



**POLITECHNIKA
GDAŃSKA**

WYDZIAŁ ELEKTRONIKI,
TELEKOMUNIKACJI I INFORMATYKI

Gdańsk, 19.01.2023 r.

Dr hab. inż. Jacek Rak, prof. PG
Wydział Elektroniki, Telekomunikacji i Informatyki
Politechnika Gdańska

Recenzja rozprawy doktorskiej

Tytuł: Traffic forecasting in optical networks with predefined traffic levels
(Prognozowanie ruchu w sieciach optycznych posiadających ustalone poziomy ruchu)

Autor: mgr inż. Daniel Szostak

Promotor: prof. dr hab. inż. Krzysztof Walkowiak

Rozprawa doktorska mgr inż. Daniela Szostaka rozpatruje zagadnienie prognozowania ruchu sieciowego w kontekście transmisji w rozległych sieciach optycznych wykorzystujących technikę zwielokrotnienia z podziałem długości fal (ang. Wavelength Division Multiplexing – WDM). Poruszany problem jest niewątpliwie istotny oraz aktualny – w szczególności w kontekście efektywnej realizacji zadań sieci transportowych zarówno w przypadku typowego scenariusza pracy (bez awarii), jak i w okresach awarii jej elementów. Prognozowanie ruchu jest ważną techniką pozwalającą na dokładniejsze oszacowanie ilości zasobów sieci (przepustowości łączy), co może być przydatne np. na etapie planowania sieci.

Oceniana rozprawa doktorska została napisana w języku angielskim i składa się z sześciu głównych (numerowanych) rozdziałów, kolejno: wprowadzenia przedstawiającego motywację pracy, tezę oraz cele rozprawy (Rozdział 1); ogólnego wprowadzenia do zagadnień związanych z ruchem w sieciach teleinformatycznych, uczeniem maszynowym, szeregami czasowymi, metodami prognozowania ruchu znanymi z literatury (Rozdział 2); definicji problemu rozpatrywanego w rozprawie (Rozdział 3); szczegółowej analizy problemu prognozowania krótkoterminowego (Rozdział 4) oraz długoterminowego (Rozdział 5); podsumowania oraz przedstawienia potencjalnych dalszych kierunków badań (Rozdział 6). Rozprawa obejmuje także kilka części nienumerowanych, w tym m.in. streszczenie, wykazy rysunków i tabel, wykazy skrótów i symboli oraz wykaz pozycji literatury.

Rak

WPŁYNĘŁO

23-01-2023

1

QDN-IT/16/2023

1. Jaki jest cel naukowy rozprawy i czy został on trafnie i jasno sformułowany?

Celem pracy było opracowanie metod krótkoterminowego i długoterminowego prognozowania ruchu w sieciach optycznych. Cel ten został jasno określony w początkowej części rozprawy. Cel rozprawy należy także uznać za trafny m.in. z uwagi na rosnące natężenie ruchu w optycznych sieciach szkieletowych, jak i znaczenie optymalizacji rozdziału zasobów sieci w realizacji usług w odniesieniu do ogromnej liczby użytkowników końcowych. Cel ten ma ścisły związek z tezą pracy której tłumaczenie (recenzenta) na język polski brzmi: *Istnieją metody krótkoterminowego i długoterminowego prognozowania ruchu w sieciach optycznych, w których transmisja bazuje na predefiniowanych poziomach ruchu.*

Teza rozprawy została zdefiniowana niewątpliwie poprawnie, choć jej forma wydaje się nieco ogólna (jej udowodnienie sprowadza się do wykazania, że wspomniane metody można opracować i mogą mieć one zastosowanie do prognozowania ruchu w sieciach teleinformatycznych). W mojej opinii nie odbiega ona jednakże znacząco formą od tez rozpraw doktorskich.

Zakres prac zdefiniowany przez Doktoranta na stronie 2 rozprawy obejmuje:

- opracowanie i implementację metod przetworzenia wstępnego historycznych strumieni danych w celu uzyskania danych wejściowych dla algorytmów uczenia maszynowego oraz metod pozostałych rozpatrywanych typów;
- opracowanie strategii krótkoterminowego i długoterminowego prognozowania ruchu przy wykorzystaniu algorytmów uczenia maszynowego i szeregów czasowych;
- zdefiniowanie metryki oceny jakości przydatnej do oszacowania wydajności proponowanych algorytmów i wynikowej jakości rezultatów prognozowania poziomu ruchu;
- szczegółową ocenę efektywności najlepszej z proponowanych metod przeprowadzoną dla założonego zbioru scenariuszy testowych;
- zebranie danych dotyczących ruchu rzeczywistego.

Ogólnie powyższy zestaw zadań należy uznać za zasadny w celu udowodnienia tezy rozprawy, a stopień złożoności powyższych prac jest adekwatny do złożoności badań naukowych na poziomie doktorskim.

