

Recenzja rozprawy doktorskiej

Prof. dr hab inż. Andrzej Chydziański
Wydział Automatyki, Elektroniki i Informatyki
Politechnika Śląska

Gliwice, 19 kwietnia 2016

Tytuł rozprawy: **Router interface as a queuing system with correlated service times**

Autor rozprawy: **mgr inż. Krzysztof Rusek**

Promotor rozprawy: **prof. dr hab. inż. Zdzisław Papir**

Promotor pomocniczy: **dr inż. Lucjan Janowski**

1 Zawartość rozprawy

Rozprawa poświęcona jest modelowaniu kolejek pakietów występujących w interfejsach routerów. Głównym celem było doprecyzowanie istniejących modeli w celu uzyskania dokładniejszych prognozowanych charakterystyk procesu strat pakietów, w szczególności współczynnika strat, czasu do przepełnienia bufora oraz chwilowych intensywności strat.

Rozprawa ma charakter eksperymentalno-teoretyczny. W części eksperymentalnej badano w laboratorium urządzenia sieciowe w celu rozstrzygnięcia, czy bufony interfejsów są wymiarowane w bajtach, czy w pakietach, co następnie pozwoliło na przyjęcie odpowiednich założeń dotyczących modelu systemu kolejkowego. W części teoretycznej podano rozwiązanie w stanie nieustalonym oraz ustalonym modelu kolejki uwzględniającego uzyskany wcześniej typ bufora, strumień wejściowy z autokorelacją czasów pomiędzy pakietami, oraz skorelowane czasy transmisji pakietów.

W rozprawie przedstawiono ponadto przykładowe wyniki obliczeń numerycznych dla rozważanego modelu, które porównano z wynikami uzyskanymi

dla wcześniejszych modeli oraz z wynikami dokładnymi, otrzymanymi dzięki symulacjom ze śladami rzeczywistego ruchu. W zdecydowanej większości przypadków otrzymano znaczącą poprawę dokładności wyników.

Poprawa ta jest konsekwencją zastosowania lepszego modelu bufora oraz uwzględnienia korelacji czasów transmisji pakietów, która rzeczywiście występuje w ruchu sieciowym.

Rozprawa zawiera pięć rozdziałów głównych oraz cztery rozdziały dodatkowe, tzn. wstęp, wnioski, oraz rozdziały dotyczące dokładności wyników eksperymentalnych i włożonych łańcuchów Markowa. Ponadto w pracy zawarto streszczenie oraz spisy: literatury, rysunków, tablic i symboli matematycznych. Rozprawa napisana jest w języku angielskim.

Rozdział pierwszy zawiera opis przeprowadzonych przez autora eksperymentów z urządzeniami sieciowymi. Eksperymenty te miały na celu zbadanie architektury pamięci buforów w celu zaproponowania poprawnego modelu tych buforów. W wyniku eksperymentów okazało się, że bufor routera powinien być modelowany jako poczekalnia mogąca pomieścić określoną liczbę pakietów, niezależnie od ich długości, a nie tak jak np. pamięć operacyjna komputera, mogąca pomieścić określoną liczbę bajtów.

W rozdziale drugim omówiono literaturę przedmiotu. W szczególności przedstawiono różne niekorzystne własności ruchu pakietów w sieciach komputerowych (spiętrzenia, autokorelację, dalekosieźność, samopodobieństwo), oraz najważniejsze modele ruchu pozwalające te własności uwzględnić. Najwięcej uwagi poświęcono modelom markowskim. W końcowej części rozdziału omówiono literaturę poświęconą dopasowaniu parametrów modelu do zapisanych śladów ruchu sieciowego.

