



WYDZIAŁ MATEMATYKI I INFORMATYKI

INSTYTUT INFORMATYKI

ul. Joliot-Curie 15

50-383 Wrocław

tel. +48 71 375 78 00 | +48 71 325 12 71

fax +48 71 375 78 01

sekretariat.ii@uwr.edu.pl | www.ii.uni.wroc.pl

Tomasz Jurdziński
Wydział Matematyki i Informatyki
Instytut Informatyki
Uniwersytet Wrocławski

Wrocław, 5.05.2025

Recenzja rozprawy doktorskiej mgr. Patryka Stopyry zatytułowanej „Random-
Access Channel in Time-Critical Ad Hoc Network Synchronization”

Niniejsza recenzja rozprawy doktorskiej mgr. inż. Patryka Stopyry zatytułowanej „Local Computation Algorithms for Hypergraph Coloring” została przygotowana w odpowiedzi na pismo Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja na Politechnice Wrocławskiej, prof. dr hab. Wojciecha Bożejki.

Główne wyniki rozprawy zaprezentowane zostały w jej rozdziałach 3 i 4. Wyniki z rozdziału 3 rozprawy zostały opublikowane w czasopiśmie Vehicular Communications (Elsevier) o wysokich wskaźnikach bibliometrycznych i wysokiej punktacji MNiSW (140 punktów). Współautorami tych rezultatów są promotor i prof. Jacek Cichoń. Natomiast, zgodnie z deklaracją autora, treść rozdziału 4 nie była wcześniej prezentowana w formie publikacji bądź manuskryptu.

W dalszej części recenzji omówię tematykę i główne wyniki rozprawy, następnie przejdę do mojej oceny, uwag i podsumowania.

Tematyka i treść rozprawy

Rozprawa poświęcona jest problemowi dostępu do dzielonego kanału (Multiple Access Channel) przez zbiór niezależnych agentów. Dokładniej przedmiotem badań jest następująco sformułowany problem, nazwany *synchronizacją*. Każdy element zbioru k niezależnych agentów ma do przesłania jedną wiadomość poprzez współdzielony kanał komunikacyjny, w którym dostęp podzielony jest na n slotów a każdy slot umożliwia przesłanie jednej wiadomości. Próba przesłania wiadomości przez więcej niż jednego agenta w slotie kończy się porażką każdego z próbujących agentów. Przyjęto też tutaj, że brak jest potwierdzeń dla agentów pozwalających ustalić, w których rundach następuje udane wysłanie wiadomości. Agenci są nierozróżnialni i autonomicznie uruchamiają ten sam algorytm. Założenia te powodują, że dla złamania symetrii konieczne jest użycie losowości. Model taki jest badany jako abstrakcja ilustrująca dostęp do wspólnego kanału komunikacyjnego w sieciach bezprzewodowych (sensorowych, Internet of Things, Vehicle to Vehicle communication itp.) a także w sieci Ethernet.

Oprócz badanego w rozdziale 3 oryginalnego problemu synchronizacji, w rozprawie został sformułowany i poddany badaniom zrelaksowany wariant problemu, w którym dopuszcza się, że pewnej liczbie $f < k$ agentów nie uda się wysłać wiadomości. Na potrzeby niniejszej recenzji nazywać będę ten problem *niepełną synchronizacją*.

WPLYNEŁO

07-05-2025

RON/IT/122/2025

Rozprawa ma bardzo jasno sformułowany plan badawczy i hipotezę. Uwaga skoncentrowana jest na dwóch prostych algorytmach realizacji powyżej opisanego zadania nazwanych BT (od Bernoulli Trials) i t SLOTS. W pierwszym algorytmie każdy agent w każdym slotcie, niezależnie od pozostałych slotów, wysyła wiadomość z ustalonym prawdopodobieństwem p . Z kolei w metodzie t SLOTS algorytm wybiera losowo t spośród n dostępnych slotów a każdy wybór jest jednakowo prawdopodobny. Podstawową miarą jakości algorytmu jest prawdopodobieństwo realizacji zadania synchronizacji przy zadanej liczbie slotów n , liczbie agentów k oraz wartości parametru p dla algorytmu BT oraz t dla algorytmu t SLOTS. W rozprawie badany jest także dobór optymalnych wartości p i t .

