

OPTIMALIZACJA PRZEPUSTOWOŚCI SIECI KOMPUTEROWYCH ZA POMOCĄ ALGORYTMÓW GENETYCZNYCH

Andrzej SZYMONIK, Krzysztof PYTEL

Streszczenie: W złożonych sieciach komputerowych istnieje problem doboru przepustowości łączy pomiędzy poszczególnymi węzłami sieci. Optymalny dobór przepustowości łączy pozwala na minimalizację opóźnienia pakietów przy jednoczesnym ograniczeniu kosztów utrzymania łączy. Zadanie to jest złożone obliczeniowo zwłaszcza dla sieci o dużej liczbie węzłów. W artykule zaproponowano wykorzystanie algorytmów genetycznych do minimalizacji średniego opóźnienia pakietu w sieci komputerowej. Dla zapewnienia większej zbieżności metody zastosowano zmodyfikowany algorytm genetyczny, w którym sterownik rozmyty ocenia poszczególnych osobników w populacji. Na podstawie oceny poszczególnych osobników, sterownik wpływa na przebieg algorytmu poprawiając jego skuteczność.

Słowa kluczowe: optymalizacja, algorytmy genetyczne, sieci komputerowe.

1. Wstęp

Optymalizacja przepustowości łączy w sieciach komputerowych posiada duże znaczenie praktyczne. Dane w sieci powinny być przesyłane możliwie szybko. Wraz ze wzrostem przepustowości łączy, rośnie również koszt ich eksploatacji. Podstawowym celem optymalizacji sieci jest zapewnienie jej sprawnego funkcjonowania przy jednoczesnym ograniczeniu kosztów eksploatacji. Dla zapewnienia większej zbieżności algorytmów zostały one zmodyfikowane poprzez zastosowanie sterownika rozmytego. Zadaniem sterownika rozmytego jest wyznaczenie kierunku, w którym powinna przebiegać ewolucja, dzięki czemu algorytm może uniknąć błędzenia w przestrzeni rozwiązań. Ponadto sterownik pozwala na wykorzystanie informacji o przebiegu optymalizacji zgromadzonych przez algorytm we wcześniejszych pokoleniach. Do testów wykorzystano zadanie minimalizacji średniego opóźnienia pakietu. Zadanie to opisane zostało w pracy [2].

W dalszych rozważaniach przyjmujemy, że dana jest sieć, reprezentowana przez graf właściwy i zorientowany $G(N_w, L, C_l)$, gdzie: N_w jest zbiorem węzłów reprezentujących routery w sieci, L jest zbiorem łuków reprezentujących łącza. Wagi wektora łuków C_l mogą reprezentować odległości, opóźnienia, przepustowości lub inną wybraną miarę.

Oznaczmy:

$(x, y) \in L$ - łuk o początku w węźle x i końcu w węźle y , gdzie $x, y \in N_w$,

$c(x, y) \in C_l$ - przepustowość łuku (x, y) ,

$f(x, y)$ - przepływ w łuku (x, y) .

2. Zadanie minimalizacji średniego opóźnienia pakietu

Oznaczmy przez $R = [r_{ij}]_{m \times m}$ macierz z zewnątrz wprowadzanych natężeń oraz r_{ij} średnie natężenie strumienia danych kierowanych z węzła i do węzła j . Poszczególne składniki r_{ij} możemy ponumerować od 1 do q . Z k -tym składnikiem związana będzie para węzłów s_k i u_k jako źródło i ujście. Przez r_k oznaczmy średnie natężenie k -tego składnika tzn. $r_k = r_{ij}$ gdzie $i = s_k$, $j = u_k$. Przepływem wieloskładnikowym w sieci S realizującym macierz z zewnątrz wprowadzanych natężeń R nazywamy zespół funkcji

$$f^k : L \rightarrow R^+ \cup \{0\} \quad k = 1, \dots, q, \quad (1)$$

których wartości $f^k(x, y)$ $k = 1, \dots, q$ przypisane poszczególnym łukom $(x, y) \in L$ spełniają następujący układ warunków:

$$\sum_{y \in A(x)} f^k(x, y) - \sum_{y \in B(x)} f^k(y, x) = \begin{cases} r_k & \text{dla } x = s_k \\ -r_k & \text{dla } x = u_k \\ 0 & \text{w pozostałych przypadkach} \end{cases} \quad (2)$$

dla każdego $x \in N$, $k = 1, \dots, q$.

