

Warszawa, dnia 23 kwietnia 2023 r.

Prof. dr hab. Oleg Tikhonenko
Uniwersytet Kardynała Stefana Wyszyńskiego
w Warszawie.
Wydział Matematyczno-Przyrodniczy. Szkoła Nauk Ścisłych.
Instytut Informatyki

RECENZJA

Rozprawy Doktorskiej Godlove Suila Kuaban
*Transient Markovian and Diffusion Approximation Models
for Performance Analysis of Computer Networks
and Battery Energy Storage Systems.*

(Modele stanów nieustalonych, oparte na łańcuchach Markowa
i aproksymacji dyfuzyjnej w analizie efektywności
sieci komputerowych i systemów magazynowania energii),
wykonanej pod kierunkiem:
prof. dra hab. inż. Tadeusza Czachórskiego

I. Przedmiot i zakres rozprawy

Celem pracy (w szerokim sensie) jest pokazanie możliwości aproksymacji dyfuzyjnej w zagadnieniach teorii kolejek związanych z analizą charakterystyk systemów kolejkowych w stanach nieustalonych, tj. charakterystyk zależnych od czasu i warunków początkowych działania systemu.

Klasyczna teoria kolejek zajmuje się przeważnie obliczaniem charakterystyk stacjonarnych, co potrzebuje założenia spełnienia warunków istnienia trybu stacjonarnego (stanu ustalonego). Charakterystyki niestacjonarne w postaci jawnej otrzymano tylko dla najprostszych systemów typu $M/M/1$, $M/M/1/N$, $M/G/\infty$. W postaci transformat Laplace'a funkcji tworzącej liczbę obecnych w systemie klientów takie charakterystyki otrzymano dla systemów typu $M/G/1$ (włącznie z systemami priorytetowymi) i systemu z podziałem procesora $M/G/1-EPS$.

Bardzo często powstaje konieczność obliczania charakterystyk niestacjonarnych zależnych od czasu, innymi słowy, potrzebna jest analiza zasadniczo niestacjonarnego zachowania systemów kolejkowych (jeżeli np. należy oszacować wpływy zatorów w sieciach lub czasu rozładowywania baterii na charakterystyki systemu). W realnych zagadnieniach analiza systemów kolejkowych jest ważna zwłaszcza w przypadku długich kolejek i długiego czasu oczekiwania; mniej ciekawy jest przypadek małych kolejek i małych opóźnień, ponieważ wtedy nie ma poważnych trudności w działaniu systemu.

Najczęściej jednak nie istnieje metod dokładnego rozwiązywania podobnych problemów w ramach klasycznej teorii. Przyczyna takiej sytuacji tkwi w „detaliczności” dokładnych metod klasycznych, w których należy uwzględniać wszystkie drobne zmiany stanów: każde przybycie klienta, każde zakończenie obsługi, skąd wynika, że wskazane metody są bardzo czułe na najmniejsze odchylenia od pewnych standardów. Na trudności dokładnej analizy analitycznej istotnie wpływają np. (nawet w przypadku obliczenia charakterystyk stacjonarnych) typy rozkładów odstępów czasu między chwilami przybycia klientów do systemu i czasu obsługi: często dokładne obliczenia są możliwe tylko dla rozkładów wykładniczych. A czasem nawet modelowanie markowskich sieci kolejkowych z dużą liczbą systemów typu $M/M/1$ staje się niemożliwe za pomocą metod klasycznych.

Modelowanie symulacyjne niestety również nie może znacząco pomóc w tej sytuacji, ponieważ dla osiągnięcia potrzebnej dokładności (jeżeli np. chcemy zbudować przedział ufności dla pewnej charakterystyki liczbowej) należy wiele razy losować zachowanie systemu niestacjonarnego w tym samym przedziale czasowym. Oprócz tego podejście symulacyjne staje się bardzo nieefektywne przy oszacowaniu prawdopodobieństw rzadkich zdarzeń, takich jak np. utrata pakietu w węźle sieci telekomunikacyjnej.

Wyjściem z opisanej sytuacji jest korzystanie ze znanych metod przybliżonych, takich jak aproksymacja płynna i aproksymacja dyfuzyjna. Metoda aproksymacji płynnej w ogóle działa na poziomie wartości średnich, tj. jest praktycznie modelem deterministycznym. Modele dyfuzyjne uwzględniają także wariancję odstępów między chwilami przybycia klientów i czasu obsługi, oni są bardziej dokładne w porównaniu z modelami aproksymacji płynnej.

