

*dr hab. inż. Antoni Zabłudowski, prof. nzw. UTP*  
*Instytut Telekomunikacji*  
*Uniwersytet Technologiczno-Przyrodniczy*  
*w Bydgoszczy*

*Bydgoszcz, 10.05.2010.*

**RECENZJA ROZPRAWY DOKTORSKIEJ  
DLA RADY WYDZIAŁU ELEKTRONIKI I TELEKOMUNIKACJI  
POLITECHNIKI POZNAŃSKIEJ**

*Tytuł rozprawy:*

**Badania algorytmów heurystycznych dla połączeń  
rozgałęźnych w sieciach pakietowych**

*Autor rozprawy:*

**mgr inż. Maciej Piechowiak**

Rozprawa doktorska Pana Macieja Piechowiaka obejmuje zakres badań, które związane są z algorytmami routingów dla połączeń rozgałęźnych (połączeń multicastowych). Z definicji połączeń multicastowych wynika, że transmisja multicastowa jest takim sposobem dystrybucji informacji, dla której istnieje jeden węzeł źródłowy oraz grupa węzłów odbiorczych składająca się z większej niż jeden, liczby odbiorców, a dane są transportowane w sieci na zasadzie jeden do wielu. W sieciach pracujących z wykorzystaniem protokołu IP cała grupa odbiorców typu multicast jest widziana przez nadawcę jako pojedynczy, odbiorca grupowy, który jest wywoływany jednym multicastowym adresem IP. Multicastowy tryb transferu danych jest realizowany w sieciach pakietowych dla świadczenia usług multimedialnych, w tym w szczególności usług przesyłania obrazu video do wielu odbiorców (np. usług IP TV, usług near VoD), usług videokonferencyjnych, czy usług dystrybucji oprogramowania w sieci. Na obecnym etapie rozwoju sieci pracujących w oparciu o protokół IP, wolumen ruchu video jest najszybciej rosnącą składową całego ruchu w takiej sieci. Z prognoz dotyczących oceny wolumenu ruchu wynika, że ruch multicastowy w sieci IP będzie przyrastał wykładniczo.

Algorytmy wyznaczania dróg dla połączeń multicastowych (algorytmy routingowe) wykorzystują rozwiązania problemu drzewa Steinera. Problem drzewa Steinera jest zdefiniowany w następujący sposób: niech będzie dany graf ważony  $G = (V, E, w)$ , gdzie  $V$  oznacza zbiór wierzchołków grafu,  $E$  zbiór krawędzi grafu, zaś  $w$  określa funkcję kosztu krawędzi postaci:  $E \rightarrow R^+$ . Niech w zbiorze wierzchołków  $V$  określony będzie podzbiór  $S \subset V$  tego zbioru. Drzewem Steinera w grafie  $G$  nazywa się drzewo, które spina wszystkie wierzchołki podzbioru  $S$  wykorzystując możliwe krawędzie i wierzchołki grafu  $G$ . W swojej rozprawie doktorskiej Pan Piechowiak zajmuje się problemem wyznaczania drzew Steinera o minimalnym koszcie. Ponieważ problem minimalnego drzewa Steinera należy do klasy zadań typu NP-zupełne, stąd Autor nie zajmował się algorytmami dla znajdowania dokładnych rozwiązań, ale główny nacisk położył na analizę algorytmów heurystycznych

Na stronie 14, we wprowadzeniu do rozprawy Autor prezentuje stwierdzenie, które faktycznie wyjaśnia, co jest zakresem badawczym Autora. Stwierdzenie to precyzuje, że *„Optymalizacja komunikacji typu multicast polega na konstrukcji efektywnych algorytmów routingu, których zadaniem jest budowa drzewa dystrybucyjnego o minimalnym koszcie między węzłem nadawczym a grupą węzłów odbiorczych.”* Dlatego rozprawa Pana Piechowiaka dotyczy właśnie badania efektywnych algorytmów rozwiązania problemu drzewa Steinera, gdyż integralną częścią algorytmu routingu multicastowego jest rozwiązanie tego zadania optymalizacyjnego. Ponieważ algorytmy routingu muszą znajdować optymalne lub prawie optymalne rozwiązanie w czasie rzeczywistym, stąd algorytmy znajdujące dokładne rozwiązania problemu Steinera nie mogą być stosowane w algorytmach routingu multicastowego i do praktycznego stosowania w algorytmach routingu multicastowego mogą być wykorzystywane jedynie algorytmy heurystyczne.

