УДК 621.396

ОБ ОДНОМ АЛГОРИТМЕ РЕШЕНИЯ ЗАДАЧИ ОПТИМАЛЬНОЙ МАРШРУТИЗАЦИИ ПО КРИТЕРИЮ СРЕДНЕЙ ЗАДЕРЖКИ

В. И. Парфенов, С. В. Золотарев

Воронежский государственный университет

Предложен новый алгоритм решения задачи оптимальной маршрутизации по критерию средней задержки, основанный на применении к информационным сетям законов Кирхгофа. Алгоритм не требует обязательного использования производных целевой функции, поэтому является весьма удобным для распределенных реализаций. Приведен пример использования алгоритма для сети простой топологии и произведено сравнение полученных результатов с теоретическими расчетами. Показано, что применение алгоритма для оптимизации сети требует небольшого числа итераций, обеспечивая при этом точность в десятые доли процента.

ВВЕДЕНИЕ

В последнее время наблюдается бурное развитие и создание новых сетей. Подобные сетевые структуры порождаются как на техническом уровне (транспортные потоки, сети связи, вычислительные комплексы), так и на информационном (потоки данных зарождаются, переплетаются, требуя обработки на промежуточных этапах). Эффективное использование сетей требует анализа их функционирования как на этапе проектирования, так и в ходе эксплуатации. Одной из важных задач, встречающихся при анализе сетей, является задача маршрутизации.

Проблемы маршрутизации присутствуют в сетях любого типа – как в сетях коммутации пакетов и сообщений, так и в цифровых сетях коммутации каналов. Конкретная реализация алгоритма маршрутизации существенно зависит от специфических особенностей сети, но в целом для различных сетей используется достаточно похожий математический аппарат — алгоритмы кратчайшего пути и потоковые алгоритмы, применяемые к потоковым моделям сетей, основанных на интенсивностях трафика, поступающего в линии связи. В потоковых моделях делается неявное предположение, что статистика трафика, поступающего в сеть, не меняется во времени. Такое допущение является разумным, когда эта статистика меняется очень медленно по сравнению со средним временем, необходимым для уменьшения очередей в сети, и когда потоки в линиях измеряются путем временного усреднения.

Под алгоритмом маршрутизации понимается правило, в соответствии с которым в каждом узле сети передачи данных осуществляется выбор линии связи для передачи блока данных (сообщения или пакета). Под фиксированной (неразветвленной) маршрутизацией будем понимать такую процедуру выбора маршрутов, при которой для передачи данных от узла-источника узлу-адресату (пара отправитель-адресат, ОА-пара) используется единственный маршрут. Если в процедуре выбора маршрутов можно использовать более одного пути, то она называется альтернативной разветвленной или многопутевой. Очевидно, что в общем случае альтернативная маршрутизация является предпочтительнее, чем фиксированная, так как она более полно использует ресурсы сети. При этом распределение сообщений по линиям связи, исходящим из узла коммутации, производится согласно правилу рандомизации, в соответствии с которым сообщение поступает в линию связи или выходит из сети с заданной вероятностью.

В сетях передачи данных с промежуточным хранением сообщений имеется много очередей на передачу, которые взаимодействуют друг с другом в том смысле, что поток, уходящий из одной очереди, поступает в одну или несколько других очередей. С аналитической точки зрения это усложняет характер процессов поступления очереди, расположенные по течению потока. Трудность состоит в том, что когда сообщения передаются за пределами первой по отношению к точке их входа в сеть очереди, то интервалы между моментами поступления сообщений становятся сильно коррелированными с их длинами. В результате невозможно выполнить точный и эффективный анализ такой сети. Для того чтобы разрешить эту трудность, Л. Клейнрок предположил, что при объединении нескольких потоков сообщений в линии связи сохраняется независимость между интервалами поступления и длинами сообщений. Было заключено, что при этом для каждой линии связи можно приближенно принять модель системы M/M/1 с пуассоновским распределением моментов поступления сообщений в линию связи и экспоненциальным временем передачи (обслуживания) сообщения в ней. В дальнейшем в качестве целевой (стоимостной) функции сети будет выступать средняя задержка сообщения в сети, полученная на основе аппроксимации независимостей Клейнрока.

