

ИСПОЛЬЗОВАНИЕ ГРАФОВ ДЛЯ СОКРАЩЕНИЯ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ (на примере анализа модели сети провайдера Интернет)

А. М. Пуртов, В. Н. Задорожный (Омск)

Совместное использование имитационных моделей (ИМ) и графовых моделей процессов позволяет повысить информативность и сократить трудоемкость имитационных экспериментов [1]. Разработанная авторами методика демонстрируется на примере моделирования сети провайдера Интернет.

1. Постановка задачи

Рассмотрим структуру технических средств провайдера, приведенную на рис. 1. За основу взята сеть провайдера Интернет для учреждений науки, образования и культуры г. Омска [2]. Для связи с первичным провайдером используются полудуплексные каналы K1 и K2. Канал K1 служит для передачи данных между провайдером и междугородной телефонно-телеграфной станцией (МТТС). Сотрудники провайдера работают в локальной сети и имеют доступ в Интернет. В одном здании с провайдером находятся организации, сотрудникам которых также предоставляется доступ в Интернет. Локальные сети этих организаций и провайдера объединены с помощью коммутатора.

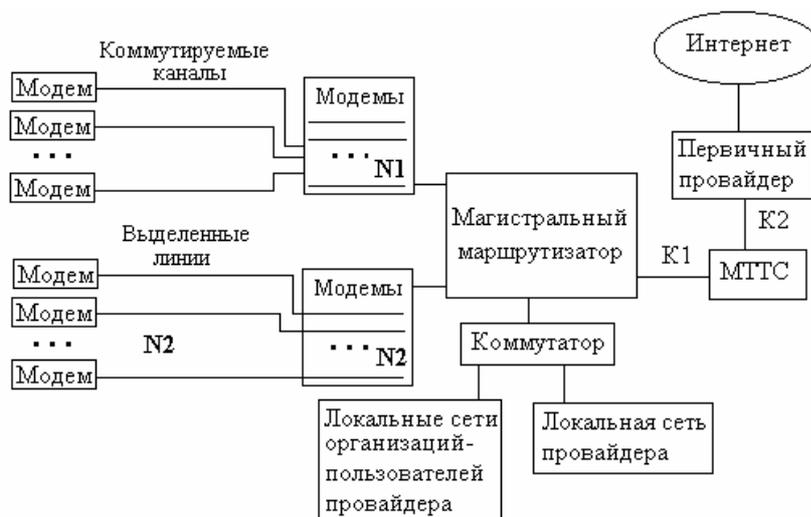


Рис. 1. Структура технических средств провайдера

Пользователи вне здания провайдера связываются с ним по выделенным линиям или коммутируемым каналам. Для обеспечения доступа пользователей в Интернет по коммутируемым каналам (режим dial-up) провайдер имеет N1 модемов. Для постоянного подключения пользователей провайдер арендует N2 выделенных линий. Трафиком, проходящим через провайдера, управляет магистральный маршрутизатор.

Требуется оценить влияние параметров системы на среднее время T получения информации из Интернет и оптимизировать параметры технических средств сети.

2. Объекты и процессы ИМ

Уровень детализации ИМ выбирается так, чтобы отразить процессы, которые заведомо наиболее существенно влияют на показатели производительности сети.

Оборудование. Моделями каналов K1, K2 и магистрального маршрутизатора являются одноканальные устройства. Время обработки запросов на МТТС и первичным провайдером считается нулевым. Считается также, что коммутатор слабо влияет на

время выполнения запросов, поэтому в ИМ он отсутствует. Моделями коммутируемых каналов с модемами являются N1 одноканальных устройств. Моделями выделенных линий с модемами являются N2 одноканальных устройств.

Сеансы связи пользователей. Все пользователи локальных сетей в здании провайдера являются источниками запросов на передачу данных в Интернет и приемниками данных из Интернет. Пользователи генерируют запросы на сеансы связи. В каждом сеансе связи запускается случайное количество параллельных процессов. Каждый процесс генерирует случайное количество запросов на получение данных из Интернет. Процесс генерирует следующий запрос только после получения ответа на предыдущий. Количество данных в ответе на запрос – случайная величина. Модель Интернет представляет собой приемник запросов и генератор ответов.