W trzecim rozdziale sformalizowano model kolejkowy interfejsu routera. Zaproponowano zastosowanie modelu typu MAP/SM/1/b w notacji Kendalla, tzn. z markowskim strumieniem zgłoszeń (MAP), półmarkowskim procesem obsługi (SM) oraz buforem wymiarowanym w pakietach. Model ten uwzględnia autokorelację strumienia wejściowego oraz korelację rozmiarów pakietów. Przedstawiono argumenty, dlaczego model tego typu jest najlepszy dla interfejsu routera.

W rozdziale czwartym podano rozwiązania analityczne modelu kolejki MAP/SM/1/b w stanie ustalonym i nieustalonym. W szczególności, przedstawiono najpierw metodę potencjału, która była wcześniej wykorzystywana do rozwiązywania modeli z nieskorelowanymi czasami obsługi, a którą autor uogólnia na przypadek badanego modelu interfejsu routera. Następnie wyznaczono rozkład czasu do przepełnienia bufora (twierdzenie 1) oraz średnią liczbę strat pakietów w przedziale czasu $(0, t]$ (twierdzenie 2). Pokazano też, jak przy pomocy twierdzenia 2 obliczać współczynnik strat pakietów w stanie ustalonym. W końcowej części rozdziału dyskutowane są aspekty numeryczne

wyznaczania charakterystyk kolejkowych przy pomocy twierdzeń 1 i 2.

Rozdział piąty zawiera wyniki numeryczne. W pierwszej części przedstawione są parametryzacje ruchu uzyskane na podstawie dwóch śladów: klasycznego śladu Bellcore oraz śladu ruchu z domów studenckich Akademii Górniczo-Hutniczej. Następnie przedstawiono szereg charakterystyk procesu strat w stanie ustalonym dla tych dwóch śladów. Za każdym razem porównywano charakterystykę uzyskaną dla dokładnego śladu, charakterystykę uzyskaną dla wcześniej używanych modeli kolejkowych (MAP/G/1/b) oraz charakterystykę uzyskaną dla modelu MAP/SM/G/1/b. W dalszej części przedstawiono interesującą analizę wpływu rozmiaru bufora i autokorelacji ruchu na dokładność wyników. Wreszcie, w końcowej części rozdziału, podano kilka przykładów charakterystyk procesu strat w stanie nieustalonym.

2 Opinia o rozprawie

Po pierwsze należy podkreślić, że rozprawa dotyczy ważnej problematyki oceny wydajności sieci przy pomocy modeli kolejek pakietów występujących w węzłach. Modele takie były już proponowane i rozwiązywane przez wielu naukowców, jednak na tle wcześniejszych prac oceniana rozprawa wyróżnia się pozytywnie silnym powiązaniem założeń modelu z wynikami przeprowadzonych eksperymentów z urządzeniami sieciowymi. W wielu poprzednich pracach założenia dotyczące modelu kolejkowego były formułowane na podstawie ogólnych przesłanek dotyczących konstrukcji sieci pakietowych, stosowanych urządzeń oraz protokołów, jednak bez dogłębnej weryfikacji tych założeń w eksperymentach laboratoryjnych.

Za główne osiągnięcia przedstawione w rozprawie uznaję:

- Zaproponowanie i przeprowadzenie eksperymentów laboratoryjnych mających rozstrzygnąć, jaki model bufora najlepiej odwzorowuje konstrukcję interfejsu routera.
- Zaproponowanie do modelowania interfejsu routera modelu kolejkowego MAP/SM/1/b, który pozwala na poprawę dokładności uzyskiwanych charakterystyk dzięki odpowiednim założeniom dotyczącym bufora oraz uwzględnianiu korelacji czasów transmisji pakietów.
- Uogólnienie metody potencjału służącej do rozwiązywania systemów kolejkowych na przypadek z półmarkowskim procesem obsługi.
- Uzyskanie przy pomocy opracowanej metody charakterystyk procesu strat dla systemu kolejkowego MAP/SM/1/b w stanie nieustalonym i ustalonym.