Zbiór $A(x) = \{y : y \in N \text{ i } (x, y) \in L\}$ jest zbiorem węzłów sieci, do których prowadzą łuki wychodzące z węzła x , zbiór $B(x) = \{y : y \in N \text{ i } (x, y) \in L\}$ jest zbiorem węzłów sieci, do których prowadzą łuki skierowane do węzła x . Problem minimalizacji średniego opóźnienia pakietu polega na minimalizacji funkcji

$$T = \frac{1}{\gamma} \sum_{(x,y) \in L} \frac{f(x, y)}{c(x, y) - f(x, y)}, \quad (3)$$

gdzie:

$\gamma = \sum_{i=1}^q r_i$ - jest natężeniem sumarycznego strumienia pakietów wprowadzanych do sieci, przy spełnieniu warunków:

$$\sum_{y \in A(x)} f^k(x, y) - \sum_{y \in B(x)} f^k(y, x) = \begin{cases} r_k & \text{dla } x = s_k \\ -r_k & \text{dla } x = u_k \\ 0 & \text{w pozostałych przypadkach} \end{cases} \quad (4)$$

dla każdego $x \in N$, $k = 1, \dots, q$,

$f^k(x, y) \geq 0$ dla każdego $(x, y) \in L$, $k = 1, \dots, q$,

$f(x, y) \leq c(x, y)$ dla każdego $(x, y) \in L$.

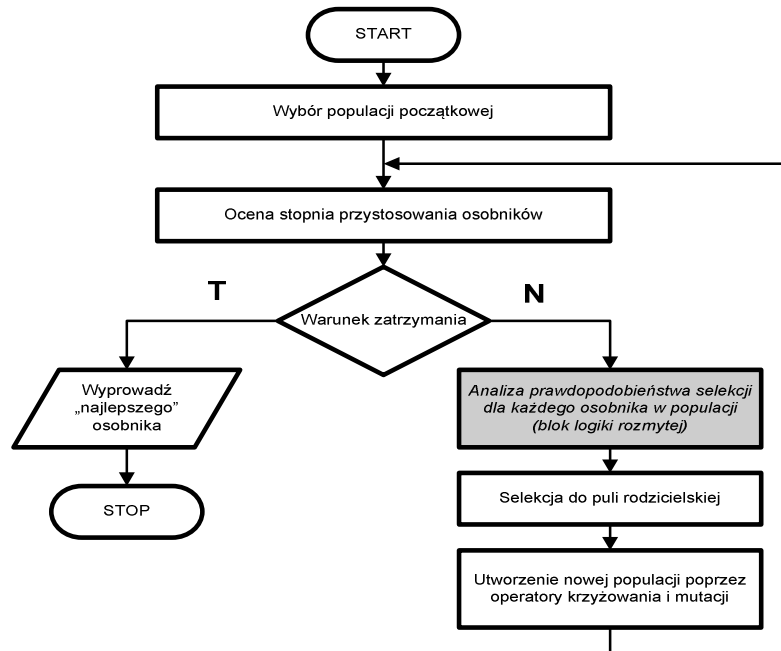
Rozwiązanie zadania polega na wyszukaniu optymalnych ścieżek, po których przesyłane będą poszczególne składniki przepływu. Wybór odpowiednich ścieżek powinien uwzględniać wielkości składników przepływu, przepustowości poszczególnych łączy oraz założenie, że w danym momencie z każdego łącza może korzystać co najwyżej jeden składnik przepływu. Poszczególne składniki przepływu mogą wzajemnie oddziaływać na siebie. Zmiana trasy jednego z nich, może powodować konieczność zmiany również innych składników. Rozwiązanie zadania za pomocą klasycznych metod algorytmicznych jest trudne również ze względu na dużą przestrzeń rozwiązań. Ponadto optymalizowana funkcja jest nieciągła ze względu na skokowe zmiany wartości przepustowości łączy.

