

УДК 004.7

С.Д. Погорілий, Р.В. Білоус
Київський національний університет імені Тараса Шевченка
sdp77@i.ua, romanrpd@gmail.com

Застосування паралельних версій генетичних алгоритмів в комп'ютерних мережах

Запропоновано підходи до створення паралельних версій генетичних алгоритмів балансування навантаження та маршрутизації в комп'ютерних мережах. Описано особливості застосування генетичних операцій кросоверу та мутації а також стратегій відбору. Формалізовано загальну схему розподіленого генетичного алгоритму та описано особливості застосування операції міграції. Проведено аналіз та одержано асимптотичну оцінку складності паралельного генетичного алгоритму.

Генетичний алгоритм, мережа, маршрутизація, балансування навантаження в мережі

Вступ

Однією з найпоширеніших функціональних задач мереж є задача про найкоротші шляхи: пошук шляху між двома визначеними вершинами графа, який відповідає найменшому значенню певного функціонала за визначеним критерієм. Це фундаментальна задача багатьох застосувань у сферах транспорту, маршрутизації, комунікації. Існує низка класичних алгоритмів розв'язання цієї задачі (Белмана-Форда, Дейкстри) [1, 2].

Класичні алгоритми розв'язання задач маршрутизації оперують лише одним параметром оптимізації – вагою (ціною) шляху, що виражає сукупність його адитивних характеристик. Проте зазвичай існує декілька параметрів, які характеризують кожен гілку мережі, наприклад пропускна здатність, затримка, швидкість передачі, навантаження, надійність які можна поділити на адитивні та неадитивні. Таким чином у сучасних мережах з'явилась необхідність розв'язання задачі про найкоротші шляхи з кількома критеріями оптимізації [3]. Класичні алгоритми не застосовні до задач такого типу, що належать до класу NP-складних. Обчислювальні затрати на розв'язання таких задач експоненційно зростають із ростом розмірності оброблюваних графів. Тому виникає необхідність формування нових підходів та алгоритмів розв'язання задач пошуку оптимальних шляхів з багатьма критеріями, одним із яких є генетичний підхід.

В роботах [4, 5] запропоновано підходи використання генетичних алгоритмів для розв'язання задач маршрутизації та розподілення навантаження в комп'ютерних мережах. Описані підходи дозволяють застосувати комплексні метрики маршрутів в мережі а їх робота базується на сучасних алгоритмах маршрутизації. Адаптивні імовірності мутації та кросоверу а також особливості процедури відбору найкращих особин дозволяють ефективно використовувати запропоновані алгоритми в сучасних

комп'ютерних мережах.

В роботі [6] проведені теоретичні оцінки складності генетичного алгоритму та експериментально перевірено залежність його збіжності від вхідних параметрів задачі. Запропоновано чисельні оцінки ефективності генетичного підходу в порівнянні з класичними алгоритмами маршрутизації в мережах.

В результаті проведених експериментальних досліджень запропоновано найбільш ефективні значення параметрів генетичного алгоритму: розмір популяції, імовірності кросоверу та мутації залежно від розміру вхідних даних задачі.

В мережах з великою кількістю вузлів та зв'язків виникає необхідність покращення часових характеристик роботи запропонованих генетичних алгоритмів Одним із найбільш доцільних і перспективних шляхів вирішення цього завдання є формування їх паралельних версій.

Постановка задачі

Особливістю роботи генетичних алгоритмів в мережах є їх децентралізованість за даними. Кількість обчислювальних вузлів в таких системах може співпадати з кількістю точок обміну трафіком. Обчислювальна система є розподіленою і в загальному випадку слабко-зв'язаною, оскільки ніяких додаткових умов щодо характеристик мережових з'єднань не накладається. Розглянемо більш детально модель генетичного алгоритму та особливості його застосування в розподілених системах. Схема роботи послідовного генетичного алгоритму може бути описана як:

Початок

Формування початкової популяції;

Повторити

Випадковий вибір М особин;

Застосування операцій кросоверу та мутації;

Формування наступного покоління:

- обчислення функції пристосування особин популяції;
- визначення «найкращих особин»;
- застосування операції відбору.

Поки не досягнуто умови завершення алгоритму

Кінець.

