

ФРАКТАЛЬНЫЕ МЕТОДЫ ПРОГРАММНОЙ МАРШРУТИЗАЦИИ В 3G - СЕТЯХ

Интенсивность поступления трафика в любой сети носит фрактальный характер: по мере увеличения масштабов сети пики и спады трафика не сглаживаются, как можно было бы предположить, а, наоборот, складываются в еще более крупные пики и спады. Это означает, что запас производительности сети должен измеряться не десятками процентов, а десятками раз. Таким образом, без интеллектуального управления трафиком проблематично обеспечить требуемое качество сетевых услуг, особенно в интегрированных сетях. Обеспечение качества услуг в современных сетях (особенно мобильных) является серьезной проблемой, которая должна решаться интенсивными методами.

1 Сбор и обработка статистических данных

Многие наблюдения различных процессов приводят к временным зависимостям или рядам измерений. В данной работе исследуются пульсации трафика различного происхождения, поступающего на маршрутизирующий сервер через множество сетевых подключений и интерфейсов мультимедийных служб 3G-сетей (рис. 1). Кроме сетевой, в сетях третьего поколения используется маршрутизация на прикладном уровне, параметры которой определяются, исходя из содержимого трафика. Трафик передается на прикладной уровень [1] с помощью протоколов SOAP/HTTP по интерфейсам MM1, MM3, MM4 и MM7, и затем из общего потока вычлняются отдельные объекты: текст, речь, потоковый звук, фотографии, битовые карты, векторные рисунки, видео и т.д. Они анализируются по многим критериям: подозрение на спам, продвижение по каналам с заданным уровнем безопасности и качества обслуживания, содержимое текста и изображений. Также поток тестируется на наличие корреляции данных для увеличения скорости передачи за счет удаления дублирующихся данных.

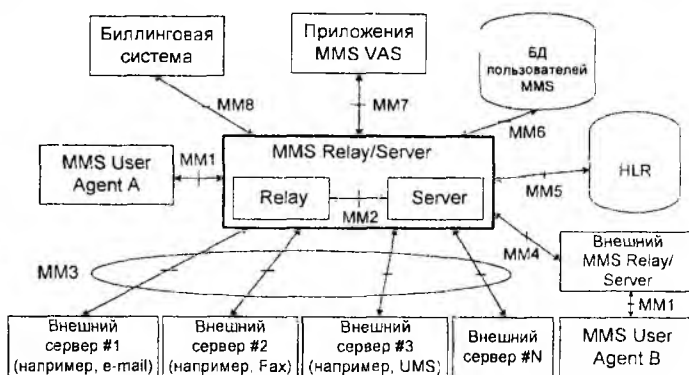


Рис. 1

Особенностью потока HTTP/SOAP является высокая вероятность прохождения по гетерогенным сетям с различными уровнями безопасности. Она обуславливается тем, что программный порт с номером 80 (ассоциированный с обработчиками HTTP), как правило, не блокируется, а поток передается в сеть после процедуры переназначения адресов согласно внутренней структуре сети. При этом, если над объектом производится декомпозиция и подьекты передаются по различным маршрутам, то можно с высокой степенью уверенности утверждать, что количество параметров, от которых зависит вероятность сборки на конечном (промежуточном) сервере, не увеличится.

Поток данных состоит из последовательности комплектов объектов $\{O_1, O_2, O_3, \dots, O_N\}$ переменной мощности, зависящих от времени, размером $\{S_1, S_2, S_3, \dots, S_2, \dots, S_N\}$. Состав комплекта фиксируется для определенного момента времени t . Последовательность комплектов фиксируется на протяжении периода τ . Тогда трафик, обслуживаемый маршрутизатором в определенный момент времени t , можно представить в виде

$$\{O_1^t, O_2^t, O_3^t, \dots, O_2^t, \dots, O_N^t, ST^t\}, \quad (1)$$

где ST – служебный трафик, необходимый для поддержания функционирования той части инфраструктуры, которая обслуживает объекты $\{O^t\}$.

Временные последовательности таких величин, как ресурсы, необходимые для поддержания функциональности интерфейсов ММ1, ММ3, ММ4 и ММ7, можно исследовать с помощью метода нормированного размаха или метода Херста [2]. Такие последовательности измерений характеризуются показателем H , называемым показателем Херста. Запись измерений представляет собой кривую фрактальной размерности $D = 2 - H$.

Оптимальным является такое выделение ресурсов, верхний порог которого никогда не превышает. В течение определенного промежутка времени t маршрутизатор принимает трафик $\xi(t)$, одновременно генерируя исходящий трафик $\langle \xi \rangle_\tau$. Исходящий трафик не обязательно должен быть идентичен по составу входящему: применение политики безопасности и добавление дополнительной информации, необходимой для маршрутизации, меняет содержимое передаваемых данных.