На стадии проектирования сети передачи данных и в процессе ее развития, задача выбора алгоритма маршрутизации является одной из основных. Цель данной работы состоит в решении задачи оптимальной маршрутизации в информационных сетях с коммутацией сообщений при помощи алгоритма, последовательное повторение которого позволяет прийти к устойчивому решению с заданной точностью. Отличие предложенного алгоритма от аналогичных состоит в том, что он не требует обязательного использования производных целевой функции.

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим следующую потоковую модель информационной сети (ИС) с альтернативной маршрутизацией сообщений. ИС состоит из N узлов коммутации и M линий связи. Предполагается, что:

1) все линии связи абсолютно надежны;

2) все линии связи помехоустойчивы;

3) узлы коммутации имеют бесконечную память;

4) время обработки в узлах коммутации отсутствует;

5) длины всех сообщений независимы и распределены по показательному закону со средним значением $1/\mu$ [байт];

6) трафик, поступающий в сеть, состоит из сообщений, имеющих одинаковый приоритет, и образует пуассоновский поток со средним значением γ_{kl} [сообщений/сек] для сообщений, возникающих в узле k и предназначенных узлу l; обозначим:

$$\gamma = \sum_{k=1}^{N}\sum_{l=1}^{N}\gamma_{kl}$$ — полный внешний трафик;

7) пропускная способность каждой линии связи, соединяющей узлы i и j, равна d_{ij} [байт/сек].

Важной характеристикой качества функционирования сети является средняя задержка сообщения в сети T. Применение формулы Литтла к сети очередей приводит к общему, чрезвычайно простому результату, впервые полученному Клейнроком:

$$T = \frac{1}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{f_{ij}}{\mu d_{ij} - f_{ij}} = \sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij}(f_{ij}), \qquad (1)$$

где f_{ij} — суммарный поток [сообщений/сек], протекающий по линии (i, j), а $T_{ij}(f) = f_{ij}/\gamma(\mu d_{ij} - f_{ij})$ является слагаемым суммы (1) и не несет в себе особого физического смысла.

Сделанные предположения и обозначения позволяют сформулировать задачу поиска таких значений потоков f_{ij}, образующих вектор $f = \{f_{ij}\}$, которые обеспечат оптимальное (наименьшее) значение величине T.

Известны:

1) топологическая структура ИС;

2) матрица входных потоков $\|\gamma_{kl}\|$;

3) пропускные способности линий связи $\|d_{ij}\|$;

4) средняя длина сообщения $1/\mu$.

Требуется найти: потоки в линиях связи f_{ij} такие, что:

$$T(f) = \frac{1}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{f_{ij}}{\mu d_{ij} - f_{ij}} \to \min \qquad (2)$$

при выполнении ограничений:

$$0 \le f_{ij} < \mu d_{ij};\ i, j = 1, 2, \ldots, N. \qquad (3)$$

Данная задача называется *задачей выбора оптимальных потоков в информационной сети по критерию средней задержки*.

В случае альтернативной маршрутизации эта задача относится к классу многопродуктовых задач с выпуклой целевой функцией и выпуклым множеством ограничений. Следовательно, существует единственный локальный минимум данной задачи, являющийся глобальным минимумом, для нахождения которого разработано достаточно большое число вычислительных методов [1, 2]. Наиболее известным методом решения данной задачи является метод отклонения (девиации) потока, предложенный в работе [1]. Данный алгоритм является частным случаем так называемого метода Франка—Вольфа [3] для решения более общих задач нелинейного программирования.