Celem rozprawy jest kompleksowa weryfikacja jasno postawionej hipotezy, że t SLOTS przewyższa BT w badanych miarach wydajności algorytmów synchronizacji dla występujących w praktyce wartości n („if n , the number of its agents, does not exceed low thousands”). Oprócz problemu wyznaczenia samego prawdopodobieństwa sukcesu w rozprawie badane są inne własności obu protokołów, w tym odporność na nieoptymalny wybór parametrów p i t , złożoność czasowa i pamięciowa implementacji obu algorytmów, wpływ wartości parametru tolerancji błędów f na wyniki obu algorytmów, skalowalność algorytmów.

Pierwsze dwa rozdziały rozprawy wprowadzają w tematykę rozprawy, zastosowania problemu synchronizacji a także terminologię i narzędzia matematyczne wykorzystywane w głównych rozdziałach 3 i 4.

Rozdział 3 w całości poświęcony jest problemowi (pełnej) synchronizacji. Z samej definicji problemu wynika wprost, że synchronizacja jest w ogóle możliwa o ile $n \geq k$. Wiadomo, że optymalne prawdopodobieństwo p w algorytmie BT wynosi $p = 1/k$ niezależnie od wartości n . Nietrudno też zapewne wykazać, że przy zmiennych wartościach n i k , optymalna wartość t dla t SLOTS jest rzędu $\theta(n/k)$. Asymptotycznie zachodzi też własność, że przy $k = o(n)$ i optymalnym doborze parametrów p i t , prawdopodobieństwo pełnej synchronizacji jest $1 - o(1)$. Niemniej, nie są znane wyrażenia w zwartej postaci pozwalające dokładnie wyznaczyć prawdopodobieństwo synchronizacji przy zadanych parametrach bądź optymalną wartość parametru t dla t SLOTS. Co więcej, choć można wyznaczyć wartości asymptotyczne, brak jest precyzyjnych oszacowań dokładnych wartości tych wielkości. A zatem aktualny stan wiedzy nie dawał możliwości weryfikacji postawionych hipotez. Wyznaczanie prawdopodobieństw synchronizacji wprost z definicji jest praktycznie niewykonalne, ponieważ rozmiar przestrzeni zdarzeń elementarnych dla obu algorytmów to odpowiednio 2^{kn} oraz $\binom{n}{t}^k$. Przybliżone informacje można uzyskać poprzez wielokrotne symulacje algorytmów i ocenę ich efektywności w oparciu o wykonane dla ich wyników testy statystyczne. W rozprawie wybrane zostało jednak dość nietypowe, inne podejście. Co prawda nie udało się wyznaczyć zwartych wyrażeń wyznaczających prawdopodobieństwa (niepełnej) synchronizacji lub ich dokładne przybliżenia, podjęta została natomiast próba uproszczenia wyrażeń wyznaczających te prawdopodobieństwa wg ich definicji do postaci, w której liczba składników i operacji arytmetycznych jest możliwa do wykonania dla niedużych wartości parametrów. W efekcie w przypadku pełnej synchronizacji udało się uzyskać wyrażenia, które pozwalają przy liczbie operacji arytmetycznych rzędu $O(k)$ dla algorytmu BT (Theorem 3.1) i rzędu $O(k \cdot t^k)$ dla t SLOTS (Theorem 3.4) wyznaczyć dokładne wartości prawdopodobieństw pełnej synchronizacji. Wykorzystując te wyniki Pan Stopyra przeprowadził dość kompleksowe porównanie BT i t SLOTS dla wybranych wartości