Средний «расход» ресурса (процессорного времени, оперативной памяти, сетевых соединений и т.д.), выделяемого для обработки трафика в системе за выбранный период времени τ , равен:

$$\langle R \rangle_\tau = \frac{1}{\tau} \cdot \sum_{t=1}^{\tau} \xi(t) \cdot r(t), \quad (2)$$

где $r(t)$ – ресурс, необходимый для обработки трафика, поступающего в момент t .

Это среднее должно равняться среднему объему высвобождаемого маршрутизатором ресурса в течение каждого такого указанного промежутка времени. Пусть $X(t)$ – накопившееся отклонение $\xi(t) \cdot r(t)$ от среднего $\langle R \rangle_\tau$:

$$X(t, \tau) = \sum_{u=1}^t \{ \xi(u) \cdot r(u) - \langle R \rangle_\tau \}. \quad (3)$$

Разность максимального и минимального выделенного X :

$$R(\tau) = \max_{1 \leq t \leq \tau} X(t, \tau) - \min_{1 \leq t \leq \tau} X(t, \tau), \quad (4)$$

где t – дискретное время, принимающее целочисленные значения; τ – длительность рассматриваемого промежутка времени.

Размах зависит от рассматриваемого периода τ , и мы ожидаем, что R растет с τ .

Для обработки разнородных данных будем использовать безразмерное отношение R/S , где S – квадратный корень из дисперсии:

$$S = \left(\frac{1}{\tau} \sum_{t=1}^{\tau} \{ \xi(t) - \langle \xi \rangle_\tau \}^2 \right)^{1/2} \quad (5)$$

Для многих временных рядов наблюдаемый нормированный размах R/S очень хорошо описывается эмпирическим соотношением

$$R/S = (\tau/2)^H. \quad (6)$$

Показатель Херста H более или менее симметрично распределен вокруг среднего значения 0,73 со стандартным отклонением, равным примерно 0,09.

Для повышения точности анализа необходимо находить все показатели для каждого из сетевых элементов и типа трафика. В реальной ситуации как правило рассматриваются группы элементов и совокупности типов трафика.

При отсутствии долговременной статистической зависимости отношение R/S должно быть асимптотически пропорционально $\sqrt{\tau}$, если временные ряды связаны со случайными процессами с независимыми значениями и конечной дисперсией $R/S = (\pi\tau/2)^{1/2}$. Основным показателем при маршрутизации является загруженность (виртуальных) каналов передачи данных, и поэтому все показатели вычислялись, исходя именно из него.

Для практических исследований по загруженности каналов использовался сервер компании «Технаучсервис». В качестве сканера использовалась программа MRTG (рис. 2, а, б, в). Графики г, д, е были получены усреднением во времени уровня загруженности канала. Данные по отдельным направлениям и протоколам обособливались дополнительными системными утилитами (снифферами).

Графики за день, неделю и месяц являются типичными и статистически подобными (при одинаковых эксплуатационных условиях оборудования и стабильности активности клиентов). Это позволяет оптимизировать процесс перераспределения нагрузки путем использования алгоритмов, основанных на использовании статистики: предсказание нагрузки путем поиска и наложения шаблонов.

Для маршрутизации исследовалась только та часть данных, которая поступала по протоколу SOAP/HTTP. Как показали наблюдения, долгосрочное прогнозирование соответствует реальной ситуации, поскольку оборудование статистически подобно реагирует на ситуации в сети и поэтому применение шаблонов, основанных на временном соответствии, корректно.

Данная статистическая информация (и производная от нее) может использоваться в дальнейшем для организации более сложных подсчетов. Возможно вычленение информации, относящейся к определенным группам сетевых элементов и маршрутизаторов и составление из них взаимосвязанных групп (кластеров). Фрактальная размерность кластера служит количественной характеристикой одной из особенностей кластера, а именно, заполнения им определенной части канала и занятия вычислительного ресурса.