Podstawą metody aproksymacji dyfuzyjnej jest ten fakt wynikający z twierdzenia granicznego teorii prawdopodobieństwa, że w warunkach „długiej” kolejki, kiedy bez przerwy trwa obsługa klientów, rozkład liczby klientów obecnych w systemie w ustalonej chwili czasu ma w przybliżeniu rozkład normalny niezależnie od typu rozkładów wskazanych odstępów i czasu obsługi. W takich warunkach, jak się okazuje, proces dyskretny opisujący liczbę klientów w systemie można aproksymować procesem ciągłym dyfuzji, którego gęstość spełnia tzw. równanie dyfuzji (równanie (1.10) na s. 17).

Dokładny cel pracy podano na s. 10. Jest nim pokazanie tego, że aproksymacja dyfuzyjna jest lepsza w porównaniu z innymi metodami modelowania (takimi jak aproksymacja płynna i łańcuchy Markowa) w stosowaniu do analizy systemów kolejkowych działających w trybie niestacjonarnym. Pokazano to na przykładach: 1) obliczania jakości transmisji w sieciach z dynamicznymi protokołami marszrutyzacji; 2) wyznaczenia opóźnień spowodowanych agregacją pakietów na granicy dwu typów sieci; 3) analizy procesu wyczerpania się energii w systemach jej magazynowania w węzłach sieci komputerowych.

Uważam, że temat rozprawy jest aktualny, realizacja badań odpowiadających jej celu jest ważnym osiągnięciem naukowym mającym wiele zastosowań w praktyce projektowania i analizy systemów komputerowych i sieci telekomunikacyjnych.

II. Opis struktury rozprawy

Rozprawa składa się ze streszczenia (Abstract), wstępu (Chapter 1), pięciu rozdziałów (Chapters 2–6), podsumowania (Chapter 7) i listy publikacji.

We Wstępie (chapter 1) podano motywację i cel pracy, dokonano przeglądu różnych metod teorii kolejek, podano krótki opis problemów rozwiązywanych w dalszych rozdziałach pracy.

W drugim rozdziale (chapter 2) Autor przedstawia podejście do modelowania mechanizmów agregacji pakietów oparte na aproksymacji dyfuzyjnej. Rozważa się tu model sieci, która składa się z sieci dostępowej (linie abonenckie, sieci lokalne, sieci bezprzewodowe, sieci mobilne, sieci dostępne Internetu itd.), sieci rdzeniowej, przez którą pakiety z sieci dostępowej są transportowane do użytkowników innych sieci dostępowych lub centrów danych, i właśnie z centrów danych.

W pracy szczegółowo opisano działanie takiej sieci. Ponieważ w sieci dostępowej generowano pakiety zasadniczo różnych typów (bardzo krótkich i bardzo długich), to pakiety użytkowników są agregowane w większe pakiety o podobnej długości, co dobrze wpływa na wydajność sieci i jej wskaźniki ekonomiczne. Różne charakterystyki rozkładów długości pakietów i prędkości ich transmisji lub przetwarzania w różnych elementach sieci wymaga ich kolejkowania w buforach o ograniczonej pojemności. W tym rozdziale pracy modelowano trzy typy takiego kolejkowania. 1) Małe pakiety w buforze agregowano w jeden duży pakiet w ciągu ustalonego czasu T , po czym ten duży pakiet wysyłano dalej. W tym przypadku pakiety o różnej długości wysyłano regularnie. 2) Małe pakiety agregowano dopóki nie wypełniono całego buforu. Wtedy wszystkie pakiety mają długości bliskie pojemności bufora N bajtów, natomiast są oni wysyłane w losowych odstępach czasu. 3) Ten typ agregacji jest mieszaniną pierwszych dwóch typów. Jest ona analizowana dla przypadku agregacji pakietów w sieci N-GREEN.

W pracy zbadano wszystkie trzy typy agregacji pakietów. Otrzymane wyniki porównano z wynikami symulacji. Takie porównanie służy dowodem słuszności podejścia opartego na aproksymacji dyfuzyjnej. Wyniki otrzymane w tym rozdziale są dobrze ilustrowane za pomocą rysunków i wykresów. Wyniki dotyczą głównie rozkładów rozmiarów zagregowanych pakietów, czasów agregowania, czasów czekania w buforze, wydajności sieci.

