УДК 519.17+519.24

Генетический алгоритм структурной оптимизации сетей с применением подхода кумулятивного уточнения границ надёжности

Д. А. Мигов, К. А. Нечунаева, А. С. Родионов *

В статье предлагается новый алгоритм оптимизации структуры сети по критерию надёжности в условиях различных ограничений. Рассматриваются сети с абсолютно надёжными узлами и ненадёжными каналами связи. Под надёжностью сети понимается вероятность связности соответствующего случайного графа, расчёт которой представляет собой NP-трудную задачу. Разработанный алгоритм оптимизации является генетическим, его операторы позволяют отсеивать непригодные особи на более ранних этапах за счёт использования кумулятивного уточнения верхней и нижней границ надёжности сети. Проведённые эксперименты подтвердили эффективность предлагаемого подхода.

Ключевые слова: надёжность сети, случайный граф, оптимизация структуры сети, генетический алгоритм, метод факторизации, диаметр сети.

1. Введение

В процессе эксплуатации элементы сетей связи могут подвергаться отказам по различным причинам: поломка, износ, природные явления, результат преднамеренных действий злоумышленника и другие. Широко распространена модель, в которой узлы принимаются абсолютно надёжными, подобные сети рассматриваются и в настоящей статье. Сеть с ненадёжными элементами обычно описывается случайным графом [1], вершины которого соответствуют узлам сети, а рёбра — каналам связи. Для каждого элемента e графа G задаётся вещественное число r_e от 0 до 1 — вероятность присутствия элемента в графе, что соответствует надёжности соответствующего элемента сети. Предполагается, как правило, что рёбра отказывают независимо друг от друга.

Наиболее распространённой мерой надёжности сети является вероятность связности соответствующего случайного графа, однако изучаются и другие, например, – средняя вероятность связности пар вершин [2], вероятность связности с ограничением на диаметр [3–5] и другие. Задача расчёта вероятности связности сети является NP-трудной [1], как и задачи точного расчёта других показателей надёжности. Далее под надёжностью сети будем подразумевать именно вероятность её связности и обозначать её как $R(G)$. Данный показатель надёжности достаточно хорошо изучен, разработано большое количество различных точных и приближённых методов расчёта [6–12].

В [13] предлагается другой подход к задаче о надёжности сети: ставится задача установить, превосходит ли надёжность сети заданное значение (порог), т.е. является ли сеть достаточно надёжной. Идея решения заключается в последовательном (кумулятивном) уточнении верхней и нижней границ надёжности сети, что можно осуществлять при помощи метода факторизации. Каждый раз, когда в процессе факторизации получается подграф, для которого можно вычислить надёжность непосредственно, мы уточняем верхнюю и нижнюю границу надёжности исходной сети и, если одно из полученных значений пересекает значение порога, принимаем

*Работа поддержана грантом РФФИ № 14-07-31069

решение о её надёжности. Таким образом, появляется возможность делать достоверный вывод о надёжности сети без исчерпывающего расчёта. Данный подход был далее развит в наших прошлых работах [14] при помощи декомпозиции по точкам сочленения и двухвершинным сечениям.

Среди оптимизационных задач на сетях важное место занимает задача синтеза максимально надёжной структуры сети в условиях заданных ограничений [15–17]. При решении подобных задач сети моделируются графами, а для решения, как правило, используются различные бионические техники [18]. Одним из подобных подходов является применение генетических алгоритмов (ГА). ГА представляют собой адаптивные методы поиска, основанные на выборе и рекомбинации перспективных решений. В большинстве случаев они успешно справляются с решениями различных оптимизационных задач на графах и сетях. Ранее авторами данной работы был разработан каркас (framework) для работы с бионическими техниками в применении к задачам структурной оптимизации сетей в условиях стоимостных, структурных и других ограничений с настраиваемыми параметрами и выбором операторов [19, 20].

Основной проблемой при разработке ГА является подбор генетических операторов (селекция, скрещивание, мутация), применение которых оставляет в нужном классе решений и подходящих представлений особей (в данном случае сетевых структур). Другим важным аспектом, определяющим эффективность работы ГА, является выбор метода вычисления или оценки функции пригодности (целевой функции). Это становится особенно актуальным в случае NP-трудности задачи расчёта целевой функции, как и для рассматриваемой в статье задачи расчёта надёжности сети. В настоящей статье предлагается использовать метод кумулятивного уточнения границ надёжности сетей для ускорения отсева непригодных особей. Как показали проведённые численные эксперименты, предлагаемый подход позволяет ускорить ГА структурной оптимизации сети по критерию надёжности.