АЛГОРИТМ УСЛОВНОЙ МИНИМИЗАЦИИ

Рассмотрим другой метод нахождения решения задачи оптимальной маршрутизации. Введем вспомогательную функцию одной переменной $g(x \mid a)$, где a является параметром, определив ее следующим образом:

$$g(f' \mid f) = \sum_{i=1}^{N}\sum_{j=1}^{N}(f_{ij}')^2 z_{ij}(f_{ij}), \qquad (4)$$

где $f' = f + \alpha(\varphi - f)$ — выпуклая комбинация допустимых потоков φ и f, а величины $z_{ij}(f_{ij}) = T_{ij}(f_{ij})/f_{ij}^{\,2} = 1/f_{ij}\gamma(\mu d_{ij} - f_{ij})$ представляют собой «веса» линий связи (i, j). Будем называть поток допустимым, если он удовлетворяет ограничениям (3). Множество допустимых потоков, проходящих через линии связи, будем обозначать F. Тогда, если f — допустимый поток, то $f \in F$. Под (4) можно понимать условную мощность сети, распределение потоков в которой задается вектором f', когда «веса» линий связи $z(f)$ полагаем постоянными (не зависящими от проходящего через них потока f'), то есть потоки f выступают в роли параметров.

Можно показать, что задача нахождения оптимального решения задачи (2) относится к классу итеративных алгоритмов для решения задачи оптимальной маршрутизации. Основной итерацией этого алгоритма является:

$$f := f + \alpha^*(\varphi - f),$$

где α^* — такой размер шага, который минимизирует $T[f + \alpha(\varphi - f)]$ по всем $\alpha \in [0,1]$, а вектор потоков φ обеспечивает наименьшую величину условной мощности $g(\varphi \mid f)$. Для оптимального размера шага α^* будем использовать метод, предложенный в [4], в соответствии с которым:

$$\alpha^* = \min_{\alpha\in[0,1]}\{T[f + \alpha(\varphi - f)]\} =$$

$$= \min\left\{1, -\frac{\sum\limits_{i,j}(\varphi_{ij} - f_{ij})\dfrac{dT_{ij}(f_{ij})}{df_{ij}}}{\sum\limits_{i,j}(\varphi_{ij} - f_{ij})^2\dfrac{d^2T_{ij}(f_{ij})}{df_{ij}^{\,2}}}\right\}. \qquad (5)$$

Следует отметить, что использование первых двух производных $dT_{ij}(f_{ij})/df_{ij}$ и $d^2T_{ij}(f_{ij})/df_{ij}^{\,2}$ не является обязательным, поскольку для нахождения оптимального размера шага α^* можно использовать любой метод одномерного поиска (например, метод Фибоначчи).

Из-за использования потоковой модели сети проведем аналогию (вполне правомерную) между информационной сетью и электрической сетью, которая имеет ту же структуру (топологию). Токи в электрических сетях всегда распределяются таким образом, чтобы суммарная мощность, выделяемая на элементах сети, принимала наименьшее значение. Тогда можно говорить, что условная мощность сети $g(\varphi \mid f)$, которая является аналогом электрической мощности, будет иметь наименьшее значение, когда вектор φ будет удовлетворять законам Кирхгофа, то есть когда информационные потоки будут распределяться как электрические токи. Для информационной сети первый закон Кирхгофа можно сформулировать следующим образом: алгебраическая сумма потоков, втекающих и вытекающих из любого узла сети, равна нулю:

$$\sum_i\sum_j \varphi_{ij} = 0. \qquad (6)$$

Первый закон Кирхгофа обеспечивает условие сохранения потока в сети. Второй закон Кирхгофа для ИС сформулируем так: алгебраическая сумма числа пакетов вдоль любого замкнутого контура информационной сети равна нулю, то есть

$$\sum_i\sum_j \varphi_{ij} z_{ij}(f_{ij}) = 0, \qquad (7)$$

где величины $z_{ij}(f_{ij}) = 1/f_{ij}\gamma(\mu d_{ij} - f_{ij})$ можно трактовать как кибернетические сопротивления линий связи (аналоги электрических сопротивлений). Потоки φ, найденные в соответствии с (6)—(7), так же как и мощность $g(\varphi \mid f)$ будем называть условно оптимальными (для данных f).

