Politechnice Poznańskiej.

2.5.2 Współpraca z otoczeniem społecznym

Doktorant brał udział w programie "Lustra" realizowanym przez telewizję TVP Poznań. Tematem wyemitowanego 03.02.2023 roku programu było bezpieczeństwo cyber-

netyczne samorządów w Polsce. W programie oprócz doktoranta brał udział także Pan Michał Łakomski (Pełnomocnik Prezydenta Miasta Poznania ds. Smart City, Kierownik Biura Cyfryzacji i Cyberbezpieczeństwa w Smart City Poznań). Doktorant ponadto brał aktywny udział w organizacji drzwi otwartych na Politechnice Poznańskiej, podczas których oprowadzał uczniów oraz prezentował sale laboratoryjne Instytutu Sieci Telekomunikacyjnych, a także Nocy Naukowców, podczas których prowadził prezentacje z bezpieczeństwa kart płatniczych.

2.6 Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę

2.6.1 Prowadzone zajęcia dydaktyczne

Doktorant od roku 2017 prowadził zajęcia dydaktyczne na studiach I i II stopnia, stacjonarnych i niestacjonarnych Politechniki Poznańskiej. Zajęcia realizowane były w ramach kierunków studiów: Elektronika i Telekomunikacja (studia stacjonarne i niestacjonarne), Electronics and Telecommunications (studia stacjonarne), Teleinformatyka (studia stacjonarne), Informatyka - specjalność cyberbezpieczeństwo (studia stacjonarne). Wśród zajęć dydaktycznych można wyróżnić:

- Koordynacja przedmiotu:
 - Digital Forensics;
- Zajęcia ćwiczeniowe.
 - Systemy Operacyjne;
 - Systemy Wbudowane;
 - Struktury i Działanie Sieci Telekomunikacyjnych;
- Zajęcia laboratoryjne:
 - Sieci Komputerowe;
 - Informatyka;
 - Lokalne Sieci Teleinformatyczne;
 - Systemy Operacyjne;
 - Security of Industrial IoT;
 - Struktury i Działanie Sieci Telekomunikacyjnych.
- Zajęcia projektowe:
 - Traffic Control.

W ramach realizowanych przedmiotów doktorant opracował autorskie materiały dydaktyczne (prezentacje multimedialne, materiały teoretyczne, instrukcje laboratoryjne, zadania, nagrania wideo z omówieniem podejmowanych zagadnień dostępne dla studentów na platformie e-learnigowej).

Jakość oraz atrakcyjność prowadzonych przez doktoranta zajęć dydaktycznych była okresowo oceniana przez studentów Politechniki Poznańskiej w anonimowych ankietach. Od roku 2017 doktorant został oceniony przez ponad 500 studentów uzyskując średnią 4,88, przy maksymalnej ocenie 5.0. Wysokie oceny studentów, które otrzymywał doktorant w tych ankietach stały się podstawą uzyskania nagrody Rektora Politechniki Poznańskiej za szczególne osiągnięcia dydaktyczne w roku akademickim 2022/2023. Dodatkowo doktorant otrzymał nagrodę Rektora Politechniki Poznańskiej za nowoczesne rozwiązania w dydaktyce.

Doktorant ponadto brał udział w opracowaniu przedmiotu IoT Security, realizowanego przy współpracy z Uniwersytetem w Mons.

2.6.2 Wybrane kursy szkoleniowe i inne zajęcia dydaktyczne

Doktorant w ramach dodatkowych aktywności realizowanych na Politechnice Poznańskiej angażował się w proces realizacji kursów dla studentów i zajęć dydaktycznych dla uczniów szkoły średniej, takich jak:

- Akademia sieci Huawei – realizacja kursów dla studentów w latach 2021-2023 w ramach kursów Huawei HCIA Datacom oraz Huawei HCIA AI – ponad 120 przeszkolonych studentów;
- Zespół Szkół Łączności (ZSŁ) w Poznaniu – realizacja zajęć z przedmiotu Informatyka w latach 2020-2022 w ramach współpracy pomiędzy ZSŁ, a Politechniką Poznańską;
- Koordynacja i realizacja zajęć w ramach szkoły letniej "IoT Security" realizowanej na Politechnice Poznańskiej dla studentów Uczelni Partnerskich EUNICE oraz Uniwersytetu w Peru – 27 uczestników.