2. Постановка задачи

Рассмотрим задачу получения наиболее надёжной структуры сети на заданном множестве узлов при некоторых ограничениях. Пусть задано множество вершин V и множество рёбер S. Для каждого ребра e задана стоимость его прокладки c_e и значение надёжности r_e. Сумму весов рёбер графа G на множествах V, S будем обозначать как $Weight(G)$. Задано также натуральное число d и вещественное положительное число C^*.

Необходимо построить связный неориентированный граф $G = (V, E)$, диаметр которого не превышает d, $E \subset S$, со следующими ограничениями:

$$\begin{cases} R(G) \to \max; \\ Weight(G) < C^*. \end{cases} \tag{1}$$

Очевидно, что точное решение задачи потребует перебора $2^{|S|}$ вариантов топологий, для каждой из которых нужно будет решить NP-трудную задачу расчёта её надёжности. Далее предлагается генетический алгоритм решения задачи.

3. Метод факторизации и кумулятивное уточнение границ надёжности сети

Для расчёта показателей надёжности используются различные методы, самый распространённый из которых – метод факторизации (ветвления) [1]. Метод заключается в рекурсивном применении формулы полной вероятности при рассмотрении в качестве альтернативных гипотез наличия либо отсутствия очередного разрешающего ребра. Для вероятности связности соответствующая формула принимает вид

$$R(G) = r_e R(G/e) + (1 - r_e) R(G \backslash e), \tag{2}$$

где G/e – граф со стянутым ребром e, $G \backslash e$ – граф без ребра e. Рекурсии продолжаются либо до получения несвязного графа, либо до получения графа малой размерности, для которого надёжность можно рассчитать непосредственно. В наших прошлых работах [10, 14] получены формулы, позволяющие останавливать рекурсии по получении 5-и вершинных графов, что позволило значительно ускорить расчёт надёжности.

Одним из последних значимых результатов в области анализа сетевой надёжности стал метод кумулятивного уточнения верхней и нижней границ надёжности сети [13]. Используя данный подход, необязательно осуществлять исчерпывающий расчёт, достаточно определить, пересекает ли верхняя или нижняя граница наперёд заданное значение, что позволит принять решения о надёжности/ненадёжности сети по отношению к заданному порогу. Для этого задаётся значение порога R_0, а значения нижней границы RL и верхней границы RU надёжности сети вычисляются последовательно таким образом, что на l-ом шаге $RL_l \geq RL_{l-1}$ и $RU_l \leq RU_{l-1}$.

Для вычисления RL и RU необходимо вычислять точные значения надёжностей некоторых оконечных подграфов, возникающих в процессе факторизации, и вероятности получения самих этих подграфов. Предположим, что в процессе факторизации получено L подграфов $G_1, G_2, \ldots, G_L$, для которых надёжность может быть легко вычислена. Пусть P_l есть вероятность получения G_l для $1 \leq l \leq L$. Тогда $\sum_{l=1}^{L} P_l = 1$, а надёжность исходного графа может быть вычислена как

$$R(G) = \sum_{l=1}^{L} P_l R(G_l). \tag{3}$$

Для каждого $1 \leq k \leq L$ имеет место следующее неравенство [13]:

$$\sum_{l=1}^{k} P_l R(G_l) \leq R(G) \leq 1 - \sum_{l=1}^{k} P_l (1 - R(G_l)). \tag{4}$$

Алгоритм кумулятивного уточнения верхней и нижней границ надёжности сети определяется непосредственно данным неравенством. Как только получен G_l для $1 \leq l \leq L$, мы уточняем RL_l и RU_l по формулам

$$\begin{aligned} RL_l &= RL_{l-1} + P_l R(G_l) \\ RU_l &= RU_{l-1} - P_l (1 - R(G_l)). \end{aligned} \tag{5}$$

RL_0 и RU_0 изначально полагаются равными 0 и 1 соответственно. С ростом l значения RL_l и RU_l приближаются к $R(G)$. При достижении значения R_0 какой-либо из границ алгоритм делает вывод о надёжности G: если RL_l достигла R_0, сеть надёжна, если RU_l достигла R_0 — ненадёжна.