Таким образом, процедура отыскания вектора потока $f := f + \alpha^*(\varphi - f)$ на каждой итерации состоит из двух этапов. На первом этапе находим вектор условно оптимальных потоков φ. Для этого необходимо решить линейную систему уравнений (6)—(7). На втором этапе из (5) находим оптимальный шаг α^*.

Ниже предлагается модель алгоритма, реализующего предложенный метод оптимизации сети. В дальнейшем будем называть его алгоритмом условной минимизации.

Описание алгоритма

Шаг 1. Произвести анализ ИС на основе законов (6)—(7) и определить аналитическое выражение для нахождения условно оптимального вектора потоков $\varphi = \varphi[z(f)]$, предполагая, что топология сети неизменна во времени.

Шаг 2. Инициализировать «веса» линий связи z_{ij}:

$$z_{ij} = 1/(\mu d_{ij});\ i,j = 1,2,..,N.$$

Шаг 3. Используя «веса» линий связи z_{ij}, определить условно оптимальные потоки φ_{ij}; $i,j = 1,2,..,N$, используя зависимость $\varphi = \varphi[z(f)]$, полученную на первом шаге алгоритма.

Шаг 4. Распределить потоки в соответствии с найденными условно оптимальными потоками:

$$f_{ij} = \varphi_{ij};\ i,j = 1,2,..,N.$$

Шаг 5. Вычислить значение целевой функции: $T_{Old} = \dfrac{1}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}\dfrac{f_{ij}}{\mu d_{ij} - f_{ij}}$.

Шаг 6. Определить веса линии связи $z_{ij}(f_{ij}) = 1/f_{ij}\gamma(\mu d_{ij} - f_{ij})$; $i,j = 1,2,..,N$.

Шаг 7. Используя известную зависимость $\varphi = \varphi[z(f)]$, определить условно оптимальные потоки φ_{ij}; $i,j = 1,2,..,N$.

Шаг 8. Найти размер шага α^*, такой что:

$$\alpha^* = \min_{\alpha\in[0,1]}\left\{\frac{1}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{f_{ij} + \alpha(\varphi_{ij} - f_{ij})}{\mu d_{ij} - f_{ij} - \alpha(\varphi_{ij} - f_{ij})}\right\}.$$

Шаг 9. Распределить потоки в соответствии с формулой:

$$f := f + \alpha^*(\varphi - f).$$

Шаг 10. Вычислить новое значение целевой функции: $T_{New} = \dfrac{1}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}\dfrac{f_{ij}}{\mu d_{ij} - f_{ij}}$.

Шаг 11. Если $|T_{New} - T_{Old}| \le \varepsilon$, то STOP (прекратить работу алгоритма), иначе положить $T_{Old} := T_{New}$ и перейти к шагу 6.

Следует заметить, что данный алгоритм предполагает допустимость входящих в сеть внешних потоков, что выражается в том, что эти потоки априори не должны вызывать перегрузку сети. Если это условие выполняется, то алгоритм позволяет получать асимптотически точное решение задачи оптимальной альтернативной маршрутизации.

ЧИСЛОВОЙ ПРИМЕР

Рассмотрим простую сеть, состоящую из двух линий связи (рис. 1), в которой узел 1 — это единственный отправитель, а узел 2 — единственный получатель сообщений. Известный входной поток γ требуется разделить на два путевых потока f_1 и f_2 так, чтобы минимизировать целевую функцию, основанную на аппроксимации линии связи моделью M/M/1:

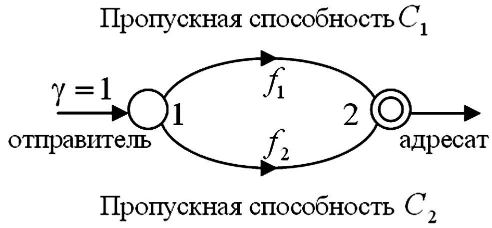

Рис. 1. Пример задачи маршрутизации в сети с одной ОА-парой и двумя путями

$$T(f) = T_1(f_1) + T_2(f_2), \tag{8}$$

где $T_i(f_i) = f_i/\gamma(C_i - f_i)$, $i = 1,\ 2$, а $C_i = \mu d_i$ является пропускной способностью линии i [сообщений/сек].