2.6.3 Odbyte kursy i szkolenia

Doktorant poszerzał także swoją wiedzę biorąc udział w następujących kursach oraz szkoleniach:

- Prince2 Foundation (poświadczenie nr: CR656217449MW);
- SANS Institute: SEC587: Advanced Open-Source Intelligence (OSINT) Gathering and Analysis;
- Check Point (ATC Partner: Monnappa K A): Threat Hunting Using Memory Forensics;
- Check Point (ATC Partner: Payatu Inc.): Android App Hacking;

- Check Point: Security Administration R81.00 (CCSA);
- Palo Alto: Cloud Security Fundamentals;
- Palo Alto: Cybersecurity Foundation;
- Palo Alto: Network Security Fundamentals;
- Palo Alto: Security Operations Fundamentals;
- Cisco: CCNA Routing & Switching;
- Huawei: HCIP Routing & Switching;
- Huawei: HCIA AI;
- Huawei: HCIA Datacom Instructor.

2.7 Załącznik 9 - Oświadczenia współautorów artykułów

Oświadczenie współautorów artykułu

[J1] S. Hanczewski, M. Stasiak, and M. Weissenberg, “*High-accuracy analytical model for heterogeneous cloud systems with limited availability of physical machine resources based on Markov Chain,*” in *Electronics*, no. 13:11:2161.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 10%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Maciej Stasiak (20%)

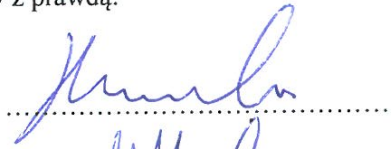
- Konsultacje naukowe związane z opracowaniem koncepcji modelu analitycznego i prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Michał Weissenberg (70%)

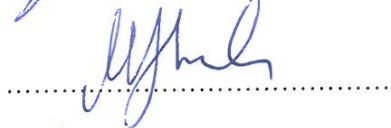
- Opracowanie koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu chmurowego,
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu,
- Nadzór nad procesem wydawniczym.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

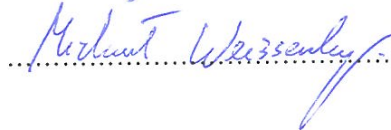
Sławomir Hanczewski



Maciej Stasiak



Michał Weissenberg



Oświadczenie współautorów artykułu

[J2] S. Hanczewski, M. Stasiak, and M. Weissenberg, "An Analytical Model of IaaS Architecture for Determining Resource Utilization," in *Sensors*, no. 9: 2758, 2024.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 10%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu,
- Nadzór nad procesem wydawniczym.

Maciej Stasiak (10%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Michał Weissenberg (80%)

- Opracowanie koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu chmurowego,
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

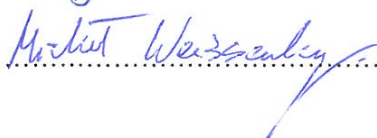
Sławomir Hanczewski



Maciej Stasiak



Michał Weissenberg



Oświadczenie współautorów artykułu

[J3] J. Weissenberg and M. Weissenberg, "Model of a Queuing System With BPP Elastic and Adaptive Traffic," in IEEE Access, vol. 10, pp. 130771-130783, 2022.

Joanna Weissenberg (Szacowany udział w tworzeniu artykułu: 50%)

- Opracowanie koncepcji modelu analitycznego,
- Udział w implementacji modelu analitycznego i symulacyjnego systemu kolejkowego z ruchem BPP typu elastycznego i adaptacyjnego,
- Udział w opracowaniu scenariuszy badawczych oraz opis rezultatów badań,
- Udział w przeprowadzeniu badań z wykorzystaniem zaimplementowanych programów,
- Udział w redakcji i korekcie artykułu,
- Nadzór nad procesem wydawniczym.

Michał Weissenberg (50%)

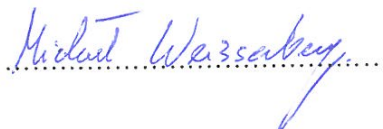
- Udział w opracowaniu koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu kolejkowego z ruchem BPP typu elastycznego i adaptacyjnego,
- Udział w opracowaniu scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań z wykorzystaniem zaimplementowanych programów,
- Wizualizacja wyników
- Udział w redakcji i korekcie artykułu.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

Joanna Weissenberg


.....