Jak pokazały porównania z wynikami symulacji, rozwiązany model kolejkowy pozwala na uzyskiwanie znacznie dokładniejszych charakterystyk procesu strat, niż modele wcześniej stosowane.

Dodatkowym ważnym osiągnięciem jest zaobserwowanie zależności dokładności wyników od kombinacji rozmiaru bufora i dopasowania autokorelacji strumienia wejściowego (sekcja 5.2.3). Autor zauważył, że precyzyjne modelowanie procesu strat wymaga tym dokładniejszego dopasowania funkcji autokorelacji, im większy jest bufor. A zatem duże bufony wymagają bardzo precyzyjnej funkcji autokorelacji, zaś małe – jedynie jej zgrubnego przybliżenia.

Rozprawa prezentuje wysoki poziom naukowy i zasługuje na bardzo wysoką ocenę z następujących powodów.

Tematyka, którą zajął się autor była już badana przez wielu naukowców, zatem uzyskanie istotnego postępu nie było łatwe. Rozpatrywane modele kolejkowe są bardzo złożone i wymagają zaawansowanego aparatu pojęciowego i matematycznego. Z kolei przeprowadzenie opisanych badań laboratoryjnych wymagało od autora ugruntowanych umiejętności praktycznych.

Nie natknąłem się w pracy na żadne istotne błędy merytoryczne, zarówno w części opisowej, jak i w części matematycznej, tzn. w przedstawionych twierdzeniach i ich dowodach.

Rozprawa jest starannie zredagowana, napisana w sposób jasny i dobrym językiem angielskim. Potknięcia redakcyjne, które wymieniam w następnej sekcji, są bardzo nieliczne. Układ treści pracy, przedstawione rysunki, tabele, rozdziały dodatkowe i spisy są przygotowane zgodnie z zasadami sztuki.

Oczywiste jest, że doktorant swobodnie porusza się we wszystkich obszarach telekomunikacji, których dotyczy rozprawa – widać, że posiada wiedzę szeroką i ugruntowaną.

Przedstawiono dobrą, starannie wyselekcjonowaną bibliografię, bezpośrednio dotyczącą tematyki rozprawy. Przedstawiona lista nie jest oczywiście wyczerpująca, gdyż być nie mogła. Pełna lista prac poświęconych modelom kolejek z markowskimi modelami ruchu i ich zastosowaniami zajmowałaby kilkadziesiąt stron.

Nie mam wątpliwości, że teza rozprawy, która brzmi¹:

Interfejs urządzenia sieciowego z buforem mieszczącym pewną liczbę pakietów niezależnie od ich wielkości może być modelowany przez system kolejkowy MAP/SM/1/b ze skorelowanymi czasami obsługi. Możliwe jest analityczne rozwiązanie tego modelu oraz obliczenie wybranych charakterystyk kolejki w stanie nieustalonym oraz ustalonym.

¹tłumaczenie recenzenta

została prawidłowo sformułowana i wykazana w treści pracy.

3 Uwagi krytyczne i dyskusyjne

Najważniejsze uwagi krytyczne dotyczące rozprawy można zawrzeć w czterech punktach.