Jednakże wątpliwości budzi umieszczenie celu „zebranie danych dotyczących ruchu rzeczywistego” na końcu listy, gdyż sugeruje jego realizację dopiero po przeprowadzeniu analizy symulacyjnej wykorzystującej dane wygenerowane. Jako że pozyskanie danych dotyczących ruchu rzeczywistego dostarczało danych wejściowych do symulacji, wydaje się zasadne umieszczenie tego celu wyżej na liście celów.

MPK

2. Na czym polega oryginalny dorobek Autora i jakie jest znaczenie poznawcze lub przydatność praktyczna dla nauki bądź techniki?

Oryginalny dorobek Doktoranta ukazany w ocenianej rozprawie moim zdaniem obejmuje:

- 1) omówienie w rozdziale 2.1 zagadnień podstawowych związanych z charakterystyką ruchu w optycznych sieciach szkieletowych wyjaśniające również przyczyny tendencji wzrostowej natężenia ruchu w ostatnich kilku latach, jak i prezentujące w sposób zwięzły możliwości transmisyjne sieci optycznych;
- 2) objaśnienie zasad działania algorytmów uczenia maszynowego oraz szeregów czasowych przedstawione w rozdziale 2.2;
- 3) przeprowadzenie relatywnie obszernego przeglądu literatury w kontekście metod prognozowania ruchu w sieciach teleinformatycznych zawartego w rozdziale 2.3 wraz z identyfikacją otwartych problemów w tym zakresie (przede wszystkim rozpatrywanego w niniejszej rozprawie problemu prognozowania w kontekście poziomego ruchu);
- 4) zdefiniowanie w rozdziale 3 problemu prognozowania ruchu oraz zdefiniowanie w rozdziale 3.2 sposobu konstrukcji zbioru danych wejściowych dotyczącego zmiennego w czasie natężenia ruchu na łączach analizowanej sieci (na bazie danych rzeczywistych pochodzących z repozytorium <https://www.seattleix.net/>);
- 5) propozycję metryki TLPQ oceny jakości prognozowania poziomego ruchu przedstawioną w rozdziale 3.4;
- 6) rezultaty analizy krótkoterminowego prognozowania ruchu opisane w rozdziale 4;
- 7) analizę w zakresie długoterminowego prognozowania ruchu przedstawioną w rozdziale 5.

3. Czy Autor rozwiązał postawiony problem i czy użył do tego celu właściwych metod?

Moim zdaniem wyniki zawarte w rozprawie doktorskiej mgra inż. Daniela Szostaka świadczą o tym, że problem badawczy zdefiniowany w rozprawie został przez Autora rozwiązany w sposób właściwy. Należy podkreślić, że przyjęta przez doktoranta metodologia jest adekwatna dla badań naukowych, a stopień złożoności odpowiada ogólnym oczekiwaniom w kontekście prac na poziomie doktorskim. Kolejnym aspektem jest praktyczne znaczenie uzyskanych rezultatów: Doktorant rozwiązał aktualny i realny problem badawczy w zakresie oszacowania natężenia ruchu w sieciach teleinformatycznych.

Wartą odnotowania jest także rozpoznawalność prac Doktoranta mierzona liczbą cytowań. Liczba prac Doktoranta związanych z rozprawą doktorską wynosi 8 (wykaz ze strony ix rozprawy). Na dzień sporządzenia tejże recenzji liczba cytowań prac Doktoranta w serwisie Google Scholar wynosiła 39.

Rak

4. Jakie są słabsze strony rozprawy?

Uwagi natury ogólnej

1. Pomimo, że praca została napisana ogólnie poprawnie pod względem językowym (praca w języku angielskim), dosyć wyraźnie widać w pracy ograniczenia w zakresie użycia poprawnych form przedimków określonych (the) i nieokreślonych (a/an). Najczęstszym problemem jest pomijanie przedimków w miejscach, gdzie powinny być one użyte.
2. Na stronie xxiii, skrót nazwy "communications service provider" powinien mieć postać CSP a nie CPS.
3. Strona xxv: "...links connected..." → "...links connecting..."
4. Kwestie gramatyki/stylu (wybrane przypadki):
 - na stronie 5:
"...often without being aware of it." → "...often without being aware of **them**."
"This facts have..." → "**These** facts have..."
 - polski szyk wyrazów w zdaniu angielskim, np. na str. 7:
"...devices have prefix "smart"..." → "...devices have the "smart" prefix..."
 - problem stylu np. na str. 11 początku zdania "It is a network, which has a cognitive process..."
 - "granulation" → "granularity" na str. 31,
 - "Datasets used for experiments **contains**..." → **contain** (na str. 32).
5. Niewłaściwe rozpoczynanie zdania od numeru pozycji literatury (np. na str. 7: "[85] presents its...") lub od liczby (np. na str. 22: "90 datasets of P...").
6. Zastosowanie potocznych określeń, np. na str. 8: "...into shorter **pieces**..." → "...into shorter **parts**..."
7. Nieprecyzyjne stwierdzenia, np. na str. 9 ".... to meet higher speed."
8. W dwóch ostatnich wierszach Tabeli 1 zawarte jest zaskakujące stwierdzenie "To be defined", mimo że rok działania (określony w tabeli jako 2020) już minął.
9. Rysunki przedstawiające różnej postaci schematy, wykresy (Rys. 3-10, 15-17, 18-23, 26-29) mają bardzo ograniczoną czytelność. Widać, że zostały wklejone jako grafika punktowa poddana silnej kompresji. Autor najwyraźniej nie zastosował w tym kontekście narzędzia do tworzenia grafiki wektorowej, bądź został popełniony błąd na etapie sposobu osadzania tychże obiektów w dokumencie tekstowym (tryb kopiowania). Skutkiem tego jest wyraźne rozmycie szczegółów obrazów skutkujące nieomalże całkowitym brakiem czytelności pewnych obszarów rysunków (np. wartości osi poziomych i pionowych na rysunkach 18-22).
10. Niejasny jest powód zastosowania odrębnego symbolu "*" (odmiennego od typowego symbolu operatora mnożenia) we wzorze (2.8) na stronie 23.

