wpł.

Recenzja rozprawy doktorskiej

Imię i nazwisko kandydatki: Monika Nycz

Tytuł rozprawy doktorskiej:

Modelowanie dynamiki transmisji internetowych za pomocą aproksymacji płynnej.

Promotor: **prof. dr hab. inż. Tadeusz Czachórski**

Recenzent: **dr hab. inż. prof. IITIS PAN Krzysztof Grochla**

1. Wybór tematu i cel pracy

Rozprawa zawiera kompleksowe podejście do problemu modelowania transmisji w sieciach internetowych za pomocą metody aproksymacji płynnej. Doktorantka rozważa problematykę dokładności modeli analitycznych, porównując model aproksymacji płynnej z modelem dyfuzyjnym i wynikami symulacji. Analizuje także możliwość zastosowania aproksymacji płynnej do modelowania bardzo dużych sieci, złożonych z milionów połączeń. Istotnym wkładem rozprawy jest zastosowanie przetwarzania współbieżnego i zrównoleglenie przetwarzania w celu umożliwienia modelowania sieci o większych rozmiarach. Modele oparte o aproksymacje płynną opisane w literaturze dotyczyły znacznie mniej złożonych przypadków, w których czas obliczeń numerycznych był stosunkowo krótki, a ilość przetwarzanych danych mieściła się w pamięci operacyjnej komputera. Doktorantka w sposób twórczy rozwinęła sposób przetwarzania modeli aproksymacji płynnej z wykorzystaniem baz danych, co pozwoliło na zwiększenie skali badanych sieci co najmniej 2 rzędy wielkości.

Temat badań podjętych w rozprawie jest aktualny, a przeprowadzone prace wykraczają poza stan wiedzy w dziedzinie. Modelowanie złożonych sieci jest istotnym wyzwaniem badań nad analizą wydajności systemów komputerowych, ponieważ wraz z rosnącą liczbą urządzeń podłączonych do sieci Internet rośnie złożoność modeli. Badanie funkcjonowania protokołów sieciowych w skali milionów lub setek tysięcy urządzeń jest bardzo istotnym problemem badawczym, co zostało dobrze umotywowane w pracy. Dlatego opracowanie przez doktorantkę nowych metod obliczeniowych i twórczy rozwój znanych metod modelowania sieci pozwalający na zwiększenie skalowalności uruchamianych modeli dobrze wpisuje się w wyzwania badawcze w telekomunikacji.

2. Ocena układu rozprawy

Rozprawa składa się z 7 rozdziałów, spisu osiągnięć naukowych i bibliografii. W rozdziale 1 zawarto krótkie wprowadzenie, omówienie celu przeprowadzonych badań oraz tezy pracy. W rozdziale 2 zawarto porównanie dokładności odwzorowania działania mechanizmu unikania przeciążeń protokołu TCP przez modele aproksymacji płynnej, aproksymacji dyfuzyjnej i symulację zdarzeń dyskretnych. W rozdziale 3 przedstawiono metody generowania topologii sieci komputerowych na potrzeby prowadzonych analiz oraz opisano scenariusze testowe. W rozdziałach 4 i 5 przedstawiono algorytmy budowy modeli aproksymacji płynnej w sposób klasyczny i z wykorzystaniem bazy danych. W rozdziale 6 doktorantka zaproponowała rozszerzenie modelu aproksymacji płynnej o możliwość zmiany parametrów transmisji w trakcie tej transmisji. Pracę zakończono wnioskami zawartymi w rozdziale 7.

Podział rozprawy na rozdziały jest poprawny pod względem merytorycznym. Rozprawa dobrze prowadzi czytelnika od definicji podstawowych pojęć, poprzez analizę ograniczeń zastosowanej metody modelowania sieci, do propozycji autorskich rozszerzeń w zakresie skalowalności i modelowania zmienności funkcjonowania sieci. Sposób organizacji tekstu jest klarowny, jednak pewne zastrzeżenie może budzić umieszczenie wyników własnych badań doktorantki nad porównaniem modeli aproksymacji płynnej i dyfuzyjnej w rozdziale 3, przed

rozdziałem 4, w którym istotną część stanowi stan wiedzy w zakresie wyznaczania modelu aproksymacji płynnej. Doktorantka jednak musiała wybrać jeden ze sposobów organizacji zawartości pracy, a przyjęta kolejność rozdziałów zapewnia, iż czytelnik płynnie przechodzi od stanu wiedzy do wyników oceny badań nad sposobami zrównoleglenia generowania modelu sieci. Przyjęta metoda organizacji pracy w sposób jednoznaczny pozwala na wyróżnienie wkładu własnego doktorantki od opisu stanu wiedzy, co jest istotne w rozprawach doktorskich.