Задача дослідження полягає у формуванні підходу до створення паралельних версій генетичних алгоритмів в гетерогонній слабко-зв'язаній обчислювальній системі, якою є комп'ютерна мережа, створенні формалізованої схеми розподіленого генетичного алгоритму, обґрунтуванні його збіжності та особливостей застосування операцій кросоверу та мутації.

Особливості генетичних операцій

При формуванні операцій кросоверу та мутації слід враховувати два аспекти:

- збереження послідовностей алелів, що відповідають позитивному приросту функції пристосування – для забезпечення збіжності алгоритму вцілому;
- ефективне створення нових послідовностей алелів, для забезпечення різноманітності хромосом в межах покоління та уникнення збіжності до локального екстремуму цільової функції.

Забезпечення описаних вище вимог відбувається за рахунок адаптивної зміни імовірностей операцій кросоверу p_c та мутації

p_m хромосом генетичного алгоритму [7].

Визначимо імовірності кросоверу та мутації довільної хромосоми Y :

$$p_c = \begin{cases} p_1 (F_{\max} - F(Y)) / (F_{\max} - \bar{F}), & F(Y) \geq \bar{F} \\ p_2, & F(Y) < \bar{F} \end{cases}$$

$$p_m = \begin{cases} p_3 (F_{\max} - F(Y)) / (F_{\max} - \bar{F}), & F(Y) \geq \bar{F} \\ p_4, & F(Y) < \bar{F} \end{cases}$$

F_{\max} – значення функції пристосування найкращої хромосоми популяції, \bar{F} – середнє значення функції пристосування хромосом в межах популяції, $0 < p_1, p_2, p_3, p_4 \leq 1$.

Для забезпечення різноманітності хромосом в межах популяції, імовірність кросоверу та мутації найменш пристосованих хромосом повинна бути високою, тому в

запропонованому алгоритмі $p_1 = p_2 = 1$,

$p_3 = p_4 = 0.5$. Такий адаптивний підхід до операцій кросоверу та мутації забезпечує еволюцію популяції та покращує збіжність алгоритму. З іншого боку, при зростанні кількості однакових хромосом в популяції імовірності мутації та кросоверу зростають, таким чином запропонований алгоритм уникає передчасної збіжності до локального екстремуму [8].

Операція відбору (відтворення) спрямована на покращення середньої якості популяції генетичного алгоритму. Така можливість досягається за рахунок того, що найкращі особини мають найбільші шанси переходу до наступного покоління. Запропонований генетичний алгоритм використовує метод турнірного відбору.

Нехай N – кількість особин популяції. Характеристика кожної i -ї особини популяції (фенотип) виражається значенням функції пристосованості $F(i)$. Для відбору певного індивіда створюється група з M ($M \geq 2$) випадково обраних особин. Індивід з найбільшою пристосованістю в групі відбирається, решта – відкидається. Перевагами такого підходу є:

- відсутність збіжності до локального екстремуму;
- відсутність стагнації;
- відбір не вимагає глобального впорядкування усіх особин покоління;
- немає необхідності явного обрахунку функції пристосованості.

Оскільки для операції відбору використовується турнірний метод немає необхідності обрахунку функції пристосованості вцілому, що значно спрощує саму процедуру відбору. Для порівняння особин в групі достатньо порівняти значення відхилення кожної особини і залишити одну хромосому, відхилення якої в групі є мінімальним.

Розподілена модель генетичного алгоритму

Користуючись описаними в [9] підходами до класифікації паралельних версій генетичних алгоритмів в роботі пропонується застосування розподіленої (острівної) моделі.

В цьому випадку популяція розбивається рівномірно на n окремих частин (ареалів) між обчислювальними вузлами системи. При такому підході кожен вузол використовує описаний послідовний генетичний алгоритм, оперуючи даними відповідного ареалу. Таким чином загальна схема генетичного алгоритму на окремому вузлі системи не змінюється, забезпечуючи його збіжність. Імовірності кросоверу та мутації залежать від характеристик

хромосом окремого ареалу.

Кожен ареал відділений від інших («географічно») і особини взаємодіють лише в його межах. Внаслідок такого відокремлення популяції виникає необхідність синхронізації генетичного матеріалу окремих ареалів для збереження збіжності алгоритму в цілому. Введемо додаткову операцію міграції, що полягає у переході (копіюванні) окремих особин різних ареалів. Оскільки обчислювальна система є гетерогенною особини можуть мігрувати між довільними ареалами.