Запрос передается в Интернет через маршрутизатор и каналы K1, K2. Время обработки запросов маршрутизатором не зависит от их длины. Время передачи запроса по каналам K1 и K2 зависит от их быстродействия и длины запроса. После поступления запроса в Интернет происходит его обработка в течение случайного промежутка времени. Затем генерируется ответ и определяется количество блоков в ответе. Первый блок ответа передается через каналы K2, K1 и маршрутизатор пользователю. Если получен не весь ответ, процесс посылает в сторону Интернет блок, разрешающий передачу из Интернет следующего блока ответа. Если ответ полностью получен, процесс может сгенерировать новый запрос. Если лимит запросов для процесса исчерпан, он закрывается. Если у сеанса связи закрыты все процессы, он завершается.

У всех пользователей, работающих по выделенным линиям, процессы выполнения запросов претерпевают на этих линиях дополнительную задержку.

Пользователи, подключаемые по коммутируемым каналам, генерируют заявки на их резервирование. Если есть свободный канал, он резервируется. В противном случае пользователь получает отказ в обслуживании. Когда канал зарезервирован, начинается сеанс связи. После окончания сеанса связи канал освобождается. Сеанс связи протекает так же, как у пользователей, работающих по выделенным линиям.

Для программирования модели использована бесплатно распространяемая студенческая версия GPSSW. Текст программы содержит около 150 блоков GPSS.

3. Граф процесса обработки запроса пользователя

Процесс обработки запроса пользователя на получение из Интернет информации (страницы HTML) можно представить графом задержек, происходящих при передаче и обработке данных (на каналах связи, в маршрутизаторе, в Интернет). На рис. 2 представлен граф обработки запроса пользователя, выходящего в Интернет по коммутируемому каналу.

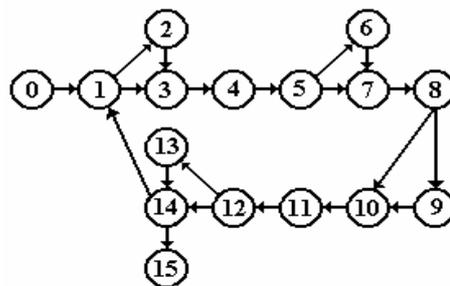


Рис.2. Граф процесса обработки запроса

Неизвестные средние задержки T_i процесса в вершинах $i = 2, \dots, 14$ и переходные вероятности P_{ij} в точках ветвления маршрутов определяются с помощью ИМ. В таблице приводится описание задержек во всех вершинах графа. Возвращение блока данных из Интернет через вершины $8 - \dots - 14 - 1$ повторяет в обратном порядке движение запроса в Интернет через вершины $1 - \dots - 7 - 8$, но характеризуется другими значениями задержек, т. к. блоки данных, идущие в Интернет, обычно короче блоков данных, поступающих из Интернета.

Описание задержек в вершинах графа

Вершина	Описание задержки
0; 15	Начало процесса и конец процесса. $T_0 \equiv 0, T_{15} \equiv 0$
1	Начало захвата коммутируемого канала для передачи блока данных. Если канал свободен, происходит переход в вершину 3, в противном случае – в вершину 2 (канал может быть занят встречным потоком данных); $T_1 \equiv 0$
2	Ожидание освобождения коммутируемого канала
3	Передача блока данных провайдеру по коммутируемому каналу
4	Обработка блока на маршрутизаторе
5	Передача блока по каналу K1 (ожидание освобождения канала K1 на графе не представлено отдельной вершиной, т. к. оно заведомо слабо влияет на время T выполнения запроса)
6	Ожидание освобождения канала K2
7	Передача блока по каналу K2
8	Обработка запроса в Интернет
9	Ожидание освобождения канала K2
10	Передача блока по каналу K2
11	Передача блока по каналу K1
12	Обработка блока на маршрутизаторе
13	Ожидание освобождения коммутируемого канала
14	Передача блока данных пользователю по коммутируемому каналу. Если при этом запрос выполнен не полностью, то выполняется переход к вершине 1, в противном случае – к вершине 15. Вероятность перехода к вершине 1 зависит от величины страницы, поступающей по запросу.

Граф с параметрами T_i и P_{ij} , найденными с помощью ИМ, обрабатывается программой редукции, которая вычисляет следующие дополнительные характеристики:

- аналитическую оценку среднего T и дисперсии D общего времени процесса;
- аналитическую оценку абсолютных и относительных коэффициентов чувствительности (КЧ) этих показателей к параметрам T_i и P_{ij} .