W trzecim rozdziale (chapter 3) przedstawiono model przełącznika sieci SDN (Software Defined Network). SDN posiada kontroler zawierający programowe realizacje algorytmów sterowania transmisją danych. Kontroler steruje ruterem, który przekazuje pakiety zgodnie z założonymi kryteriami optymalizacji. Proces przekazania zależy więc od stanu sieci w danej chwili czasu, co jest uwzględniono w algorytmach kontrolera. Na proces wpływają np. stan bezpieczeństwa sieci, zużycie energii, parametry QoS (Quality of Service) itd. Częste nadejście pakietów powoduje sytuację stałego działania systemu w trybie niestacjonarnym. Podczas odpowiedniego przetwarzania pakiet oczekuje na transmisję w kolejce bufora pakietów.

W pracy dla analizy SDN Autor korzysta z modelu jednoliniowego systemu kolejkowego typu $G/G/1/N$ (N jest ograniczeniem liczby pakietów w buforze), w którym czas obsługi jest stały (równy T) z prawdopodobieństwem p oraz z prawdopodobieństwem $1 - p$ jest dyskretny jednostajny na pewnym przedziale $[T; KT]$. Takie założenie jest potrzebne dla obliczania odpowiednich wartości oczekiwanej i wariancji czasu obsługi, co umożliwia stosowanie aproksymacji dyfuzyjnej. W pracy analizowano dwa rodzaje modelu dyfuzyjnego: model, w którym nadejście pakietów tworzy strumień stacjonarny z intensywnością $\lambda > 0$, oraz model, w którym intensywność jest zmienna (równa $\lambda(t)$). Ostatnie założenie prowadzi do równania dyfuzji ze zmiennymi (zależnymi od czasu) współczynnikami.

W pracy zbadano oba typy modelu. Wynikami są aproksymacja rozkładu liczby pakietów w systemie, przybliżona wartość prawdopodobieństwa utraty pakietu (przepełnienia buforu) i rozkład czasu oczekiwania w kolejce w trybie stacjonarnym (rys. 3.5, 3.6 i 3.7 na s. 73, 74 i 75 odpowiednio). W przypadku zmiennej w czasie intensywności nadejścia pakietów przeanalizowano wpływ zmiany intensywności λ na średnią długość kolejki (rys. 3.8 na s. 77), wpływ zmiany obciążenia systemu $\rho = \lambda/\mu$ na średni czas oczekiwania pakietu (rys. 3.9 na s. 78) oraz wpływ zmiany intensywności λ na zachowanie w czasie średniej długości kolejki. Prowadzone w tym rozdziale badania umożliwiają więc przewidywanie zależnego od czasu zachowania ważnych wskaźników wydajności, takich jak średnie opóźnienie pakietu na routerze, długość kolejki pakietów oraz prawdopodobieństwo utraty pakietu.

W czwartym rozdziale (chapter 4) Autor prezentuje model pracy sieci przełączników SDN. Sieć takiego typu składa się z oddzielnych systemów, tak że jeżeli pakiet w chwili czasu t opuszcza system o numerze i , to on skieruje się do systemu o numerze j z prawdopodobieństwem $r_{ij}(t)$. Prawdopodobieństwa $r_{ij}(t)$ pozostają bez zmian na pewnym przedziale czasowym Δ , chociaż intensywności strumieni pakietów mogą się zmieniać na mniejszych przedziałach $\delta < \Delta$. W modelu przyjęto założenie, że $\Delta = n\delta$, gdzie n może być wybrane w zależności od konkretnych warunków działania systemu.

Za pomocą modelu aproksymacji dyfuzyjnej Autor wyznacza w przybliżeniu rozkłady liczby pakietów w każdym systemie (węźle) sieci dla każdego przedziału δ , rozkład czasu pobytu (czas oczekiwania plus czas obsługi) pakietu w każdym systemie oraz prawdopodobieństwo utraty pakietu.

Otrzymane w tym rozdziale rezultaty wskazują, że proponowana metoda pozwala na uzyskanie korzystnych wyników ilościowych dla modeli sieci SDN o realistycznych wartościach parametrów. Podane przykłady liczbowe ujawniają charakter zależności długości kolejki, opóźnień kolejkowania i ich dynamiki od zmieniającej się intensywności przybycia pakietów i charakterystyk odstępów czasu między chwilami przybycia. Proponowany model pokazuje, że okresy przejściowe odgrywają istotną rolę w działaniu sieci SDN, i mogą być wykorzystywane dla analizy „dużych” sieci.