Для того, чтобы эта задача имела смысл, мы должны предположить, что γ меньше, чем максимальная скорость передачи $C_1 + C_2$, которую сеть может иметь.

Предположим, что $C_1 \ge C_2$. Тогда из соображений здравого смысла оптимальный поток f_1^* не может быть меньше чем f_2^* (не имеет смысла посылать больше трафика по линии с меньшей пропускной способностью). Данная задача может быть легко решена аналитически и нетрудно показать, что она имеет два варианта решения.

1. Оптимальные потоки распределяются следующим образом:

$$f_1^* = \gamma,\ f_2^* = 0. \tag{9}$$

Это может произойти только когда

$$\gamma \le C_1 - \sqrt{C_1C_2}. \tag{10}$$

2. Оптимальные потоки положительны и равны:

$$f_1^* = \frac{\sqrt{C_1}[\gamma - (C_2 - \sqrt{C_1C_2})]}{\sqrt{C_1} + \sqrt{C_2}},\quad f_2^* = \frac{\sqrt{C_2}[\gamma - (C_1 - \sqrt{C_1C_2})]}{\sqrt{C_1} + \sqrt{C_2}} \tag{11}$$

при

$$\gamma > C_1 - \sqrt{C_1C_2}. \tag{12}$$

Решим поставленную задачу с помощью алгоритма условной минимизации. В соответствии с шагом первым алгоритма найдем зависимость условно оптимальных потоков φ_i в линиях связи от «весов» $z_i = z_i(f_i)$ линий связи. Запишем первый закон Кирхгофа для узла-отправителя:

$$\varphi_1 + \varphi_2 = \gamma. \tag{13}$$

Применение к сети второго закона Кирхгофа позволяет получить следующее уравнение:

$$\varphi_1 z_1 - \varphi_2 z_2 = 0. \qquad (14)$$

Решив систему уравнений (13)—(14), получаем искомые зависимости:

$$\varphi_1 = \frac{\gamma}{1 + z_1 / z_2}, \qquad \varphi_2 = \frac{\gamma}{1 + z_2 / z_1}. \qquad (15)$$

Теперь определим аналитический вид «весов» $z_i = z_i(f_i)$, необходимый для выполнения второго и шестого шага алгоритма:

$$z_1(f_1) = T_1(f_1) / f_1^2 = 1/f_1\gamma(C_1 - f_1), \qquad z_2(f_2) = T_2(f_2) / f_2^2 = 1/f_2\gamma(C_2 - f_2). \qquad (16)$$

Далее определим порядок нахождения оптимального шага итерации α^*, который реализован на восьмом шаге алгоритма. Нетрудно показать, что размер шага $\overline{\alpha}$, при котором достигается минимум целевой функции, для нашей задачи равен:

$$\overline{\alpha} = -\frac{(\varphi_1 - f_1)\dfrac{C_1}{(C_1 - f_1)^2} + (\varphi_2 - f_2)\dfrac{C_2}{(!_2 - f_2)^2}}{(\varphi_1 - f_1)^2\dfrac{2C_1}{(C_1 - f_1)^3} + (\varphi_2 - f_2)^2\dfrac{2C_2}{(C_2 - f_2)^3}}.$$

Тогда оптимальное значение размера шага в соответствии с (7) равно:

$$\alpha^* = \min\left\{1,\ \overline{\alpha}\right\}. \qquad (17)$$

Далее для простоты будем опускать размерность [сообщений/сек] интенсивностей потоков и пропускных способностей линий связи. Значение средней временной задержки также будет записано без размерности, полагая, что оно измеряется в секундах.