3. Algorytm genetyczny wykorzystany do optymalizacji średniego opóźnienia pakietu

Algorytmy genetyczne umożliwiają wyszukiwanie przybliżonych rozwiązań złożonych problemów optymalizacyjnych. Swoje istnienie zawdzięczają obserwacji i próbie naśladowania naturalnych procesów zachodzących w świecie organizmów żywych - ewolucji i związanej z nią naturalnej selekcji występującej w populacjach żywych osobników. Algorytmy genetyczne operują na populacjach rozwiązań nazywanych osobnikami. Osobnik reprezentuje kompletne rozwiązanie danego problemu optymalizacyjnego. Jakość osobników oceniana jest za pomocą nieujemnej funkcji przystosowania. Osobniki o wysokim przystosowaniu spełniają warunki zadania w wyższym stopniu niż osobniki o niskim przystosowaniu. Osobniki kodowane są w struktury zwane ciągami kodowymi. Algorytm genetyczny poszukuje coraz lepszych rozwiązań, generując kolejne pokolenia osobników, poprzez przeprowadzanie operacji na ciągach kodowych. Pierwsze pokolenie tworzy się losowo lub wybiera w inny sposób. Dla każdego kolejnego pokolenia wybiera się pewną liczbę osobników (proces selekcji). Osobniki o wysokiej wartości funkcji przystosowania mają większą szansę na wybór. Osobniki te są kojarzone w pary (losowo), po czym następuje wymiana (wymieszanie się) materiału genetycznego rodziców, w wyniku czego otrzymujemy dwa nowe ciągi kodowe (odpowiadające genotypowi dzieci). Więcej informacji na temat algorytmów genetycznych można znaleźć w literaturze [1], [3].

Algorytmy genetyczne nie gwarantują znalezienia rozwiązania optymalnego, lecz zwykle pozwalają na znalezienie rozwiązania bliskiego optymalnemu. Czas działania algorytmu, zwłaszcza w przypadku konieczności przeszukiwania dużych przestrzeni rozwiązań może być długi. W eksperymencie rozwiązywano zadanie testowe przy pomocy algorytmu genetycznego wykorzystującego sterownik rozmyty do modyfikowania prawdopodobieństwa selekcji. Schemat blokowy zmodyfikowanego algorytmu genetycznego pokazany został na rys.1.

Sterownik rozmyty (na schemacie blokowym opisany kursywą na ciemnym tle) ocenia każdego osobnika w populacji i wyznacza dla niego współczynnik adaptacji prawdopodobieństwa selekcji (wps_k).



Rys.1. Schemat blokowy zmodyfikowanego algorytmu genetycznego
Źródło: opracowanie własne

Sterownik modyfikuje prawdopodobieństwo selekcji w oparciu o zasady:

- zwiększ prawdopodobieństwo selekcji dla osobników o wartości funkcji przystosowania powyżej średniej w pokoleniach, w których średnia wartość funkcji przystosowania rośnie w stosunku do pokolenia poprzedniego,
- pozostaw bez zmian prawdopodobieństwo selekcji dla osobników o wartości funkcji przystosowania równej średniej w pokoleniach, w których średnia wartość funkcji przystosowania nie zmienia się w stosunku do pokolenia poprzedniego,
- zmniejsz prawdopodobieństwo selekcji dla osobników o wartości funkcji przystosowania poniżej średniej w pokoleniach, w których średnia wartość funkcji przystosowania zmniejsza się w stosunku do pokolenia poprzedniego.