Wok

11. Co oznacza skrót ("rrd" – pisownia oryginalna z małej litery) w rozprawie na str. 32?
12. Forma powołania się w zdaniach na sekwencję rysunków np. na stronach 34, 38, 40 nie jest poprawna. Zamiast pisać np. "Figures Figure 10 – Figure 14" powinien być zastosowany krótszy "Figs. 10-14".

Uwagi szczegółowe oraz polemiczne

1. Ograniczeniem pracy jest wykorzystanie jedynie niewielkiego zbioru danych rzeczywistych dotyczących ruchu (w odniesieniu do węzła Internet Exchange Point w Seattle, USA z dnia 27.12.2020) do oszacowania charakterystyk ruchu na łączach symulowanej sieci Euro 28 (jak zostało przedstawione na str. 32 rozprawy). Problem ten można częściowo uzasadnić ograniczonym zakresem udostępniania przez operatorów danych dotyczących ruchu rzeczywistego. Niemniej powyższy stan rzeczy sprawia, że wartość praktyczna wyników może odbiegać od rezultatów możliwych do uzyskania na podstawie szerszego zbioru danych rzeczywistych.
2. Mimo, że proponowana przez Autora metryka TLPQ oceny jakości prognozowania poziomu ruchu jest niewątpliwie potrzebnym składnikiem pracy, w rozprawie brakuje analizy właściwości tejże metryki. Przeprowadzenie takiej analizy wydaje się być niezbędne w odniesieniu do każdej metryki przed jej zastosowaniem.
3. Autor stwierdza na stronie 2 rozprawy, że zaproponowane rozwiązania mogą być zastosowane praktycznie dla każdej architektury sieci. Niemniej rozprawa porusza w zasadzie wyłącznie kontekst sieci optycznych. Nasuwa się więc pytanie o celowość takiego zabiegu. Ponadto, odmiennie niż w przypadku sieci szkieletowych, w obszarach sieci bliższych użytkownikowi końcowemu (w obszarach dostępowych, sieciach lokalnych) wzorzec ruchu może być zupełnie odmienny od relatywnie „gładkich” charakterystyk czasowych natężenia ruchu spodziewanych w obszarze rdzenia sieci.
4. Rozszerzając powyższą uwagę, należałoby się również zastanowić jaka jest górna granica stopnia „porowatości” charakterystyki czasowej natężenia ruchu, do której rozwiązania Doktoranta są użyteczne. W tym kontekście ważne byłoby zastanowienie się nad wpływem czasu obliczeń na użyteczność rezultatów (co może być niewątpliwie problemem w przypadku sieci o dużej dynamice zmienności natężenia ruchu). Ponadto, czy dodatkowe charakterystyki środowiskowe (np. mobilność węzłów bezprzewodowych) mogłyby mieć wpływ na te charakterystyki?
5. Użyteczne dla operatorów byłoby przedyskutowanie możliwości wykorzystania rozwiązań proponowanych w niniejszej rozprawie w zadaniu wymiarowania sieci w przypadku różnych strategii doboru tras (w tym uwzględniających ochronę transmisji poprzez trasy zabezpieczające) o niewątpliwie zwiększonym (w porównaniu z przypadkiem pojedynczych tras transmisji) zapotrzebowaniu na zasoby sieci.

Powyższe uwagi nie wpływają na moją pozytywną łączną ocenę rozprawy.

5. Do której z następujących kategorii Recenzent zalicza rozprawę:

- a/ nie spełniająca wymagań,
- b/ wymagająca wprowadzenia poprawek i ponownego recenzowania,
- c/ spełniająca wymagania,
- d/ wykraczająca ponad poziom zadawalający (spełniająca wymagania z nadmiarem),
- e/ wybitna?

Podsumowując, moja recenzja pracy doktorskiej mgr inż. Daniela Szostaka jest pozytywna. Wartość merytoryczna rozprawy jak i sposób prezentacji wyników kwalifikują w mojej ocenie rozprawę mgr inż. Daniela Szostaka do kategorii c/ spełniająca wymagania ustawowe stawiane rozprawom doktorskim.

Wnoszę o dopuszczenie rozprawy doktorskiej mgr inż. Daniela Szostaka do publicznej obrony.