Пусть входной поток равен $\gamma = 1$, а пропускные способности линий связи равны $C_1 = 3$, $C_2 = 1$. Тогда будет выполнено условие (10) и теоретические значения оптимальных потоков равны $f_1^* = 1$, $f_2^* = 0$. Соответствующее значение целевой функции, в роли которой выступает среднее время задержки сообщения в сети, в соответствии с (8) равно $T(f^*) = 1/(C_1 - \gamma) = 0,5$.

Результаты решения данной задачи с помощью алгоритма условной минимизации с точностью $\varepsilon = 0,001$ (0,2 %) приведены в таблице 1.

Как видно из таблицы 1, решение задачи достигается за пять итераций алгоритма.

Пусть входной поток равен $\gamma = 3$, а пропускные способности линий связи равны $C_1 = 3$, $C_2 = 1$. Тогда будет выполнено условие (12) и

Таблица 1

Результаты анализа сети при входном потоке $\gamma = 1$

Итерация (номер) k	0	1	2	3	4
f_1	0,955	0,978	0,989	0,995	0,9974
f_2	0,045	0,022	0,011	0,005	0,0026
Целевая функция $T(f)$	0,5143	0,5061	0,5028	0,5014	0,5007
Точность $\varepsilon = T(f) - T(f^*)$	0,0143	0,0061	0,0028	0,0014	0,0007

теоретические значения оптимальных потоков в соответствии с (11) равны $f_1^* = 2,366$, $f_2^* = 0,634$. Соответствующее значение целевой функции, в роли которой выступает среднее время задержки сообщения в сети, в соответствии с (8) равно $T(f^*) = 1,8214$.

Результаты решения данной задачи с помощью алгоритма условной минимизации с точностью $\varepsilon = 0,001$ (0,05 %) приведены в таблице 2.

Таблица 2

Результаты анализа сети при входном потоке $\gamma = 3$

Итерация (номер) k	0
f_1	2,362
f_2	0,638
Целевая функция $T(f)$	1,8215
Точность $\varepsilon = T(f) - T(f^*)$	0,0001

Как видно из таблицы 2, решение задачи достигается всего за одну итерацию.

Таким образом, алгоритм условной минимизации позволяет успешно решать задачу оптимальной альтернативной маршрутизации по критерию средней задержки, затрачивая при этом небольшое число итераций. При этом точность решения составляет десятые доли процента.

ЗАКЛЮЧЕНИЕ

Для реальных сетей передачи данных предложенная модель сети, основанная на аппроксимации независимостей Клейнрока, не всегда позволяет получить точные результаты анализа. Поэтому используемый алгоритм условной минимизации, как и другие подобные алгоритмы оптимизации сети по критерию средней задержки, в действительности позволяет получить квазиоптимальное решение. Действительно оптимальное решение можно получить лишь в том случае, если удастся найти точную математическую модель линии связи, которая в общем случае является системой массового обслуживания типа G/G/1. На основе этой модели задается целевая функция сети. Однако к настоящему времени эта задача не всегда может быть решена аналитически.

ЛИТЕРАТУРА

1. *Fratta L., Gerla M., Kleinrock L.* The flow deviation method: An approach to store and forward communication network design // Networks. — 1973. — V. 3, № 2. — P. 97—133.

2. *Schwartz M., Cheung C. K.* The gradient projection algorithm for multiple routing in message-switched networks // IEEE Trans. On Commun. — 1976. — V. 25. — P. 100—127.

3. *Frank M., Wolfe P.* An algorithm for quadratic programming // Naval Research Logistic Quarterly. — 1956. — № 3. — P. 95—110.

4. *Бертсекас Д.* Сети передачи данных / Д. Бертсекас, Р. Галлагер: Пер. с англ. — М.: Мир, 1989. — 544 с.