Zmodyfikowane prawdopodobieństwo selekcji osobnika k obliczane jest na podstawie wzoru:

$$ps'_k = ps_k * wps_k \quad \text{dla } k = 1, \dots, N \quad (5)$$

gdzie:

ps'_k - zmodyfikowane prawdopodobieństwo selekcji osobnika k ,

ps_k - prawdopodobieństwo selekcji osobnika k ,

wps_k - współczynnik prawdopodobieństwa selekcji osobnika k wyznaczony przez sterownik rozmyty.

Więcej informacji o algorytmach genetycznych, w których zastosowano sterownik rozmyty do modyfikacji prawdopodobieństwa selekcji osobników można znaleźć w pracy [4], [5].

4. Eksperyment

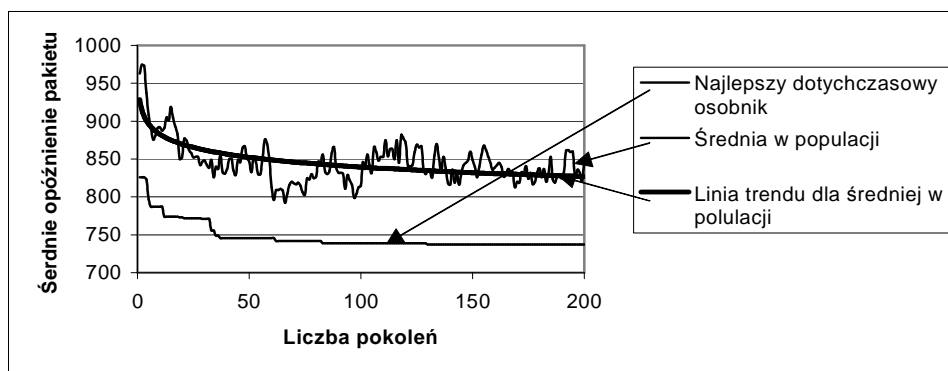
Eksperyment polegał na optymalizacji przepływu wieloskładnikowego bez rozgałęzień. Jako kryterium optymalizacji przyjęto średnie opóźnienie pakietu. W zadaniu testowym optymalizowano sieć składającą się z 58 węzłów. Łuki pomiędzy węzłami posiadają różną przepustowość przy czym nie wszystkie węzły połączone są bezpośrednio. Z zewnątrz wprowadzono do sieci obciążenie z 9 składników o różnym natężeniu. Dla każdego ze składników zdefiniowano po 8 dopuszczalnych tras.

Zadanie testowe zostało rozwiązane za pomocą algorytmu przeszukującego wszystkie możliwe kombinacje rozwiązań (metoda „brute force”). Wyniki uzyskane przez ten algorytm posłużyły jako punkt odniesienia i umożliwiły porównanie z wynikami uzyskanymi przez zaproponowany, zmodyfikowany algorytm genetyczny. Zadanie rozwiązano również przy pomocy klasycznego algorytmu genetycznego, nie korzystającego ze sterownika rozmytego. Pozwoliło to na ocenę skuteczności modyfikacji algorytmu.

W eksperymencie przyjęto następujące parametry algorytmów genetycznych:

- geny reprezentowane są przez ciągi 9 liczb z przedziału $<0, 7>$,
- funkcja przystosowania dla osobnika obliczana na podstawie zależności (3),
- prawdopodobieństwo krzyżowania 0,8,
- prawdopodobieństwo mutacji 0,15,
- liczebności populacji 25.

Wykres na rys. 2. pokazuje wartość funkcji przystosowania osobnika (średniego opóźnienia pakietu) oraz średnią wartość przystosowania w populacji dla pierwszych 200 pokoleń algorytmu.