3. Analiza źródeł i stanu wiedzy

Doktorantka dokonała dogłębnego przeglądu literaturowego i bardzo szczegółowo odniósła się do stanu wiedzy. W rozprawie zacytowano 205 prac, a zacytowane prace są związane z tematem rozprawy i adekwatne do treści pracy. Przegląd literatury jest wyczerpujący i dowodzi, że doktorantka dogłębnie przenalizowała postęp badań oraz zgłębiła istniejące prace z zakresu badania wydajności sieci komputerowych za pomocą aproksymacji płynnej i aproksymacji dyfuzyjnej. Odniesienia do literatury są dobrze umiejscowione w tekście pracy, a cytowane prace są adekwatne do treści poszczególnych fragmentów rozprawy. Co istotne, autorka na początku każdego z rozdziałów opisujących wyniki naukowe pracy przygotowała odniesienie do aktualnego stanu prac naukowych związanych z prezentowanymi badaniami (np. opis nowych modeli aproksymacji płynnych w rozdział 6.1.1 lub odniesienie do stanu wiedzy na początku rozdziału 5).

W rozprawie zawarto także zestawienie publikacji naukowych, w których opublikowano wyniki badań doktorantki. Wyniki prac naukowych doktorantki zostały już opublikowane w 4 czasopismach naukowych z współczynnikiem wpływu, w tym w 2 artykułach w czasopismach o Impact Factor powyżej 2. Dodatkowo doktorantka jest autorką 16 artykułów opublikowanych w materiałach konferencyjnych lub w czasopismach o mniejszym zasięgu. Potwierdza to, że wysoka jakość realizowanych badań została zweryfikowana przez recenzentów w kraju i za granicą.

4. Wartość naukowa rozprawy

Praca stanowi istotny przyczynek do rozwoju dziedziny nauki, jaką jest badanie wydajności sieci komputerowych. Doktorantka w sposób twórczy rozwinęła istniejące algorytmy generowania modeli analitycznych oraz przebadała, jak na wydajność takich algorytmów wpływa architektura systemu równoległego. Rozprawa w usystematyzowanych etapach prowadzi do osiągnięcia celu jakim jest modelowanie rozległych i złożonych sieci komputerowych za pomocą aproksymacji płynnej wspartej narzędziami programowymi. Modelowanie to umożliwia analizę wybranych scenariuszy dynamiki natężenia transmisji internetowych. Tezy szczegółowe rozprawy zostały udowodnione krok po kroku, a wszystkie cele pośrednie osiągnięte, co udokumentowano licznymi wynikami badań zaprezentowanymi w postaci wykresów bądź tabel.

Najbardziej istotnym wynikiem rozprawy, w mojej ocenie, jest opracowanie nowych metod generowania i analizowania modeli aproksymacji płynnej dla sieci złożonych z milionów urządzeń. Problem badania wydajności systemów teleinformatycznych ogromnej skali jest bardzo istotnym zagadnieniem badawczym, a stale rosnąca liczba urządzeń korzystających z sieci Internet oraz zwiększająca się złożoność infrastruktury sieciowej powoduje, że badanie wydajności sieci jest coraz trudniejsze. Doktorantka opracowała nową metodę badania modeli aproksymacji płynnej z wykorzystaniem baz danych, nazwaną SQL-FFA: Bazodanowa Aproksymacja Płynna. Metoda opracowana przez doktorantkę pozwala zwiększyć o co najmniej 2 rzędy wielkości złożoność modeli. W dotychczasowych pracach równania opisujące działanie protokołu TCP było możliwe do rozwiązania dla sieci złożonych co najwyżej z tysięcy urządzeń. Autorska metoda doktorantki pozwala wykonać badania dla sieci złożonych setek tysięcy węzłów, z ponad milionem przepływów (np. wyniki numeryczne przedstawione w rozdziale 6.1.3 zostały wykonane dla topologii sieci złożonej ze 134 023 węzłów i 1 352 081 przepływów). Doktorantka także znalazła rozwiązanie dla kolejnego problemu badawczego – podczas analizy złożonych sieci generowane są bardzo duże zbiory danych opisujące bieżący stan sieci oraz wyniki działania modelu. Zastosowanie baz danych w sposób opisany w rozprawie pozwoliło na efektywne przechowywanie i przetwarzanie tych zbiorów danych.