W swojej rozprawie Pan Piechowiak rozważa dwa typy problemów optymalizacyjnych, mianowicie problem wyznaczania najtańszego drzewa Steinera bez ograniczeń oraz problem wyznaczania najtańszego drzewa z ograniczeniami, które określają maksymalne opóźnienie transferu informacji na ścieżce od węzła nadawczego do węzła odbiorczego.

Formułując problem badawczy Autor nie wyjaśnia jednak jak należy rozumieć koszt krawędzi dla rozważanych sieci. Opisując problem Steinera (str. 21 rozprawy) Autor przyjął, że *„każdemu łączu  $e_{ij} \in E$  przypisane są dwa parametry: koszt  $c_{ij}$  oraz opóźnienie  $d_{ij}$ ”*, ale nie zdefiniował, jak w praktycznych aplikacjach należy rozumieć koszt krawędzi (łącza)  $c_{ij}$ . W większości problemów sieciowych koszt łącza sieci jest przyjmowany jako proporcjonalny do długości tego łącza. Koszt łącza stanowi bardzo istotny parametr na etapie projektowania oraz



budowy sieci, gdyż ściśle wiąże się z sumarycznym kosztem inwestycyjnym tej sieci. Trudno jednak wyobrazić sobie rzeczywistą sytuację, że po zbudowaniu sieci, na etapie jej użytkowania, w algorytmach wyboru drogi rozważany jest koszt łącza jako parametr wpływający na wybór drogi w sieci. W rozległych sieciach pakietowych jako parametr kosztu dla jakiegoś łącza zwykło się najczęściej przyjmować opóźnienie przesyłania pakietów w tym łączu i właśnie dlatego ten parametr jest traktowany jako koszt łącza w sieci. Jednak w algorytmach rozważanych w rozprawie opóźnienie jest traktowane jako dodatkowe ograniczenie problemu znajdowania minimalnego drzewa Steinera, zaś koszt krawędzi (łącza)  $e_{ij}$  jest parametrem łącza zupełnie niezależnym od opóźnienia informacji w tym łączu. Brak wyjaśnień przez Autora, co uznaje On za koszt łącza w rzeczywistych sieciach, utrudnia interpretację wyników badań sieci zaprezentowanych w rozdziale 6 rozprawy, gdzie do generowania struktur topologicznych sieci przyjęto, że koszt każdego z łączy sieci przyjmuje wartość z zakresu od 10 do 1000 nie wyjaśniając jednak, co ten koszt oznacza.

O ile rozważanie kosztu drzewa Steinera w odniesieniu do aspektów wyznaczania optymalnych połączeń w układach scalonych lub w projektach płytek drukowanych modułów elektronicznych jest w pełni uzasadnione (kosztem krawędzi jest jej długość), o tyle dla rozległych sieci WAN jedynym rozsądnym, a zarazem łatwym do zinterpretowania, przypadkiem kosztu ścieżki jest koszt ścieżki liczony liczbą krawędzi wchodzących w skład tej ścieżki. Przypadek ten można określić następująco:  $\forall i, j \in V$  koszt  $c_{ij} = \infty$ , gdy krawędź  $e_{ij}$  nie występuje w grafie pierwotnym lub  $c_{ij} = 1$ , gdy krawędź  $e_{ij}$  występuje w grafie pierwotnym. Rozważania dotyczące kosztów łączy są bardzo istotne dla współczesnych sieci IP, które muszą być projektowane w sposób zapewniający obsługę wykładniczo rosnącego wolumen ruchu. Oznacza to, że zarówno warstwa szkieletowa sieci IP, jak i jej warstwa dostępową muszą udostępniać łącza o bardzo dużej wydajności i trudno jest, w takim przypadku, wydzielić część zasobów łącza, która jest wykorzystywana do obsługi jakiejś klasy ruchu.