Rys. 2. Wartość funkcji przystosowania osobników (średniego opóźnienia pakietu) w kolejnych pokoleniach algorytmu

Źródło: opracowanie własne

5. Wnioski

Algorytm genetyczny znalazł rozwiązanie testowego problemu. Pomimo, iż zwykle algorytmy genetyczne pozwalają tylko na znalezienia rozwiązania bliskiego optymalnemu, zaproponowane zadanie testowe okazało się na tyle łatwe, że algorytm genetyczny znalazł rozwiązanie optymalne (takie samo jak algorytm „brute force”).

Średnia wartość średniego opóźnienia pakietu (funkcji optymalizowanej) w pierwszym pokoleniu (wybrany losowo) wynosiła 963, a wartość optymalna 732. Oznacza, że

algorytm genetyczny na drodze ewolucji znalazł rozwiązanie o około 24% lepsze od rozwiązania początkowego (wartość odniesiona do wartości początkowej).

Algorytmy genetyczne mogą być zastosowane do zadań optymalizacji sieci o większej liczbie węzłów i składników, lecz wymaga to dłuższego czasu działania. Możliwe jest również korzystanie z innych kryteriów optymalizacji, np. przepustowości sieci, kosztu budowy, kosztu korzystania z usług komunikacyjnych itp.

Problem optymalizacji przepływu wieloskładnikowego jest zadaniem trudnym do rozwiązania metodami tradycyjnymi ze względu na bardzo dużą przestrzeń rozwiązań. Dla przyjętych w eksperymencie warunków czas potrzebny na obliczenie wszystkich możliwych rozwiązań (metodą „brute force”) wynosił około 3,5 godziny. Zastosowanie zmodyfikowanego algorytmu pozwoliło na skrócenie tego czasu do kilkunastu sekund. Oba algorytmy uruchomione były na tym samym komputerze.

Przepustowości łącza oferowanych przez firmy telekomunikacyjne są wielkościami zmieniającymi się skokowo, ponadto poszczególne składniki przepływu mogą wykorzystywać te same łącza, co powoduje, że funkcja optymalizowana nie jest funkcją ciągłą. Dlatego w eksperymencie zastosowano model mutacji, w którym mutowany gen przyjmuje dowolną spośród dopuszczalnych wartości genu z jednakowym prawdopodobieństwem.

Zastosowanie sterownika rozmytego modyfikującego prawdopodobieństwo selekcji pozwoliło na zmniejszenie o około 15% liczby pokoleń potrzebnych do uzyskania rozwiązania, co w konsekwencji spowodowało skrócenie czasu działania algorytmu o około 6,5% w stosunku do klasycznego algorytmu genetycznego (nie korzystającego ze sterownika rozmytego).

Zaproponowana metoda może być zastosowana do innych zadań związanych z optymalizacją przepływów, np. towarów, wyrobów itp.

Literatura

1. Goldberg D.E.: Algorytmy genetyczne i ich zastosowania. Wydawnictwa Naukowo Techniczne, Warszawa 1998.
2. Kasprzak A.: Rozległe sieci komputerowe z komutacją pakietów. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1999.
3. Michalewicz Z.: Algorytmy genetyczne + struktury danych = programy ewolucyjne, Wydawnictwa Naukowo Techniczne, Warszawa 1999.
4. Pytel K., Kluka G.: Wykorzystanie logiki rozmytej do wspomagania ewolucji osobników w algorytmach genetycznych. w Studia z Automatyki i Informatyki tom 27, Poznańskie Towarzystwo Przyjaciół Nauk, Poznań 2002.
5. Pytel K., Kluka G., Kisilewicz J.: Zastosowanie algorytmów genetycznych do optymalizacji kosztów eksploatacji sieci komputerowych. Proceedings of 1st Polish and International PhD Forum-Conference on Computer Science, Łódź 2005.

Dr hab. inż. Andrzej SZYMONIK prof. PŁ
Katedra Zarządzania Produkcją
Wydział Organizacji i Zarządzania
Politechnika Łódzka
ul. Wólczańska 215, 90-924 Łódź
Dr inż. Krzysztof PYTEL
Akademia Humanistyczno-Ekonomiczna w Łodzi