4. Генетический алгоритм

Генетические алгоритмы являются разновидностью эволюционных алгоритмов, т.е. оптимизационных техник, основанных на принципах естественной эволюции. Теория, которую Чарльз Дарвин представил в [21] в 1859 году, послужила основой для развития подобных алгоритмов.

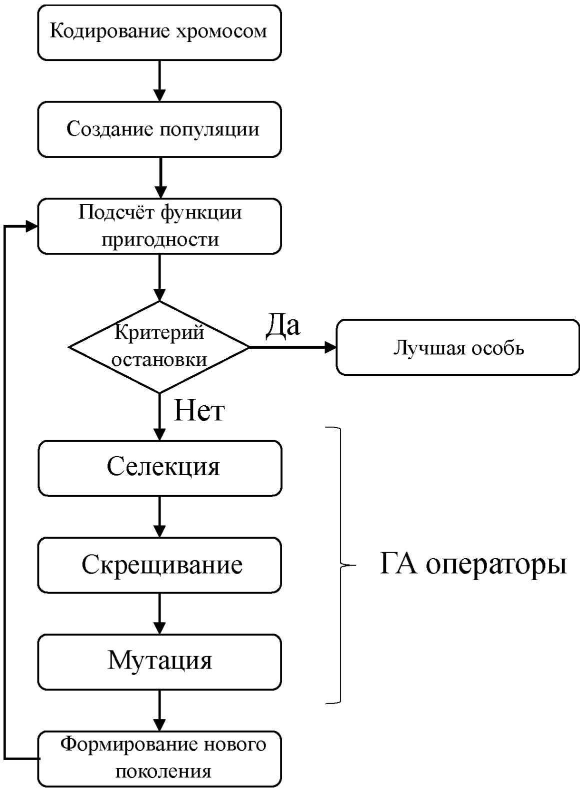

Рис. 1. Схема ГА

Опишем в общих чертах процесс оптимизации генетическим алгоритмом. Сначала определяются особи (или хромосомы), они представляются, например, битовыми строками или упорядоченной последовательностью элементов. Каждая особь представляет собой некоторое решение. Из этих особей произвольно набирается популяция. Далее определяются самые сильные и подходящие особи посредством функции пригодности. С помощью оператора селекции определяются пары родителей (например, турнирной селекцией) и происходит скрещивание (оператор "crossover"). После скрещивания новое поколение подвергается мутации. Новое поколение формируется из самых пригодных особей текущей популяции и новых сформированных потомков. Отсев самых непригодных особей и добор новых до размера популяции делается для того, чтобы алгоритм мог выйти из локального оптимума. Критерием остановки алгоритма может быть количество поколений, временное ограничение, или алгоритм может прекратить работу, если в течение нескольких поколений не происходило улучшения решения.

При оптимизации структуры сети по критерию надёжности предлагается использовать генетические операторы, ускоренные при помощи кумулятивных оценок границ надёжности. Кумулятивные оценки позволяют генетическим операторам работать быстрее за счёт отбрасывания заранее плохих хромосом с низким значением функции пригодности, т.е. недостаточно надёжные реализации сети. Операторы в таком случае приобретают нижеследующий вид.

Оператор мутации. Пусть A_0 — изначальная хромосома с известной надёжностью R_0. Проверка на пригодность новой хромосомы A_1 будет проходить по следующему правилу:

$$Feas(A_1, P_0) = \begin{cases} 1, \text{если} R(A_1) > R_0; \\ 0, \text{иначе.} \end{cases} \tag{6}$$

Оператор скрещивания. Пусть A_0 и A_1 — выбранные родители для скрещивания. Значение надёжности для наиболее надёжного из родителей обозначим как R_{max}. Проверка на пригодность потомка A_2 проводится по следующему правилу:

$$Feas(A_2, R_{max}) = \begin{cases} 1, \text{если} R(A_2) > R_{max}; \\ 0, \text{иначе}. \end{cases} \quad (7)$$

Таким образом, при вычислении функции пригодности осуществляется уточнение границ надёжности новой особи. Если верхняя граница пересекла пороговое значение (R_0 или R_{max}), то расчёт останавливается и принимается решение о непригодности данной особи. Если нижняя граница пересекла пороговое значение, то принимается решение о пригодности данной особи и значение надёжности досчитывается до конца.