W piątym rozdziale (chapter 5) przedstawiono modele zużycia energii w systemach jej magazynowania oraz ataków na urządzenia Internetu Rzeczy w celu pozbawieniu ich energii.

Na początku Autor podaje szczegółowy opis analizowanych procesów związanych z zużyciem energii i atakami na urządzenia. W dalszym ciągu podano model, który opisuje odpowiedni procesu dyfuzji na przedziale $[0; B]$, gdzie B jest pojemnością baterii. Czas

zużycia definiuje się jako czas potrzebny do wyczerpania zgromadzonej w niej energii. Model dyfuzyjny daje rozkład czasu zużycia baterii (wzór (5.36) na s. 111). Wynik ten porównano z rezultatem otrzymanym za pomocą techniki łańcuchów Markowa. Jak pokazano na rys. 5.7 (s. 112), oba podejścia dają prawie ten sam wynik.

W modelu dyfuzyjnym oblicza się również prawdopodobieństwo tego, że bateria jest zupełnie zużyta w ustalonej chwili czasu t (wzór (5.41) na s. 113). Przeanalizowano także przypadek szczególny, dla którego prawdopodobieństwo tego, że bateria nigdy nie rozładuje się jest większe od zera. Jest to możliwe, jeżeli $\lambda > \mu$, gdzie λ jest intensywność poboru energii, a μ – intensywność jej zużycia. Zbadano również zależność wartości wskazanego prawdopodobieństwa od wielkości $\rho = \lambda/\mu$.

W tym rozdziale prowadzono również (na przykładach obliczeniowych) badania wpływów ataków polegających na wyczerpywaniu baterii. Wskazane ataki manipulują urządzeniem Internetu Rzeczy w celu znacznego zwiększenia zużycia energii. Podane przykłady umożliwiają zbadanie wpływu ataków polegających na wyczerpywaniu baterii na rozkład czasu życia urządzenia Internetu Rzeczy.

W modelu możliwe jest wprowadzenie progu energetycznego, po przekroczeniu którego należy wymienić baterię urządzenia, aby mieć pewność, że bateria nie rozładuje się całkowicie przed jej wymianą. Czas, po którym należy wymienić baterię, także może być oszacowany za pomocą modelu. Przybliżenie dyfuzyjne można więc wykorzystać do wygodnego modelowania procesu wyczerpywania energii baterii urządzenia Internetu Rzeczy. Znając pojemność baterii, średnie zużycie energii i jego wariancję można wyznaczyć gęstość, wartość średnią, wariancję „czasu życia” baterii.

W szóstym rozdziale (chapter 6) prowadzono modelowanie wydajności systemu magazynowania energii baterii dla stacji bazowej sieci komórkowej.

W rozdziale tym przeanalizowano wydajność systemu BESS (Battery Energy Storage System), która jest stale ładowana przez odnawialne źródła energii i rozładowywana w procesie zasilania stacji bazowej. Opisano architekturę stacji bazowej (Green Mobile Network Base Station). W celu analizy wydajności akumulatorowego systemu magazynowania energii w stanach ustalonym i przejściowym Autor korzysta z modeli Markowa i aproksymacji dyfuzyjnej. Za ich pomocą wyznaczono czas, po jakim akumulatorowy system magazynowania energii jest całkowicie rozładowany lub w pełni naładowany.

Prowadzono porównanie wyników otrzymanych w obu modelach przy założeniu, że czasy pozyskiwania i zużycia energii mają rozkład wykładniczy. Okazało się, że otrzymane wyniki są w obu modelach takie same.

W Podsumowaniu (chapter 7) Autor wnioskuje o wartości poszczególnych modeli analitycznych stosowanych w prowadzonych badaniach i wskazuje dalsze kierunki badań związanych z tematem rozprawy. W rozdziale tym Autor przelicza także oryginalne wyniki swojej pracy.

Bibliografia pracy zawiera 261 tytułów, wśród których znajdują się także publikacje z udziałem Autora rozprawy

III. Ocena ogólna rozprawy

Przedstawiona do recenzji Rozprawa Doktorska jest kompletną pracą naukową. Badania przeprowadzone w rozdziałach 2–6 potwierdzają, iż cel Rozprawy został zrealizowany. Przedstawione i opracowane przez Autora modele analityczne (markowskie i dyfuzyjne) są skutecznymi narzędziami badania sieci komputerowych i telekomunikacyjnych, które pozwoliły na adekwatne opisanie procesów w systemach i sieciach kolejkowych, co prowadzi do dobrego ich zrozumienia, daje możliwość dostatecznie dokładnego oszacowania charakterystyk oraz porównania wartości stosowanych modeli analitycznych.