W kontekście prezentowanych w niniejszej recenzji uwag jawią się dwa dosyć zasadnicze pytania: po pierwsze, czy w obecnych warunkach technologicznych (wykorzystanie w warstwie fizycznej sieci wydajnych urządzeń optycznych) za kryterium wyboru ścieżki dla realizacji usługi w sieci powinno się przyjmować koszt ścieżki, a po drugie: czy w aktualnych warunkach nie należałoby rozważać routingów w sieci, które optymalizują zasoby węzłów i nie zależą od wyboru krawędzi.

Zanim dokonam oceny rozprawy Pana Macieja Piechowiaka chciałbym zaprezentować swoje polemiczne uwagi, co do zakresu oraz możliwości wykorzystania

proponowanych w rozprawie tez oraz stwierdzeń odnosząc je do sposobu działania pakietowych sieci WAN.

Moim zdaniem, wykorzystanie problemu Steinera jako bazy dla tworzenia tablic routingu multicastowego ma sens wtedy, gdy rozważane sieci charakteryzują się niehierarchiczną strukturą logiczną, a ponadto łąca pomiędzy węzłami sieci mają bardzo ograniczoną i niewielką przepustowość (tak jak to było historycznie w sieci ARPA). Obecna architektura sieci IP wymusza zupełnie inne podejście do transmisji multicastowej, bo operatorskie sieci IP charakteryzują się hierarchiczną strukturą logiczną, w której przede wszystkim daje się wyróżnić bardzo wydajną warstwę szkieletową oraz wydajną warstwę dostępową. Obie te warstwy logiczne sieci służą do oferowania usług użytkownikom sieci (odbiorcom usług). To właśnie użytkownicy generują ruch w sieci i węzły szkieletowe służą jedynie tylko do transferu informacji od serwerów usług do użytkowników. Zatem również ruch multicastowy jest przenoszony przez routery warstwy szkieletowej do routera dystrybucyjnego (agregacyjnego), do którego dołączeni są użytkownicy znajdujący się w warstwie szkieletowej. Ponieważ warstwa dostępową zbudowana jest jako struktura gwiazdy lub gwiazdy hierarchicznej, stąd w warstwie tej do każdego użytkownika istnieje tylko jedna ścieżka, co oznacza, że tablica routingu jest trywialna i do jej wyznaczenia nie wymagane są żadne złożone metody wyznaczania drzewa multicastowego.

Podobne, jak dla warstwy dostępowej, rozważania dotyczą warstwy szkieletowej sieci. Przy dużej wydajności współczesnych routerów w warstwie szkieletowej sieci IP zainstalowanych jest stosunkowo niewiele routerów, bo współczesna technologia pozwala obsługiwać sumaryczny ruch szkieletowy wykorzystując, kilka lub co najwyżej kilkanaście, routerów, dla których drzewa multicastowe mogą być z góry predefiniowane. W świetle powyższych uwag badanie wydajności algorytmów multicastowych dla sieci zawierających kilkaset, czy nawet kilka tysięcy węzłów, jak to Autor czyni w rozdziale 6 swojej rozprawy jest nie do końca uzasadnione.

Moje zaprezentowane wyżej uwagi dotyczą oceny przydatności prezentowanych algorytmów wyznaczania multicastowych drzew dystrybucyjnych w zakresie ich stosowania we współczesnych sieciach IP. Uwagi te posiadają charakter uwag bardziej ogólnych i mają służyć jedynie do przemyślenia oraz uwzględnienia ich przez Doktoranta, na etapie dalszych badań. W niczym więc nie dotyczą oceny rozprawy w zakresie jej zawartości. Jeśli przyjmemy założenie, że badania Autora nie koncentrują się na sieciach IP, a dotyczą w ogólności abstrakcyjnych sieci pakietowych, dla których transfer ruchu w łącu generuje



zarówno pewien koszt związany z oferowaniem usługi, jak i opóźnienie transferu tej informacji, wówczas informacje to zawarte w rozprawie nabierają zupełnie innego znaczenia.