5. Эксперименты

Были проведены численные эксперименты для выявления эффективности предложенного алгоритма по сравнению с аналогичным алгоритмом без использования кумулятивных оценок [19]. Оптимальные структуры искались для трёх сочетаний количества узлов, количества рёбер и значения диаметра. Соответствующие значения приведены в верхней строке таблицы. Для каждого из этих трёх случаев было осуществлено по 10 запусков обоих алгоритмов. В каждом из десяти таких запусков для обоих алгоритмов выбирались одинаковые параметры: размер популяции, вероятность мутации, количество поколений, турнир. В таблице приведено среднее время расчёта для каждого алгоритма для каждого сочетания количества узлов, рёбер и значения диаметра.

Таблица 1.

Размер Диаметр	$\vert V\vert = 10$, $\vert E\vert = 20$ $d = 4$	$\vert V\vert = 20$, $\vert E\vert = 30$ $d = 5$	$\vert V\vert = 50$, $\vert E\vert = 70$ $d = 8$
Генетический алгоритм			
Время	8 с 411 мс	4 м 46 с 662 мс	1 ч 52 м 58 с 643 мс
Генетический алгоритм + кумулятивные оценки			
Время	8 с 240 мс	25 с 374 мс	1 м 40 с 607 мс

Из таблицы видно, что во всех случаях модифицированный генетический алгоритм работает быстрее. Разница во времени счёта составила примерно 2 %, 90 %, и 98 % для входных данных, приведённых во второй, третьей и четвертой колонке соответственно. Лучшие результаты ГА с применением кумулятивных оценок для функции пригодности показал на разреженных графах (98 %). Так как при скрещивании и мутации производится много неподходящих потомков, отсев таких решений происходит намного быстрее, чем полный подсчёт их вероятности связности. Для расчётов использовался компьютер с четырёхъядерным процессором Intel(R) Core(TM) i3-4130 CPU 3,40 GHz, RAM=6 Gb.

6. Заключение

В работе предложены модифицированные операторы мутации и скрещивания генетического алгоритма для решения задач структурной оптимизации сети по критерию надёжности в условиях различных ограничений. Новые операторы позволяют прекратить подсчёт функции пригодности для новых особей за счёт кумулятивных оценок вероятности связности сети и отбросить непригодные индивидуумы на более раннем этапе работы алгоритма. Таким образом, генетический алгоритм для задачи оптимизации сети с модифицированными операторами значительно улучшил время своей работы. В дальнейшем планируется применить подобный подход для сетевой оптимизации по другим показателям надёжности, а также дальнейшее повышение эффективности операторов для соответствующих генетических алгоритмов.