Michał Weissenberg


.....

Oświadczenie współautorów artykułu

[J4] S. Hanczewski, M. Stasiak, and M. Weissenberg, "A Multiparameter Analytical Model of the Physical Infrastructure of a Cloud-Based System," in IEEE Access, vol. 9, pp. 100981-100990, 2021.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 10%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Maciej Stasiak (20%)

- Konsultacje naukowe związane z opracowaniem koncepcji modelu analitycznego i prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Michał Weissenberg (70%)

- Opracowanie koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu chmurowego,
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu,
- Nadzór nad procesem wydawniczym.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

Sławomir Hanczewski

Maciej Stasiak

Michał Weissenberg

Oświadczenie współautorów artykułu

[J5] S. Hanczewski, M. Stasiak, and M. Weissenberg, "Modeling of Complex Non-Full-Availability Systems," Journal of Telecommunications and Information Technology (JTIT), no. 2, pp. 70–77, Jun. 2020.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 20%)

- Konsultacje naukowe związane z opracowaniem koncepcji modelu analitycznego i prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu,
- Nadzór nad procesem wydawniczym.

Maciej Stasiak (10%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Michał Weissenberg (70%)

- Opracowanie koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu chmurowego,
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu,
- Nadzór nad procesem wydawniczym.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

Sławomir Hanczewski

Maciej Stasiak

Michał Weissenberg

Oświadczenie współautorów artykułu

[C1] S. Hanczewski and M. Weissenberg, "A Study on The Properties of Complex Non-Full-Availability Systems," in 39th International Business Information Management Association Conference IBIMA 2022, 30-31.05.2022, Granada, Spain.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 20%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Michał Weissenberg (80%)

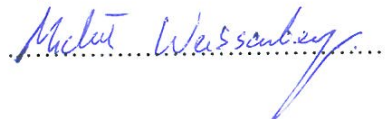
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu,
- Nadzór nad procesem wydawniczym.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

Sławomir Hanczewski



Michał Weissenberg



Oświadczenie współautorów artykułu

[C2] S. Hanczewski, M. Stasiak, and M. Weissenberg, "The Analytical Model of Complex Non-Full-Availability System," in Choraś, M., Choraś, R. (eds) Image Processing and Communications. IP\&C 2019. Advances in Intelligent Systems and Computing, vol 1062. Springer, Cham.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 20%)

- Konsultacje naukowe związane z opracowaniem koncepcji modelu analitycznego i prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu,
- Nadzór nad procesem wydawniczym.

Maciej Stasiak (10%)

- Konsultacje naukowe prac badawczych związanych z przygotowaniem artykułu,
- Udział w redakcji i korekcie artykułu.

Michał Weissenberg (70%)

- Opracowanie koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu chmurowego,
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

Sławomir Hanczewski

Maciej Stasiak

Michał Weissenberg

Oświadczenie współautorów artykułu

[C3] S. Hanczewski and M. Weissenberg, "Concept of an Analytical Model for Cloud Computing Infrastructure," 2018 11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), Budapest, Hungary, 2018, pp. 1-4.

Sławomir Hanczewski (Szacowany udział w tworzeniu artykułu: 40%)

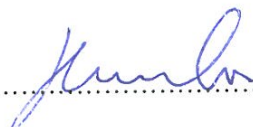
- Opracowanie koncepcji modelu analitycznego,
- Udział w opracowaniu scenariuszy badawczych,
- Udział w redakcji i korekcie artykułu,
- Nadzór nad procesem wydawniczym.

Michał Weissenberg (60%)

- Udział w opracowaniu koncepcji modelu analitycznego,
- Implementacja modelu analitycznego i symulatora systemu chmurowego,
- Opracowanie scenariuszy badawczych oraz opis rezultatów badań,
- Przeprowadzenie badań,
- Redakcja i korekta artykułu.

Niniejszym oświadczam, że przedstawiony podział prac jest zgodny z prawdą:

Sławomir Hanczewski



Michał Weissenberg