Abstrahując zatem od aktualnej specyfiki sieci IP, zaprezentowane w rozprawie Pana Piechowiaka rozważania stanowią znaczący przyczynek w zakresie analizy, ogólnych aspektów sieci pakietowych, gdyż prezentowana w rozprawie tematyka, stanowi interesujący wynik w zakresie szeroko rozumianej teorii sieci. Za samodzielny oraz oryginalny dorobek Autora należy uznać:

1. Własne algorytmy heurystyczne:

- a. heurystyczny algorytm STA (*Switched Tree Algorithm*), który służy do znajdowania najtańszego drzewa Steiner'a bez ograniczeń dotyczących opóźnienia transferu informacji. Ciekawym pomysłem Autora było wykorzystanie, na zasadzie przełączania, dwóch znanych algorytmów: SPT oraz PPH.
- b. Algorytm PLRA (*Path's Lagrange Relaxation Algorithm*), służący do rozwiązania problemu *Delay Constrained Least Cost*, w którym na bazie algorytmu LARAC (*Lagrange Relaxation based Aggregation Cost*) zaprezentowanego w pracy Juttnera, Szwiatowskiego, Mecsa, Rajko zaproponowano nowy sposób modyfikacji parametru  $\lambda$  relaksacji Lagrange'a, co pozwoliło zmniejszyć liczbę iteracji i przyspieszyć tym samym działanie algorytmu.
- c. Algorytm MLRA (*Multicast Lagrange Relaxation Algorithm*), który służy do dla wyznaczania drzewa najkrótszych ścieżek pomiędzy węzłem nadawczym oraz każdym węzłem ze zbioru węzłów multicastowych. Algorytm MLRA wykorzystuje algorytm PLRA.
- d. Algorytm KSPMA (*K-Shortest Paths Multicast Algorithm*), który podobnie jak algorytm MLRA, służy do wyznaczania drzewa najkrótszych ścieżek pomiędzy węzłem nadawczym oraz węzłami zbioru multicastowego. Algorytm KSPMA wykorzystuje algorytm znajdowania K najkrótszych ścieżek.
- e. Algorytm RDCMA stanowiący modyfikację algorytmu DCMA (*Fast Delay-Constrained Multicast Routing Algorithm*), który został zaprezentowany w pracy Hanga, Krunza i Chena.

- f. Algorytm OPTDCMA (*Optimal DCMA*), który służy do wyboru optymalnego rozwiązania z rozwiązań znalezionych zgodnie z algorytmami DCMA oraz RDCMA.
2. Przejrzysta oraz klarowna prezentacja poszczególnych algorytmów służących do znajdowania rozwiązań problemu Steinera, przedstawiona z wykorzystaniem pseudokodu.
  3. Rzeczowa oraz metodologicznie uzasadniona prezentacja metod modelowania sieci teleinformatycznych. Klarowna prezentacja metody generowania badanych, w ramach prowadzonych symulacji, struktur topologicznych sieci. Metoda generacji struktur topologicznych sieci pozwala wygenerować dowolną strukturę sieci, jednak trzeba podkreślić, że badając fizyczne struktury rzeczywistych sieci powinno się wprowadzić dodatkowe ograniczenie generowania jedynie struktur grafów planarnych.
  4. Oryginalna metoda rozmieszczania węzłów odbiorczych (*GroupRadius* oraz *GroupHighDegree*) zaprezentowana we wcześniejszych pracach Autora.
  5. Przedstawienie oraz przedyskutowanie w rozdziale 6 wyników badań symulacyjnych analizowanych sieci, które zostały już zaprezentowane w opublikowanych pracach Autora ([87, 88, 91, 92, 96]) oraz wyników badań symulacyjnych sieci dotychczas nigdzie nie publikowanych.
  6. Stworzenie oryginalnych własnych narzędzi i aplikacji dla symulacji badanych algorytmów w sieciach.