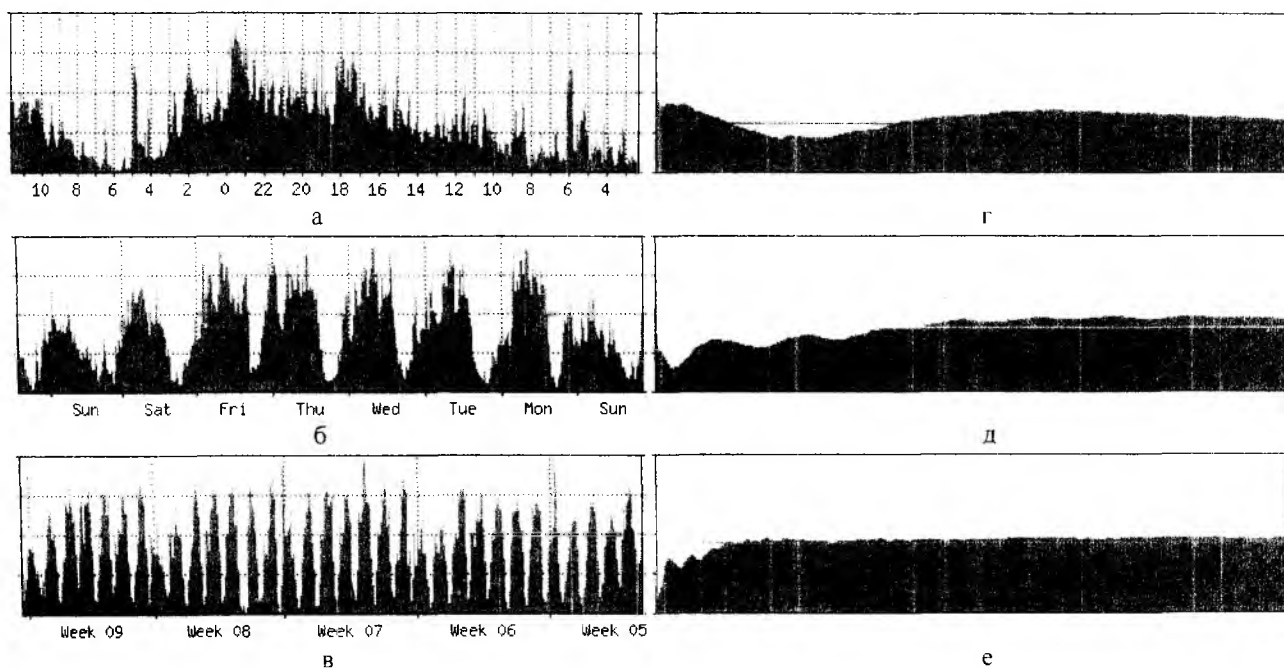


Рис. 2

Оценка ситуации производится с помощью показателя H . Для белого гауссовского шума этот показатель равен $1/2$. Одновременно анализируются участки статистики нескольких временных масштабов. Если для наибольшего масштаба показатель H начинает изменяться (что свидетельствует об изменении корреляции), то система управления маршрутизатором

восстанавливает участки для меньшего масштаба, производит композицию и отслеживает изменение показателя на смежных и подобных участках. Анализ производится с учетом ситуации на нескольких маршрутизаторах и сетевых элементах. При совпадении текущей ситуации (с установленным порогом соответствия) в сети с ранее происходившими событиями выносится решение о дальнейшем управлении в соответствии ранее предпринятыми мерами. При этом, если меры, предпринятые ранее, не способствовали равномерному рассредоточению нагрузки в сети, принимаются корректирующие действия, направленные на улучшение решения. В этом выражается обучаемость маршрутизатора.

Одной из проблем, возникающих при функционировании системы по таким алгоритмам, является вычислительная сложность анализа. Это вызвано как значительным количеством элементов в сети, так и необходимостью анализировать продолжительные временные интервалы. Поэтому для реализации алгоритмов необходимо использовать оптимизационные подходы, например, использовать потоковые инструкции процессоров и распределение вычислительной нагрузки между процессорами (если используются мультипроцессорные системы).

Потоковые инструкции позволяют обрабатывать массив данных за одну инструкцию (технология SIMD – Single Instruction Multiple Data). Поскольку в алгоритмах используются операции усреднения, сложения, умножения и сравнения над большими блоками данных, происходит значительное увеличение их скорости работы. Для проводившихся исследований было получено ускорение функционирования алгоритмов: при использовании MMX-инструкций – в среднем в 1,9 раз; SSE2-инструкций – в 3,7 раза.

Хранение данных на маршрутизирующем сервере и обмен данными с элементами сети предусматривает их предварительное сжимающее кодирование. В системе используются два типа кодирования статистической последовательности: кодирование, основанное на сегментированных итерационных функций (PIFS) [3] и основанное на поиске по словарю. Первый способ используется для формирования словаря и заключается в разбиении интервала выборки на участки и нахождении на них подобия. Кодирование по словарю заключается в индексации фрагментов статистики (восстановление осуществляется при помощи алгоритмов прямого или сужающегося поиска). Здесь обнаруживается существенный недостаток фрактального статистического анализа – необходимость в накоплении значительного количества данных для выявления точных зависимостей. В некоторых случаях для достижения необходимой точности требуется сбор данных о работе системы в течение нескольких дней или недель.