Литература

1. *Colbourn Ch. J.* The combinatorics of network reliability. N.Y.: Oxford Univ. press, 1987. 160 p.
2. *Родионов А. С., Родионова О. К.* Кумулятивные оценки средней вероятности связности пары вершин случайного графа // Проблемы информатики. 2013. № 19. С. 3–12.
3. *Canale E., Cancela H., Robledo F., Rubino G., Sartor P.* On computing the 2-diameter-constrained k-reliability of networks // International Transactions in Operational Research. 2013. V. 20(1). P. 49–58.
4. *Мигов Д. А.* Расчет надежности двухполюсной сети с ограничением на диаметр с использованием сечений // Проблемы информатики. 2011. № 11. С. 4–9.
5. *Мигов Д. А., Нестеров С. Н., Родионов А. С.* Методы ускорения расчета надежности сетей с ограничением на диаметр // Вестник СибГУТИ. 2014. № 1. С. 49–56.
6. *Родионов А. С., Сакерин А. В., Мигов Д. А.* Некоторые вопросы реализации полного перебора (на примере задач надёжности сетей) // Вестник СибГУТИ. 2014. № 4. С. 79–84.
7. *Shooman A. M., Kershenbaum A.* Methods for communication-network reliability analysis: probabilistic graph reduction // Proc. of Reliability and Maintainability Symposium. IEEE, 1992. P. 441–448.
8. *Соколова Э. С., Капранов С. Н., Балашова Т. И.* Поиск полного множества минимальных сечений в сетях передачи данных // Научно-технический вестник Поволжья. 2013. № 6. С. 431–434.
9. *Лотоцкий А. Д.* Исследование точных методов вычисления структурной надёжности компьютерных сетей // Инновации на основе информационных и коммуникационных технологий. 2014. № 1. С. 233–235.
10. *Мигов Д. А.* Формулы для быстрого расчета вероятности связности подмножества вершин в графах небольшой размерности // Проблемы информатики. 2010. № 6. С.10–17.
11. *Цициашвили Г. Ш., Осипова М. А., Лосев А. С.* Асимптотика вероятности связности графа с низконадёжными ребрами // Прикладная дискретная математика. 2013. № 1. С. 93–98.
12. *Цициашвили Г. Ш., Осипова М. А., Лосев А. С.* Вывод асимптотических констант для вероятности несвязности планарного взвешенного графа // Прикладная дискретная математика. 2014. № 2. С. 97–100.
13. *Won J.-M., Karray F.* Cumulative update of all-terminal reliability for faster feasibility decision // IEEE transactions on reliability. 2010. V. 59, № 3. P. 551–562.
14. *Rodionov A., Migov D., Rodionova O.* Improvements in the efficiency of cumulative updating of all terminal network reliability // IEEE Transactions on Reliability. 2012. V. 61, № 2. P. 460–465.
15. *Koide T., Shinmori S., Ishii H.* Topological optimization with a network reliability constraint // Discrete Applied Mathematics. November 2001. V. 115, issues 1–3. P. 135–149.
16. *Liu B., Iwamura K.* Topological optimization models for communication network with multiple reliability goals // Computers & Mathematics with Applications. April 2000. V. 39, issues 7–8. P. 59–69.
17. *Балашова Т. И.* Обеспечение отказоустойчивости сети повышением надёжности её топлогии // Современные проблемы науки и образования. 2014. № 6. (`www.science-education.ru`/120-16846)

18. *Мочалов В. А.* Гибридный бионический алгоритм синтеза структуры беспроводной сенсорной сети // T-Comm: Телекоммуникации и транспорт. 2013. Т. 7. № 10. С. 72–77.
19. *Нечунаева К. А.* Оптимальное по показателям связности объединение сетей в условиях структурных и стоимостных ограничений // Проблемы информатики. 2010. № 2. С. 18–26.
20. *Rodionov A., Nechunaeva K.* Network structure optimization: genetic operators: mutation and crossover // Proc. of the 7th Int. Conference on Ubiquitous Information Management and Communication (ACM ICUIMC 2013). Kota Kinabalu, Malaysia, 2013. ACM New York, USA. Article No. 52.
21. *Darwin Ch.* On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London: John Murray, 1859. 502 P.

Статья поступила в редакцию 21.10.2015;
переработанный вариант – 20.11.2015.

Мигов Денис Александрович

к.ф.-м.н., научный сотрудник лаборатории моделирования динамических процессов в информационных сетях Института вычислительной математики и математической геофизики СО РАН (630090, Новосибирск, пр. Лаврентьева, 6), тел. (383) 332-69-49,
e-mail: `mdinka@rav.sscc.ru`.

Нечунаева Ксения Александровна

младший научный сотрудник лаборатории моделирования динамических процессов в информационных сетях Института вычислительной математики и математической геофизики СО РАН (630090, Новосибирск, пр. Лаврентьева, 6), тел. (383) 330-96-43,
e-mail: `ksu.nech@gmail.com`.

Родионов Алексей Сергеевич

д.т.н., профессор кафедры вычислительных систем СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. (383) 332-69-49, e-mail: `alrod@sscc.ru`.

Genetic algorithm for network structure optimization using network reliability cumu-lative updating

D. Migov, K. Nechunaeva, A. Rodionov

This paper presents a new algorithm for network structure optimization via reliability criterion under conditions of various constraints. Networks with unreliable communication channels and perfectly reliable nodes are investigated. We consider network reliability as connection probability. The problem of computing this characteristic is known to be NP-hard. The proposed algorithm for optimization is genetic one. Cumulative updating of network reliability bounds lets the algorithm's operators work faster due to early cutting down bad chromosomes (with worse fitness function value). The numerical experiments show the advantages of the proposed method.

Keywords: network reliability, random graph, network structure optimization genetic algorithm, factoring method, diameter constraint.