Описана операція міграції характеризується низкою параметрів [10]:

- топологією зв'язків між ареалами;
- розміром, що визначає кількість особин N_m , які мігрують між ареалами;
- схемою міграції, що визначає які саме особини мігрують (найкращі, найгірші, випадкові), а також особини які вони заміщують в іншому ареалі;
- інтервалом T_m , що визначає частоту міграцій.

Важливим параметром операції міграції є також характер взаємодії ареалів: синхронний чи асинхронний. Оскільки обчислювальна система є принципово гетерогенною доцільним є використання асинхронної взаємодії – при цьому уникається залежність від найменш продуктивних обчислювальних вузлів.

В запропонованому підході для застосування операції міграції обираються найкращі особини певного ареалу. Така стратегія перемішування хромосом поточного покоління дозволяє значно зменшити імовірність збіжності генетичного алгоритму до локального екстремуму а також покращити часові характеристики роботи алгоритму завдяки поширенню послідовностей алелів, що відповідають високому значенню функції пристосування. Решта параметрів міграції обирається залежно від навантаження на канали зв'язку та часові затрати на обмін хромосомами між окремими обчислювальними вузлами. Ці параметри змінюються динамічно залежно від поточних вимог алгоритму та стану системи.

Схема розподіленого генетичного алгоритму виглядає наступним чином [11, 12]:

Початок

Формування початкової популяції;

Розподіл популяції на ареали;

Повторити

Для кожного ареалу:

Випадковий вибір M особин;

Застосування операцій кросоверу та мутації;

Формування наступного покоління:

- обчислення функції пристосування особин популяції;

– визначення «найкращих особин»;

– застосування операції відбору.

Асинхронна операція міграції;

Поки не досягнуто умови завершення алгоритму

Кінець.

Оцінка складності розподіленого генетичного алгоритму

Однією із найважливіших характеристик паралельного алгоритму є його ефективність в порівнянні з послідовним. Відносно покращення часу роботи алгоритму S визначається наступним чином:

$$S = \frac{T_s}{T_p}, \quad (1)$$

де T_s та T_p визначають час роботи послідовного та паралельного алгоритмів відповідно. Обчислювальні затрати паралельного генетичного алгоритму можна розподілити на такі частини:

λ – послідовні обчислення, що не підлягають розпаралелюванню (формування початкового покоління, обчислення критеріїв зупинки алгоритму);

γ – послідовні обчислення, що підлягають розпаралелюванню (генетичні операції кросоверу, мутації та відбору);

δ – додаткові часові затрати, пов'язані із взаємодією обчислювальних вузлів (міграція).

Нехай сумарний розмір популяції рівний N . Позначимо кількість обчислювальних вузлів системи як p . Введемо деякі додаткові позначення:

N_m – розмір міграції,

ω_m – частота міграції.

Користуючись описаними позначеннями рівність (1) може бути уточнена наступним чином:

$$S(N, p) = \frac{\lambda(N) + \gamma(N)}{\lambda(N) + \gamma(N) / p + \delta(N_m, \omega_m, p)}$$

Висновки

Запропоновано підхід до формування паралельних версій генетичних алгоритмів в комп'ютерній мережі. Описано особливості та обґрунтовано застосування генетичних операцій кросоверу та мутації, механізмів відбору. Для розподіленого генетичного алгоритму сформовано стратегію взаємодії ареалів його популяції за рахунок введення додаткової генетичної операції міграції. Описано параметри міграції та запропоновано особливості її використання в складних гетерогенних обчислювальних системах. Запропоновано схему

розподіленого генетичного алгоритму та його зв'язок із послідовним генетичним алгоритмом. Використання гомогенного підходу дозволяє застосовувати однакову репродуктивну схему для кожного окремого ареалу, що в свою чергу значно спрощує створення реалізації паралельної версії генетичного алгоритму.

Оцінено асимптотичну складність паралельного генетичного алгоритму та описано її залежність від параметрів генетичних операцій та кількості обчислювальних вузлів системи.