4. Расширенный метод редукции графов с анализом КЧ

Метод редукции графов для определения среднего T и дисперсии D времени процесса предложен Байцером [3]. Феррари [4] предлагает дополнить его анализом чувствительности величин T и D к параметрам графа, вычисляя соответствующие частные производные методом малых приращений. В работе [5] метод редукции графа расширен набором подстановок и пересчетных соотношений, позволяющих получать в однократном процессе редукции сразу все необходимые КЧ, вычисленные по точным формулам аналитического дифференцирования.

Метод редукции графа заключается в последовательной замене типовых фрагментов графа на более простые фрагменты, эквивалентные в смысле времени прохож-

дения через них исследуемого процесса. Пересчетные формулы, по которым определяются параметры вершин и дуг заменяющего фрагмента через параметры заменяемого, достаточно легко выводятся в предположении статистической независимости всех задержек в вершинах и случайных переходов. Обычно редукция завершается стягиванием графа в одну вершину, параметры которой T и D являются искомыми характеристиками длительности всего процесса.

Метод редукции, предложенный в [5], содержит расширенный набор типовых замен. Наряду со «свертыванием цепочки вершин», «размыканием петли», «параллельным объединением» и т. д. он позволяет трансформировать некоторые циклы специального вида, присущие процессам в вычислительных системах и характеризующиеся *зависимостью* переходов. Кроме того, в этом варианте метода состав всех пересчетных соотношений расширен дополнительными соотношениями, по которым рассчитываются частные производные параметров заменяющего подграфа по всем параметрам заменяемого подграфа. Эти дополнительные соотношения получены путем формального дифференцирования основных пересчетных соотношений. Они позволяют на каждом шаге редукции (с учетом правил дифференцирования суперпозиции функций многих переменных) вместе с параметрами нового подграфа вычислять их частные производные по изначально заданным параметрам исходного графа.

Таким образом, в результате редукции графа вместе с показателями T и D определяются значения их частных производных по всем n параметрам исходного графа (абсолютные КЧ), а также относительные КЧ, имеющие вид:

$$Kb[y_k, x_m] = Ka[y_k, x_m] \cdot x_m / y_k, \quad (k = 1, 2; m = 1, \dots, n), \quad (1)$$

где y_k – k -й выходной параметр (т. е. T или D),

x_m – m -й исходный параметр графа (задержка или переходная вероятность),

$Ka[y_k, x_m] = \partial y_k / \partial x_m$ – абсолютный КЧ показателя y_k к параметру x_m .

Метод эффективно реализован программой COIN1 (COefficients of INfluence – коэффициенты влияния), которая написана на Borland C++ 3.1. Она позволяет выполнять редукцию полных графов с отмеченными начальной и конечной вершинами. Входные данные для программы представляются в текстовом файле, выходные данные отображаются в графической форме и выводятся в текстовый файл.

5. Описание имитационных экспериментов

В первом прогоне ИМ задавались следующие значения основных параметров:

- число модемов для коммутируемых каналов $N1 = 5$;
- скорость передачи по коммутируемому каналу – 30 000 бит/с;
- скорость обработки пакетов на маршрутизаторе – 1000 пак/с;
- скорость передачи данных по каналу $K1$ – 1 000 000 бит/с;
- скорость передачи данных по каналу $K2$ – 500 000 бит/с;
- максимальная длина блока данных, передаваемого из Интернет – 80 000 бит;
- время обработки блока данных в Интернет составляет в среднем 1с;
- количество информации, поступающей из Интернет по одному запросу, в среднем равно 1 600 000 бит.

Две последние величины распределены по экспоненциальному закону.

Графовая модель (рис. 2) с параметрами, найденными с помощью ИМ, обрабатывалась программой COIN1, которая выдала коэффициенты $Kb[T, T_i]$, графически представленные на рис. 3. Поскольку T является линейной комбинацией параметров T_i (при фиксированных P_{ij}) то, как следует из (1), сумма этих КЧ должна быть равна единице.

По данным ИМ среднее время выполнения процесса T составило 177 с, по графу – 190 с. Близость экспериментальной и аналитической оценок свидетельствует о том, что граф удовлетворительно аппроксимирует моделируемую сеть, хотя в ней задержки (и переходы) в общем случае не являются независимыми, как предполагается в графе.