Badania są poprawnie zaplanowane i mają spójny charakter, a poszczególne elementy pracy dobrze łączą się ze sobą. W pracy postawiono 2 tezy:

(1) Metoda aproksymacji płynnej, po opracowaniu metod i narzędzi programowych, jest efektywnym sposobem badania sieci rozległych z uwzględnieniem występujących w nich stanów nieustalonych
(2) Przetwarzanie współbieżne można zastosować do numerycznego rozwiązywania równań stanowiących model aproksymacji płynnej

Treść rozprawy w pełni udowadnia postawione 2 tezy oraz w sposób wyczerpujący prezentuje analizę zastosowania przetwarzania współbieżnego i metod bazodanowych do rozwiązywania modeli sieci komputerowych o złożonych topologiach. Doktorantka przygotowała i przebadała działanie narzędzi programowych pozwalających na badanie sieci złożonych z stu tysięcy węzłów, tym samym wykazała, że możliwe jest badanie za pomocą aproksymacji płynnej bardzo złożonych sieci. Doktorantka także opracowała metody zrównoleglenia przetwarzania modelu aproksymacji płynnej, pozwalające z istotny sposób skrócić czas obliczeń i zwiększyć wielkość przetwarzanych modeli.

5. Poprawność redakcyjna rozprawy

Rozprawa została przygotowana w sposób bardzo staranny i nie zawiera znaczących błędów językowych lub redakcyjnych. W tekście rozprawy występują jedynie drobne usterki edycyjne i błędy interpunkcyjne, nie mające znaczącego wpływu na merytoryczną ocenę rozprawy, jak np.

- niepotrzebny przecinek na str. 6 w zdaniu „Wyniki czasowe, okazały się niezadowalające"
- str 11 - "przepustowośc"
- str 13 - "Badania wskazują, że optymalną wartości waga osiąga" -> powinno być „wartość"

- str 47 - "Model aproksymacji płynnej definiuje się jako układ równaniach różniczkowych stochastycznych" -> powinno być „równań"
- str. 64 - rozdział 4.2.2 w pierwszej linii niepotrzebny przecinek przed odwołaniem do pracy [158]
- str. 64 - niepotrzebna spacja przed znakiem kropki na str. 64 w zdaniu kończącym się „linii pamięci podręcznej ."
- str. 160 – pierwsze zdanie: „Modyfikacji pozwala" – powinno być „Modyfikacja pozwala"

6. Uwagi krytyczne

Na wykresach przedstawiających rozmieszczenie wątków na poszczególnych procesorach logicznych w rozdziale 4 przedstawiono zmienność przydziału liczby wątków do poszczególnych procesorów w czasie modelowania. Na podstawie tych wykresów trudno jednak zidentyfikować, jaki jest bezpośredni wpływ zmian w czasie liczby wątków na każdym z rdzeni na wydajność pracy. Intuicyjnie większa zmienność tej liczby może świadczyć o mniej efektywnym zarządzaniu zasobami przez system operacyjny, to bardziej wartościowe byłoby zestawienie miar zmienności (np. wariancji) z wynikową wydajnością. Czy autorka spróbowała zidentyfikować powiązanie pomiędzy w/w miarami, a wydajnością przetwarzania modelu aproksymacji płynnej?