Розроблені методики доцільно

використовувати при створенні нових мережевих алгоритмів маршрутизації та балансування навантаження, що базуються на генетичних алгоритмах. Збіжність алгоритму залежить від параметрів операції міграції та генетичних операцій послідовної схеми.

Сформовані підходи дозволяють значно спростити (а для деяких окремих випадків є єдиним варіантом) розв'язання задач маршрутизації та балансування навантаження у складних комп'ютерних та транспортних системах.

Список літератури

1. Седжвик Р. Фундаментальные алгоритмы на C++. Алгоритмы на графах / Р. Седжвик; пер. с англ. – С-Пб : ООО «ДиаСофтЮП», 2002. – 496 с.
2. Кормен Т. Алгоритмы: построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест. – М: Центр непрер. математического образования, 2000. – 960 с.
3. P. Sateesh Kumar Genetic zone routing protocol / P. Sateesh Kumar, S. Ramachandram // Journal of Theoretical and Applied Information Technology. – 2008. – P. 162-171.
4. Погорілий С.Д. Генетичний алгоритм розв'язання задачі маршрутизації в мережах / С.Д. Погорілий, Р.В. Білоус // Проблеми програмування. – 2010. – № 2-3; Спец. вип. – С.171-178.
5. Girish M K Formulation of the load engineering problems in MPLS based IP networks / M K. Girish, B. Zhou, Hu J Q // The Fifth IEEEISCC. Antibes, France. – 2000. – P. 214-219.
6. Білоус Р.В. Особливості прикладного застосування генетичного алгоритму пошуку оптимальних шляхів на графі / Р.В. Білоус, С.Д. Погорілий // Реєстрація, зберігання і оброб. даних. – 2010. – Т. 12, № 2. – С. 81–87.
7. Srinivas M. Adaptive probabilities of crossover and mutation in genetic algorithms / Srinivas M, Patnaik L. M. // IEEE Transactions on Systems, MAN and Cybernetics. – 1994. – 24 (4). – P. 656-667.
8. Gonen. B. Genetic Algorithm Finding the Shortest Path in Networks / B. Gonen. – Reno: University of Nevada, 2006.
9. Nowostawski M. Parallel Genetic Algorithm Taxonomy. Proceedings of the Third International conference on knowledge-based intelligent information engineering systems / M. Nowostawski, R. Poli. – 1999. – P.88-92.
10. E. Alba Heterogeneous Computing and Parallel Genetic Algorithms / E. Alba, A.J. Nebro, J.M. Troya // Journal of Parallel and Distributed Computing. – 2002. – 62. – P.1362-1385.
11. Cantu-Paz E. A Survey of Parallel Genetic Algorithms / E. Cantu-Paz. // Calculateurs Paralleles, Reseaux et Systems Repartis. – 1998. – V. 10 – No. 2. – P. 141-171.
12. Lee Y. A constrained multipath Traffic engineering scheme for MPLS networks. ICC'02 / Lee Y, Seok Y, Choi Y. – New York, 2002. – P. 2431-2436.

Надійшла до редколегії 10.10.2011

С. Д. ПОГОРЕЛЫЙ, Р. В. БИЛОУС

Киевский национальный университет имени Тараса Шевченко

Применение параллельных версий генетических алгоритмов в компьютерных сетях

Предложены подходы к созданию параллельных версий генетических алгоритмов балансировки нагрузки и маршрутизации в компьютерных сетях. Описаны особенности применения генетических операций кроссовера и мутации, а также стратегии отбора. Формализована общая схема распределенного генетического алгоритма и описаны особенности применения операции миграции. Проведен анализ и получена асимптотическая оценка алгоритма.

генетический алгоритм, сеть, маршрутизация, балансировка нагрузки в сети

S. D. POGORILY, R. V. BILOUS

Kyiv National Taras Shevchenko University

Application of Parallel Versions of Genetic Algorithms in Computer Networks

New approaches for creating parallel versions of genetic algorithms for load balancing and routing in computer networks were proposed. Features of genetic operations of crossover and mutation and selection strategies were described. The overall scheme of distributed genetic algorithm was formalized and features of migration operation were described. Analysis was performed and asymptotic bounds on the complexity of parallel genetic algorithm were obtained.

genetic algorithm, network, routing, network load balancing