

## Генетический алгоритм балансировки нагрузки в сети

Предложен генетический алгоритм решения задачи балансировки нагрузки на каналы связи в компьютерных и транспортных сетях как многопараметрической задачи оптимизации. Обоснованы особенности представления хромосом и генетических операций кроссовера, мутации и отбора. Сформированы подходы использования описанного алгоритма в существующих сетях с использованием современных протоколов маршрутизации.

Запропоновано генетичний алгоритм розв'язання задачі балансування навантаження на канали зв'язку в комп'ютерних та транспортних мережах як багатопараметричної задачі оптимізації. Обґрунтовано особливості представлення хромосом та генетичних операцій кросоверу, мутації та відбору. Сформовано підходи використання описаного алгоритму в існуючих мережах із використанням сучасних протоколів маршрутизації.

Задача оптимизации транспорта данных маршрутизация и балансировка нагрузки на каналы связи стала существенной составляющей современных компьютерных сетей и распределенных вычислительных систем. Ограничения существующих алгоритмов маршрутизации не позволяют достаточно эффективно решить эту задачу в динамических гетерогенных транспортных системах. В современных сетях появилась необходимость решения задачи оптимизации передачи данных трафика по многим критериям. Динамика сетей требует учета изменения нагрузки на отдельные каналы связи и формирование механизмов управления процессом маршрутизации в сети в целом. Новые подходы к распределенным вычислениям, облачные вычисления и создание сложных кластерных высокопроизводительных архитектур *НРС* выдвигают дополнительные требования к механизмам маршрутизации и балансировки нагрузки в таких системах. Классические алгоритмы не применимы к задачам такого типа, принадлежащим к классу *NP* сложных. Возникает необходимость формирования новых подходов и алгоритмов балансировки нагрузки на каналы связи в сети. Один

из таких подходов — создание генетического алгоритма

□

В задаче балансировки нагрузки сеть представлена в виде взвешенного ориентированного графа  $G = (V, E, C)$ , где  $V$  — множество вершин,  $E \subseteq V \times V$  — множество ребер графа,  $C_{ij} \in E \rightarrow R$  — пропускная способность каждого соединения ребра графа. Информация о потоке между каждой парой вершин задается функцией  $F_{ij}$ , где  $i$  — исходная вершина,  $j$  — вершина назначения. Предложенный алгоритм может быть применен и для отдельной части сети. В таком случае используется определенный подграф графа  $G$ .

Генетический алгоритм балансировки нагрузки использует таблицу маршрутизации сети, содержащую  $m_{ij}$  альтернативных путей между каждой парой вершин  $i, j$ . Эта таблица хранит все промежуточные вершины пути  $i \rightarrow j$ , а также цену каждого маршрута исходя из алгоритма маршрутизации. Такой подход позволяет применять генетический алгоритм в современных сетях, используя маршруты между двумя вершинами с различной метрикой и административной дистанцией.

Во время функционирования сети каждый маршрут между вершинами  $i$  и  $j$  используется с определенной вероятностью  $k_{ij}$ , причем

□ : маршрутизация генетический алгоритм задача балансировки нагрузки в сети

$$k \quad i \quad j \quad k \quad \dots m_{i,j}$$

Степень использования пропускной способности определенного соединения можно вычислить так

$$P_{l,m} = \frac{X_{l,m}^k(i,j)}{C_{l,m}(i,j)}$$

где

$$X_{l,m}^k(i,j) = \begin{cases} l & \text{если } l < m \\ m & \text{если } l > m \end{cases}$$

Пусть  $P_{l,m}$  тогда задача балансировки нагрузки сводится к минимизации наибольшего потребления канала передачи данных в сети т.е. к поиску такой конфигурации маршрутов передачи данных для которой  $P$  будет наименьшим

□

Хромосома определенного поколения генетического алгоритма  $Y$  состоит из последовательности аллелей  $A_{i,j}$  которые в свою очередь содержат наборы коэффициентов  $k(i,j)$  для каждой пары вершин  $i,j \in V$  т.е.

$$A_{i,j} = k(i,j)$$

Множество  $V \times V$  состоит из всех пар вершин для которых применяется алгоритм балансировки. Для повышения эффективности работы генетического алгоритма и уменьшения затрат на вычисление целесообразно также ограничить множество связей сети ребер графа  $G$  для которых применяется генетический алгоритм. Обозначим это множество  $E$ . Такое упрощение позволяет выделить критические элементы сети и значительно улучшить динамические характеристики алгоритма. Предложенное представление обеспечивает одинаковое количество и положение аллелей в хромосомах произвольного поколения рис