Rozdział 3 traktuje o topologiach sieciowych. We wstępie rozdziału nie ma konkretnej informacji, która z przytaczanych w kolejnych podrozdziałach topologia sieci jest wykorzystywana w dalszych badaniach, za wyjątkiem lakonicznego stwierdzenia, że "Niniejsza praca koncentrowała się wyłącznie na wyborze topologii, która swoim rozmiarem sięga skali Internetu i będzie mogła służyć do potwierdzenia tez pracy". W podrozdziale 3.2.1 można przeczytać, iż " jedynie dla mapy z 2003 roku dane źródłowe zostały wówczas udostępnione publicznie, dlatego to właśnie ta topologia została wykorzystana w niniejszej rozprawie". Jakie jest powiazanie pomiędzy pozostałymi topologiami prezentowanymi w rozdziale 3, a resztą pracy? Czy opracowane metody generacji topologii mogą znaleźć szersze zastosowanie i czy autorka rozważała udostępnienie opracowanych narzędzi do generacji topologii w formie np. oprogramowania Open Source?

W rozdziale 6.2 przedstawiono wyniki opracowanego przez autorkę rozszerzenia metody aproksymacji płynnej o możliwość modyfikacji parametrów węzłów sieci, opublikowanej w pracy
M. Nycz. Modeling of Computer Networks Using SAP HANA Smart Data Streaming. In A. Kwiecie«, P. Gaj, and P. Stera, editors, Computer Network, volume 1039 of Communications in Computer and Information Science, pages 48-61. Springer, 2019.
Opis metody, zawarty na stronie 160 jest jednak mało szczegółowy i sprowadza się do opisu trybów, pomiędzy którymi jest przełączana konfiguracja. Na czym polega zmiana sposobu przetwarzania modelu aproksymacji płynnej z punktu widzenia oprogramowania realizującego obliczenia numeryczne i czym zastosowane przez Doktorantkę podejście różni się od ponownego uruchomienia przetwarzania z warunkami początkowymi wyliczonymi wedle wyników działania modelu dla wcześniejszego zestawu parametrów?

W rozdziale 2 zastosowano taki dobór parametrów aproksymacji płynnej, aby uzyskać model kolejki typu „Odrzucanie Ostatniego (ang. Drop Tail). Spowodowało to wprowadzenie do modelu gwałtownego skoku prawdopodobieństwa odrzucenia pakietu z 0 na 1, co widać na wykresach 2.1 i 2.7. W mechanizmie RED, który jest typowo modelowany za pomocą aproksymacji płynnej ten skok nie jest tak gwałtowny, co może powodować, że błąd modelu aproksymacji płynnej jest mniejszy dla mechanizmu RED, niż dla mechanizmu Drop Tail przebadanego w rozdziale 2.

7. Podsumowanie i ocena końcowa

Praca doktorska pt. „Modelowanie dynamiki transmisji internetowych za pomoci aproksymacji płynnej" została przygotowana przez Panią Magister Monikę Nycz bardzo rzetelnie i wykazuje zdolność kandydatki do samodzielnego prowadzenia pracy naukowej. Praca jest poprawna pod względem metodologicznym, a opracowane metody przetwarzania równoległego i metody realizacji obliczeń numerycznych dla złożonych modeli analitycznych sieci komputerowych wykraczają poza stan wiedzy i stanowią istotny przyczynek do rozwoju dziedziny analizy efektywności systemów informatycznych. Badania zostały poprawnie zaplanowane i zrealizowane, a ich wyniki szczegółowo i precyzyjnie opisane. Praca tym samym potwierdza, iż kandydatka posiada wymaganą wiedzę teoretyczną w dyscyplinie Informatyka Techniczna i Telekomunikacja. Szczególnie istotnym elementem rozprawy jest opracowana metoda realizacji modelu aproksymacji płynnej z wykorzystaniem silnika bazy danych, pozwalająca na analizę modeli sieci złożonych z setek tysięcy węzłów. Dlatego wnioskuję o wyróżnienie rozprawy.

Ja, niżej podpisany stwierdzam, że recenzowana rozprawa doktorska Moniki Nycz spełnia warunki określone w art. 187 ust. 1 i 2 Ustawy z dnia 20 lipca 2018r. Prawo o szkolnictwie wyższym i nauce (Dz.U. 2018 r., poz. 1668 z późn. zmianami) i wnioskuję do Rady Dyscypliny Informatyka Techniczna i Telekomunikacja o dopuszczenie Moniki Nycz do dalszych etapów przewodu doktorskiego.

17.11.2023

.....

data sporządzenia recenzji

podpis recenzenta