- Zabrakło w rozprawie analizy długości kolejki w modelu MAP/SM/1/b, szczególnie w stanie nieustalonym. Opóźnienie kolejkowania wynikające z rozmiaru kolejki jest bardzo ważną charakterystyką wydajności węzła, wpływającą na wydajność całej sieci. Wydaje się, że wykorzystując metodę potencjału można było obliczyć długość kolejki w bardzo podobny sposób, jak charakterystyki procesu strat, uzyskując wyniki zarówno dla stanu ustalonego, jak i nieustalonego. Mam nadzieję, że doktorant będzie kontynuował badania w tym kierunku.
- W rozdziale pierwszym porównywane są charakterystyki strat pakietów uzyskane w symulatorze z charakterystykami uzyskanymi w eksperymentach z routerami i komputerami PC służącymi jako generatory i analizatory ruchu. Brakuje jednak pogłębionej dyskusji potencjalnych błędów spowodowanych wykorzystaniem komputerów PC w roli generatorów i analizatorów. Skąd autor wie, z jaką dokładnością odstępów czasów były wysyłane pakiety z generatora? Dokładność ta może (choć nie musi) być w generatorze PC niska, co zależy od wielu szczegółów implementacyjnych generatora oraz systemu operacyjnego.
- Dopasowanie autokorelacji półmarkowskiego modelu obsługi do śladu Bellcore nie jest najlepsze (zobacz rysunek 5.1 (c)). Czym to było spowodowane i czy mogło wpłynąć istotnie na dokładność wyników prezentowanych na stronie 53?
- Autor uzyskał przy pomocy modelu MAP/SM/1/b znacznie dokładniejsze wyniki niż te, które otrzymywano przy pomocy modeli z nieskorelowanymi czasami obsługi. Wciąż jednak wyniki te są mało dokładne dla większych rozmiarów buforów. Zabrakło w pracy szerszej dyskusji przyczyn tych niedokładności. Czy kluczowe jest tu dobre dopasowanie parametrów procesu półmarkowskiego? Czy może bardziej istotna jest korelacja odstępów czasów z rozmiarami pakietów, której nie uwzględnia model badany w rozprawie? Nie jest też jasne, jaki wpływ na dokładność wyników może mieć jakość dopasowania rozkładu brzegowego czasu pomiędzy pakietami, czy choćby jego wariancji. Autor nie podaje w rozprawie rozkładów brzegowych, ani ich wariancji.

Inne, mniej istotne uwagi:

- Na stronie 21 autor wymienia dwie przyczyny występowania spiętrzeń w ruchu, które tak naprawdę nie są przyczynami, tylko własnościami probabilistycznymi ruchu. Faktyczne przyczyny, wynikające w uwarunkowań technicznych, głównie konstrukcji protokołów sieciowych, można znaleźć np. w pracy H. Jiang et. al., *Source-Level IP Packet Bursts: Causes and Effects*, Proc. ACM SIGCOMM 2003.
- Brak wielu szczegółów dotyczących wykorzystanego śladu ruchu AGH, tzn. jaka była średnia intensywność ruchu, rozmiary pakietów, rozkład czasu pomiędzy pakietami, itd.
- Na rysunkach 5.2-5.2 autor użył dystrybuant do porównania dwóch modeli i symulacji. Do celów poglądowych znacznie lepiej stosować wykresy gęstości rozkładu prawdopodobieństwa, a nie dystrybuanty, które są mniej czytelne.
- Kilka zauważonych błędów redakcyjnych:
 - (a) niekonsekwentne używanie angielskich końcówek „zation” i „sation”. Czasami autor pisze: „uniformization”, „parameterization”, „clusterization”, zaś w innych miejscach „utilisation”, „parametrisation”, „organisation”;
 - (b) wielokrotnie powtarzany błąd „lenghts”;
 - (c) piszemy „markowski”, a nie „Markowowski”;
 - (d) piszemy „półmarkowski”, a nie „semi-Markowowski”;
 - (e) str. 56: powinno być „The relation (...) exists (...)” zamiast „exist”.

4 Wnioski

Żadna z wymienionych wyżej uwag nie obniża w sposób znaczący wysokiej oceny rozprawy wyrażonej w drugiej części recenzji. Doktorant przedstawił oryginalne rozwiązanie problemu dotyczącego modelowania i oceny wydajności sieci telekomunikacyjnych, co wymagało kreatywności naukowej oraz zaawansowanej wiedzy w zakresie telekomunikacji i metod matematycznych.

Uważam, że rozprawa mgra Krzysztofa Ruska w pełni spełnia wymagania stawiane rozprawom doktorskim i wnioskuję o dopuszczenie do jej publicznej obrony.