$n \leq 64$ i niedużych k . W tych porównaniach wykorzystane zostały uzyskane uproszczone wyrażenia na prawdopodobieństwa synchronizacji a także przeprowadzone były symulacje i testy statystyczne, które w pełni potwierdziły poprawność formuł wyznaczonych analitycznie. Dla większych wartości n i k (największe to $n = 8000$ i $k = 500$) przeprowadzone zostały natomiast tylko badania symulacyjno-statystyczne. Co ciekawe wszystkie uzyskane wyniki potwierdziły postawioną hipotezę o „wyższości” algorytmu tSLOTS nad BT. W szczególności, przy optymalnym dla danych n i k doborze parametrów p i t prawdopodobieństwo synchronizacji tSLOTS przewyższa prawdopodobieństwo synchronizacji BT. Co więcej, gdy k jest stosunkowo duże względem n (w efekcie prawdopodobieństwo synchronizacji obu algorytmów staje się małe), różnice stają się dość znaczne. Ponadto, w przypadku nieoptymalnego doboru parametrów (co jest naturalne np. w środowisku sieci ad hoc, gdzie rzeczywista wartość k nie jest znana agentom), tSLOTS wykazuje znacznie większą odporność niż algorytm BT. Innym trochę nieoczywistym wnioskiem z przeprowadzonych badań jest fakt, że optymalne wartości parametrów dla tSLOTS dają nieco mniejszy koszt energetyczny związany z transmisjami niż w przypadku BT. Okazuje się mianowicie, że podczas gdy optymalna wartość $p = 1/k$ oznacza wartość oczekiwaną liczby transmisji agenta równą n/k w BT, optymalna wartość parametru t w tSLOTS jest w niektórych badanych przypadkach mniejsza od $\lfloor \frac{n}{k} \rfloor$. Dodatkowym atutem jest tutaj fakt, że liczba transmisji agenta w tSLOTS nie jest zmienną losową a wartością z góry zadaną.

W rozdziale 4 uwaga przenosi się na bardziej skomplikowany w analizie problem niepełnej synchronizacji. Podążając ścieżką wyznaczoną w rozdziale 3 główny nakład pracy koncepcyjnej włożony został w uzyskanie uproszczonych formuł na wartości prawdopodobieństw niepełnej synchronizacji przy zadanych wartościach parametrów n, k i f , gdzie f oznacza maksymalną liczbę agentów, dla których algorytm zakończy się „błędnie”. Uwzględnienie parametru f znacznie utrudnia problem co znalazło odzwierciedlenie w uzyskanych rozwiązaniach (Corollary 4.2 dla BT oraz Corollary 4.4 dla tSLOTS). Również uzyskane wyrażenia są bardziej skomplikowane i wyznaczenie ich wartości wymaga nieco większej ilości obliczeń. Niemniej, również tutaj liczba obliczeń nie jest istotnie zależna od n (n występuje jedynie jako argument podstawowych operacji arytmetycznych). Ponieważ optymalność doboru parametru $p = 1/k$ w algorytmie BT niezależnie od wartości f jest w przypadku problemu niepełnej synchronizacji mniej jasna, Pan Stopyra udowodnił formalnie tę własność. Korzystając z uzyskanych uproszczonych wyrażeń dla prawdopodobieństw niepełnej synchronizacji przeprowadzone zostały kompleksowe obliczenia dla wybranych wartości parametru $n \leq 64$ oraz wybranych niedużych wartości k i f . Wyniki te zostały też z sukcesem potwierdzone obliczeniami symulacyjno-statystycznymi. W przypadku większych wartości n i k wykonane zostały tylko obliczenia statystyczne dla wyników przeprowadzonych symulacji. Uzyskane wyniki ponownie potwierdziły wyższość algorytmu tSLOTS nad BT we wszystkich badanych aspektach. W szczególności np. gdy akceptowana liczba błędów wynosi $f = 300$ dla $n = 8000$ i $k = 2000$, różnice są ogromne.

Rozprawę zamyka rozdział 5 podsumowujący uzyskane rezultaty i wskazujący możliwe kierunki badawcze wynikające z wyników badań naukowych zaprezentowanych w rozprawie.

Ocena rozprawy

W mojej ocenie najciekawszą i stanowiącą największe wyzwanie koncepcyjną częścią rozprawy są uproszczenia wyrażeń dla dokładnych wartości prawdopodobieństw synchronizacji i niepełnej synchronizacji obu algorytmów. W przypadku algorytmu BT jak się wydaje kluczowa była znajomość odpowiednich narzędzi kombinatorycznych, szczególnie w przypadku niepełnej synchronizacji. Oszacowania dla algorytmu tSLOTS opierają się na pomysłowej reprezentacji przestrzeni zdarzeń elementarnych odpowiadających różnym wynikom działania algorytmu przy pomocy wielomianów wielu zmiennych, w których każdemu zdarzeniu odpowiada osobny jednomian a odpowiednio zastosowane operacje ekstrakcji prowadzą do wyselekcjonowania wyłącznie elementów odpowiadających zdarzeniom odpowiadającym udanej (niepełnej) synchronizacji. Umiejętność swobodnego i skutecznego zastosowania tej techniki prowadzącego do stosunkowo zwartych postaci wyrażeń zasługuje na uznanie, tym bardziej, że nie jest to typowe narzędzie w wykształceniu informatyka.