W swojej Rozprawie Doktorant odwołuje się do 128 prac, przy czym znacząca większość z nich, to są prace opublikowane w czasopiśmie naukowych lub opublikowana w materiałach konferencyjnych. Doktorant sam jest autorem lub współautorem 24 prac, opublikowanych w czasopiśmie krajowych lub prezentowanych na konferencjach krajowych i międzynarodowych. Wśród wszystkich prezentowanych prac jedynie 7 prac (poza własnymi pracami Autora) jest opublikowanych po roku 2005. Z tych 7-miu prac, dwie to: dokumentacja MySQL oraz draft standardu RFC4760, trzy to: monografie oraz jedna to: artykuł opublikowany w periodyku *Nature*. Tak mała liczba aktualnych prac z zakresu tematyki badawczej Autora może oznaczać, że tematyka ta nie znajduje się w aktualnym nurcie badań teorii sieci pakietowych, co może potwierdzać słuszność moich rozważań dotyczących usług multicastowych oferowanych w sieciach IP.



Praca jest napisana bardzo starannie, poprawnym i jasnym językiem. Czytając uważnie prace znalazłem jedynie dwie usterki językowe, mianowicie: na stronie 14 – umieszczone jest słowo *zralizowana*, a powinno być *zrealizowana*, zaś na stronie 78 – umieszczone jest słowo *koszów*, zamiast *kosztów*.

Oprócz drobnych wymienionych wyżej usterek językowych znalazłem w rozprawie Pana Piechowiak dwie niejasności. Pierwsza niejasność związana jest z oceną efektywności algorytmu CAO, zaś druga z wyznaczeniem złożoności obliczeniowej algorytmu DCSP.

Co do pierwszej niejasności, to jak należy rozumieć stwierdzenie zamieszczone na stronie 38 „*spośród prezentowanych w pracy algorytmów heurystycznych najlepszą efektywnością charakteryzuje się algorytm CAO (Constrained Adaptive Ordering)*” w kontekście informacji przedstawionej w tabeli 4.1. na stronie 57, gdzie podano, że algorytm CAO posiada złożoność wykładniczą.

Co do drugiej, to na str. 43 znajduje się stwierdzenie: „*złożoność obliczeniowa algorytmu DCSP wynosi  $O(K_T^2 |V|^2)$ . Jeżeli  $K_T$  przyjmie wartość maksymalną  $|K_T| = |V|$ , wtedy złożoność algorytmu wynosi  $O(|V|^3)$ ” Moim zdaniem powinno być  $O(|V|^4)$ .*

Porównanie wyników uzyskanych dla opracowanych przez Autora nowych algorytmów rozwiązywania problemu Steinera, które zostały zamieszczone w rozdziale 6 rozprawy prowadzi do wniosku, że proponowane przez Autora algorytmy generalnie nie wykazują przewagi nad znanymi już wcześniej algorytmami opracowanymi innych Autorów. Algorytm SPT uzyskuje jedynie najmniejszy koszt ścieżki w drzewie transmisji grupowej w zależności od liczby węzłów sieci (rysunek 6.6) oraz w zależności od współczynnika rozproszenia grupy (rysunek 6.7). Wyniki uzyskane z symulacji algorytmów RDCMA oraz OPTDCMA pokazują, że algorytmy te są generalnie gorsze od algorytmów KPP oraz BSMA. Jedynie algorytmy MLRA oraz KSPMA generują rozwiązania (część 6.4 rozprawy), które są lepsze od rozwiązań uzyskanych z wykorzystaniem innych algorytmów.

Wszystkie uwagi przedstawione w mojej recenzji miały jedynie charakter dyskusyjny a ich zasadniczym celem było usprawnienie treści recenzowanej rozprawy. Uwagi te w niczym nie umniejszają wartości całej pracy. Dlatego, po wnikliwym przestudiowaniu rozprawy Pana Piechowiaka, mogę z całą stanowczością stwierdzić, że rozprawa spełnia, nawet z pewną nadwyżką, wymagania określone przez odpowiednie przepisy i tym samym wnioskuję o dopuszczenie jej do publicznej obrony.