Из рис. 3 видно, что наибольший относительный вклад в T формирует параметр T_{14} – среднее время передачи данных пользователю по коммутируемому каналу. С этим же каналом связаны задержки в вершинах 2, 3, 13. Провайдер не может регулировать эти задержки, т. к. они связаны со скоростью работы модемов пользователей и количеством получаемой ими информации из Интернет по одному запросу, т. е. относятся к внешней среде провайдера. С ресурсами провайдера связаны задержки в вершинах 6 и 9, также сильно влияющие на T и определяемые скоростью канала K_2 . Увеличиваем скорость канала K_2 вдвое (дихотомический шаг), т. е. до 1 Мбит/с. Снова запускаем на выполнение ИМ, потом программу COIN1 и получаем КЧ, представленные на рис. 4.

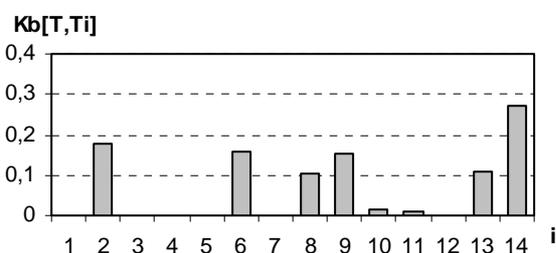


Рис. 3. Относительные КЧ, прогон 1

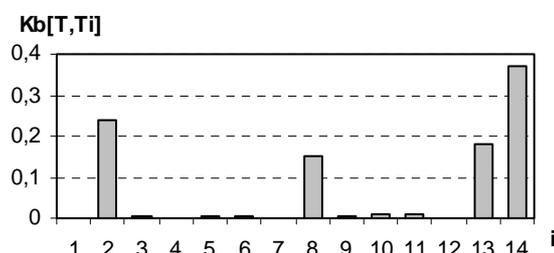


Рис. 4. Относительные КЧ, прогон 2

Из рис. 4 видно, что во втором прогоне модели наибольшее влияние на время T оказывают задержки T_2 , T_{13} , T_{14} (все – на коммутируемом канале) и задержка T_8 , зависящая от времени обработки запроса в Интернет. По результатам ИМ среднее время выполнения процесса сократилось до значения $T=143$ с, по графовой модели – до 138 с.

Таким образом, благодаря КЧ понадобилось всего два прогона имитационной модели, чтобы подобрать рациональные параметры сети провайдера. Близость к нулю коэффициентов $K_b[T, T_6]$ и $K_b[T, T_9]$ говорит о том, что в существующей внешней среде дальнейшее повышение скорости канала K_2 практически не скажется на времени T .

6. Анализ КЧ и оптимизация

Разработанную методику можно распространить и на решение задач оптимизации в классической постановке, если аргументы целевой функции f включают показатели времени ответа T и D , а параметры T_i и R_{ij} выражаются в явном виде (точно или с помощью аппроксимаций) через технические параметры сети. Поскольку редукция графа реализует суперпозицию функций, то ее можно достроить начальным и финальным пересчетами, чтобы сразу вычислялись частные производные функции f по всем контролируемым техническим параметрам. Это позволит использовать в имитационном моделировании сетей массового обслуживания градиентные методы оптимизации.

Концептуально использование для оптимизации сети наряду с ИМ ее высокоаналитичной и быстрой графовой аппроксимации можно уподобить применению квадратичной аппроксимации функций при поиске их локальных экстремумов методом Ньютона.

Литература

1. **Пуртов А. М.** Анализ производительности сетей ЭВМ на графах и имитационных моделях.: Автореф. дисс. на соиск. учен. степени канд. техн. наук (05.13.16)/ Науч. рук. В.А.Шапцев. – Омск: ИИТПМ СО РАН, 1995. –17 с.
2. **Алгазин В. А.** Создание высокоскоростной СПД для информационного обмена между членами научных коллективов (в рамках проекта КС ОКНО)//RELARN – 2001: Матер. VIII конф. представителей регион. научно-образов. сетей. – Петрозаводск: Изд-во ПетрГУ, 2001. –С. 101–105.
3. **Байцер Б.** Микроанализ производительности вычислительных систем: Пер. с англ./Под ред.В. В. Мартынюка. – М.: Радио и связь, 1983. – 360 с.
4. **Феррари Д.** Оценка производительности вычислительных систем. – М.: Мир, 1981. – 576 с.
5. **Задорожный В. Н., Мызникова Т. А.** Рекурсивный анализ чувствительности для метода Байцера. – Деп. в ВИНТИ, 1988. N 5490-B88.