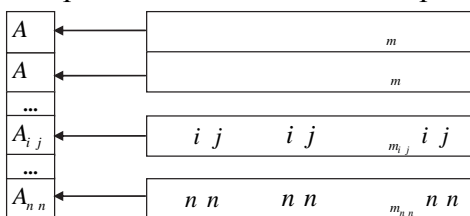


Рис Представление хромосомы

□

Формирование начальной популяции хромосом зависит от выбранного алгоритма маршрутизации. Пусть для каждой пары вершин  $i,j \in V$  существует  $m_{i,j}$  альтернативных маршрутов  $i \rightarrow j$ . Пользуясь метрикой алгоритма маршрутизации каждому маршруту  $i \rightarrow j$  ставится в соответствие определенная величина  $Q_{i,j}^k$  характеризующая цену этого пути с учетом алгоритма маршрутизации. При применении нескольких алгоритмов маршрутизации эта величина зависит от метрики маршрута и административной дистанции протокола маршрутизации. Коэффициенты  $k(i,j)$  начальной популяции генетического алгоритма выбираются так

$$k(i,j) = \frac{Q_{i,j}^k}{Q_{i,j}^k}$$

Следует заметить что применяя алгоритм уместно использовать определенные дополнительные условия для соотношения коэффициентов  $k(i,j)$

$$k(i,j)$$

Предложенные дополнительные условия накладывают ограничения на отношение величин потоков данных для различных маршрутов в сети. Однако такие ограничения необходимы для уменьшения возможного негативного влияния алгоритма балансировки на динамику передачи данных между определенными точками в сети

□ □

Операция мутации для предложенного алгоритма заключается в случайном изменении определенного аллеля хромосомы  $A_{i,j}$  одного члена мутация или изменении нескольких аллелей хромосомы одновременно. Многоточечная мутация. Операция мутации аллелей в хромосоме должна удовлетворять условиям и

Для выбора точек мутации используется эвристический подход. Поскольку задача алгоритма состоит в минимизации наибольшего потребления каналов в сети лишь изменение величины нагрузки на канал с наибольшим по

треблением может улучшить решение задачи в следующем поколении генетического алгоритма. Пусть в пределах текущего поколения наибольшее потребление канала наблюдается для канала  $l, m \in E$ . Следовательно мутации подлежат только те аллели  $A_{i,j}$  для которых  $k$   $m_{i,j}$  такое что

$$k \cdot i \cdot j \quad \text{и} \quad X_{l,m}^k \cdot i \cdot j$$

Аналогично выбираются и точки кроссовера. В силу предложенного способа представления решения длина хромосомы останется постоянной. Поэтому в описанном алгоритме операция кроссовера заключается в обмене частями или отдельными аллелями родительских хромосом.

При формировании операций кроссовера и мутации следует учитывать два аспекта:

сохранение последовательностей аллелей соответствующих положительному приросту функции приспособленности для обеспечения сходимости алгоритма в целом;

эффективное создание новых последовательностей аллелей для обеспечения разнообразия хромосом в пределах поколения и избежания сходимости к локальному экстремуму целевой функции.

Обеспечение описанных требований осуществляется путем адаптивного изменения вероятностей операций кроссовера и мутации хромосом генетического алгоритма.

Определим вероятности кроссовера и мутации произвольной хромосомы  $Y$ :

$$P_c = \frac{p \cdot F \cdot F Y / F \cdot F Y \cdot F Y \cdot \bar{F}}{p \cdot F Y \cdot \bar{F}}$$

$$P_m = \frac{p \cdot F \cdot F Y / F \cdot F Y \cdot F Y \cdot \bar{F}}{p \cdot F Y \cdot \bar{F}}$$

$F$  — значение функции приспособленности лучшей хромосомы популяции;  $\bar{F}$  — среднее значение функции приспособленности хромосом в популяции;  $p$  — вероятность кроссовера и мутации наименее приспособленных хромосом.