Wyniki prowadzonych badań opublikowano w 11 artykułach w wysoko punktowanych czasopismach międzynarodowych (3) i zbiorach prac prestiżowych konferencji (8). Oprócz artykułów wykonanych w procesie pracy nad Rozprawą Dokorską opublikowano 11 innych artykułów w czasopismach i zbiorach prac konferencji, wśród których większa część są publikacje wysoko punktowane.

Wysokie kompetencje naukowe Autora potwierdza jego udokumentowane zwycięstwo w konkursie na najlepszego studenta zagranicznego w Polsce w kategorii „studia doktoranckie”.

Stwierdzam, że wartość naukowa przedstawionych w pracy wyników jest wysoka i odpowiednia do uzyskania stopnia naukowego doktora nauk technicznych w dyscyplinie „Informatyka Techniczna i Telekomunikacja” oraz, że praca nie wymaga uzupełnień, które byłyby istotne merytorycznie dla jej treści.

IV. Uwagi szczegółowe

Strona edytorska pracy nie budzi zastrzeżeń. Szata graficzna pracy jest przejrzysta. Tekst napisano językiem klarownym i profesjonalnym. Materiał graficzny (rysunki i wykresy) jest ściśle powiązany z materiałem merytorycznym. Tekst Rozprawy jest bardzo dobrze opracowany. Mogę tylko wskazać na niektóre drobne niedoskonałości.

1. Ciężko zrozumieć czym jest „probability of the lifetime of an IoT device” (s. 120). Ponieważ „lifetime” jest zmienną losową, to chodzi tu chyba o jej rozkład.
2. Numeracja rysunków w rozdziale 6 nie odpowiada linkom do nich w tekście, co utrudnia lekturę rozprawy.
3. Dla ułatwienia lektury rozprawy należałoby także umieścić w niej tabelę użytych skrótów.
4. Prawdopodobieństwo utraty pakietu w opracowanych modelach zależy oczywiście od jego rozmiaru, co w rozprawie nie bierze się pod uwagę. Myślę, że jest to wnioskiem niedoskonałości metody aproksymacji dyfuzyjnej (każda metoda ma swoje wady).

Powyższe uwagi nie są istotne i pozostają bez znaczącego wpływu na moją pozytywną merytoryczną ocenę Rozprawy. Podobnie, przytoczone uwagi nie podważają ważnych naukowych wyników pracy. W mojej opinii praca odpowiada wymaganiom stawianym Rozprawom Doktorskim.

V. PODSUMOWANIE

Wyniki rozważań przedstawione w pracy upoważniają do stwierdzenia, że zarówno cel główny pracy, jak i wszystkie cele pośrednie, zostały osiągnięte. Przedstawiony materiał świadczy o wniesieniu do badanego obszaru (badanie systemów i sieci komputerowych za pomocą metod modelowania dyfuzyjnego i łańcuchów Markowa) wielu elementów nowości i o ich pozytywnym zweryfikowaniu. Autor dowiódł, iż swobodnie porusza się po zagadnieniach związanych z teorią kolejek i jej zastosowaniami w analizie systemów i sieci telekomunikacyjnych, krytycznie podchodzi do metod i wyników zawartych w literaturze naukowej oraz twórczą je uzupełnia i rozwija.

Na podstawie przytoczonych faktów stwierdzam, że opiniowana Rozprawa Doktorska „*Transient Markovian and Diffusion Approximation Models for Performance Analysis of Computer Networks and Battery Energy Storage Systems*” spełnia wymagania Ustawy o tytule naukowym i stopniach naukowych i wnioskuję o dopuszczenie mgr inż. Godlove Suila Kuaban do dalszych etapów przewodu doktorskiego.

Oleg Tikhonenko

Uważam, że duża waga naukowa recenzowanej Rozprawy Doktorskiej oraz wysoki poziom naukowy publikacji związanych z jej tematem powoduje pytanie o jej wyróżnienie.

Ja, niżej podpisany recenzent prof. dr hab. Oleg Tikhonenko, wnioskuję o wyróżnienie Rozprawy Doktorskiej „*Transient Markovian and Diffusion Approximation Models for Performance Analysis of Computer Networks and Battery Energy Storage Systems*”