Rozprawa została bardzo dobrze przygotowana od strony formalnej, nie ma w niej właściwie poważnych błędów czy pomyłek, zadbano o kompletność dowodów, notacji, odwołań do literatury. Wykresy przedstawiające wykonane obliczenia dobrze ilustrują stawiane tezy. Pozytywnie też należy ocenić fakt, że wszystkie utworzone na potrzeby obliczeń elementy oprogramowania zostały udostępnione poprzez github, a rozprawa zawiera odpowiednie odnośniki do nich.

Podsumowując sądzę, że Pan Stopyra dobrze wywiązał się z postawionego przed nim zadania badawczego. Zweryfikował też w kompleksowy sposób postawioną w rozprawie hipotezę badawczą. Realizacja tego zadania wymagała umiejętności informatycznych jak i wykorzystania zaawansowanych narzędzi matematycznych.

Uwagi

Omówienie wcześniejszych wyników w tematyce rozprawy mogłoby moim zdaniem zawierać konkretne odwołania do asymptotycznych oszacowań sukcesu dla problemu synchronizacji badanych algorytmów. Wg mojej wiedzy wyniki takie pojawiały się w różnych publikacjach, niekoniecznie jako główne ich rezultaty. Pozwoliłoby to też omówić optymalny dobór parametrów algorytmów w sytuacji gdy znana jest wartość k , dla dowolnie dużych n i k .

Drobne uwagi techniczne (z pominięciem językowych):

- Co prawda nie powoduje to trudności w rozumieniu prezentowanych rozumowań matematycznych, ale brakuje w nich czasem wprowadzenia oznaczeń, których czytelnik musi „domyślić” się z przeprowadzanych rozumowań. Jako przykłady posłużyć tutaj może użycie symboli M i F w dowodzie Theorem 3.4 czy symboli Y_σ, Y_β w dowodzie Theorem 4.4.
- Sugerowane w rozdziale 4 stosowanie zaproponowanych rozwiązań w wielu n -slotowych rundach z wyłączaniem agentów, którzy z sukcesem wysłali wiadomości w rundach poprzedzających wymaga modelu z feedbackiem, który zasadniczo nie jest przedmiotem rozprawy (można w nim bardziej efektywne rozwiązania uzyskać również w jednej rundzie).
- Wydaje mi się, że ciekawe byłoby w rozdziale 4 dokładniejsze zbadanie jak wartość parametru f wpływa na optymalną wartość t w algorytmie tSLOTS.

- Sekcja 1.4: jaka jest różnica między wektorami a krotkami w notacji?
- Wyrażenie w sumie w (1.10): indeksy powinny być do r , nie do k
- Mam wrażenie, że Fact 5 i Fact 8 to w rzeczywistości definicje.
- Nie do końca rozumiem dlaczego backoff nie jest „(for) time-critical synchronization”.
- Def. 2.8: fragment $p \in \text{Uniform}(0,1)$ może błędnie sugerować, że p wybierane jest losowo. Podobne wątpliwości mam w ostatnim zdaniu za definicją.
- Def. 2.10: liczba agentów k , nie n .

Podsumowanie

Uwzględniając wiedzę, umiejętności a także pewną pomysłowość i sprawność techniczną zaprezentowane przez Pana Stopyrę w rozprawie uważam, że zgodnie z zapisami ustawowymi jego *rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie informatyka i demonstrowa umiejętności samodzielnego prowadzenia pracy naukowej lub artystycznej*. A także, że rozprawa przedstawia *oryginalne rozwiązanie problemu naukowego w niej postawionego*.

W związku z powyższym rozprawa spełnia wymagania stawiane dla uzyskania stopnia *doktora*. Wniosuję o pozytywne zaopiniowanie rozprawy i dopuszczenie do dalszych etapów postępowania w sprawie o nadania stopnia doktora.