Для обеспечения разнообразия хромосом в пределах популяции вероятность кроссовера и мутации наименее приспособленных хромосом должна быть высокой. Поэтому в предложенном алгоритме  $p$  — вероятность кроссовера и мутации наименее приспособленных хромосом. Такой

адаптивный подход к операциям кроссовера и мутации обеспечивает эволюцию популяции и улучшает сходимость алгоритма. С другой стороны при росте количества одинаковых хромосом в популяции вероятности мутации и кроссовера растут. Таким образом предложенный алгоритм позволяет избежать преждевременной сходимости к локальному экстремуму.

Поскольку равенство  $\frac{F Y}{l, m \in E} = \frac{P l m}{P l m}$  соответствует задаче поиска минимума функции приспособленности хромосомы  $Y$  может быть вычислена как

$$F Y = \frac{P l m}{l, m \in E}$$

Алгоритм использует пропорциональный метод отбора в сочетании с методом сохранения лучших особей популяции — элитизмом. Значит вероятность отбора определенной хромосомы пропорциональна значению функции приспособленности этой хромосомы. Пусть размер популяции  $N$ , значение функции приспособленности хромосомы  $Y_i$  равно  $F Y_i$ , тогда вероятность отбора  $Y_i$  вычисляется так:

$$p_i^s = \frac{F Y_i}{\sum_i F Y_i}$$

Поскольку при описанном подходе существует вероятность статистической ошибки и потери лучших хромосом в предложенном алгоритме применяется метод сохранения лучших решений. При этом гарантируется переход лучших особей следующего поколения генетического алгоритма и сходимость алгоритма в целом.

После завершения работы генетического алгоритма балансировки нагрузки его результаты должны быть использованы в процессе маршрутизации сети. В современных протоколах процесс маршрутизации потока данных между вершинами  $i$  и  $j$  на промежуточных вершинах маршрутизаторов базируется на информации о вершине назначения  $j$  пакета данных, который проходит через промежуточную вершину. Поскольку алгоритм балансировки базируется на информации о всех промежуточных вершинах

нах маршрутов следует расширить классический формат таблицы маршрутизации и алгоритма маршрутизации в целом

Рассмотрим таблицу маршрутизации на промежуточной вершине  $k$  сети. В предлагаемом подходе каждая запись таблицы соответствующая маршруту  $i \rightarrow j$  который проходит через вершину  $k$  дополнительно содержит информацию о начальной вершине  $i$ . Если существует несколько альтернативных путей между вершинами  $i$  и  $j$  то запись также содержит вероятность выбора определенного маршрута. Пусть существует  $m$  путей  $i \rightarrow j$  проходящих через вершину  $k$  тогда вероятность выбора того маршрута определяется так

$$P_{i,j} = \frac{P_{i,j}^s}{\sum_{s=1}^m P_{i,j}^s}$$

Таким образом результаты работы генетического алгоритма балансировки нагрузки в сети интегрируются в дополнительные записи таблицы маршрутизации в каждой промежуточной вершине обмена трафиком. Алгоритм маршрутизации пакета данных в данном случае выглядит так

поиск набора маршрутов соответствующих информации о вершине назначения пакета данных

отбор маршрутов соответствующих информации о начальной вершине

применение определенного маршрута на основе вероятности его выбора

Существующие записи таблицы маршрутизации не меняются обеспечивая работу сети в

отсутствие алгоритма балансировки и в случае когда промежуточная вершина или определенный маршрут не используются алгоритмом балансировки. В таком случае описанный алгоритм выбора маршрута соответствует классическому алгоритму.

Описанный подход применим для построения и оптимизации транспортных и компьютерных сетей. Его внедрение позволит более эффективно использовать ресурсы сетей и значительно улучшить динамические характеристики систем балансировки нагрузки в распределенных вычислительных системах.

*Girish M. Zhou B. Hu J.*

Р

*Погорелый С.Д. Білоус Р.В.* Генетичний алгоритм розв'язання задачі маршрутизації в мережах. Проблеми програмування № С

*Олифер В.Г. Олифер Н.А.* Компьютерные сети. Принципы технологии протоколы. СПб. Питер С

*Lee Y. Seok Y. Choi Y.*

*Goldberg D.E.*

*Srinivas M. Patnaink L.*

24

Поступила

Тел. для справок

Киев

*E-mail: sdp@univ.kiev.ua, romanrpd@gmail.com*

С.Д. Погорелый, Р.В. Білоус