2 Многокритериальная маршрутизация

Несмотря на множество математических моделей маршрутизации в телекоммуникационных сетях обстановка показывает, что методы решения были не идеальными. В большинстве систем сбор данных, необходимых для принятия решения, производится с помощью оперативной аналитической обработки данных (online analytical processing, OLAP), основывающейся на методах «грубого» разведочного анализа. Для нее присуще использование концепции усреднения по выборке, приводящее к операциям над фиктивными величинами (например, средней интенсивности трафика узла сети).

Методы Data Mining [4] (наиболее популярный метод – анализ на основе нейронных сетей) показывают много дополнительных скрытых статистических зависимостей, однако для их реализации требуется гораздо больше вычислительных ресурсов. Это обусловлено необходимостью анализа данных за продолжительный период времени и значительных пересчетах при поступлении новых данных.

Фрактальные методы статистики позволяют выявить все те же свойства элементов, поступающих в потоке данных, что и Data Mining, однако в то же время требуют значительно меньше ресурсов. Сравнение различных методов приведено в таблице.

	OLAP	Data Mining	Фрактальные методы
Выявление закономерностей:			
ассоциация	да	да	да
последовательность	слабо	да	да
классификация	да	да	да
кластеризация	нет	да	да
прогнозирование	слабо	да	да
Размер выборки	малый	средний	большой
Скорость принятия решения	средняя	низкая	высокая

Маршрутизация на прикладном уровне является многокритериальной и может производиться по следующим параметрам: исходя из загруженности серверов и сетевых элементов, по пункту назначения, по типу трафика, согласно принятой политике безопасности, по временным показателям.

Поскольку маршрутизация осуществляется на прикладном уровне при установленных соединениях, то возможно определить множество путей от отправителя к получателю. Пере-направление трафика по смежной ветви задается с помощью условной вероятности отказа в обслуживании p . Данная вероятность вычисляется, исходя из сложившейся обстановки в сети, которая оценивается маршрутизаторами

$$p = E(\xi(t), Q), \quad (7)$$

где E – функция оценки возможности обслуживания трафика смежным маршрутизатором, Q – вектор параметров, задающий дополнительные характеристики качества обслуживания при взаимодействии двух сетей.

Функция оценки складывается из оценок возможности удовлетворения вышеназванным параметрам:

$$\begin{cases} \xi(t) \leq c, \\ \xi(t) \cdot r(t) \leq R, \\ \text{CONTENT}(\xi(t)) \in \text{CONTENT}, \end{cases} \quad (8)$$

где c – пропускная способность исходящих каналов; R – свободный ресурс на маршрутизаторах; $\text{CONTENT}()$ – функция проверки параметров (например, политики безопасности) на допустимость, использующая в качестве параметров содержимое заголовков сопроводительной информации или вектор Q ; CONTENT – допустимое содержимое.

Оценка производится по всем возможным путям следования трафика. При этом маршрутизатор получает сведения только лишь для смежных ветвей, и, таким образом, отпадает необходимость в сведениях о структуре сети без потери информативности.

После получения вектора вероятностей становится возможным вычленение тех путей, вероятность прохождения трафика по которым низка. Разбиение трафика на порции производится, исходя из нескольких правил: 1) неделимость простейших (атомарных) объектов; 2) приоритетность трафика; 3) возможность промежуточной сборки разделенных объектов; 4) использование свойств триадного канторовского множества [2, 5].

Использование четвертого правила позволяет сбалансировать нагрузки в сети: трафик передается не только по наименее загруженным маршрутам, но также и пропорционально распределению ресурса, необходимого для обслуживания нагрузки в сети. Одним из важнейших свойств канторовских множеств, которое используется маршрутизатором, является свойство сохранять некоторые характеристики целого при его разбиении на части. Это свойства равенства фрактальной размерности, размерности подобия и размерности кластера (массы).

Обмен статистикой в разнородной среде производится с помощью технологии XML. Это вызвано универсальностью технологии и малым (сжатым) объемом данных (коэффициенты преобразований и параметры неподвижной точки). Последовательность данных при обмене определяется XML-схемами.

Выводы

Для повышения устойчивости работы современных сетей необходима разработка нового программного обеспечения, особенно программных маршрутизаторов. Анализ содержимого должен проводиться не только общепринятыми методами математической статистики, но и методами, позволяющими найти более сложные закономерности, выделить приоритетные группы и предсказать поведение сети для значительного интервала времени. Необходим анализ, позволяющий обнаружить шаблоны, отражающие фрагменты многоаспектных взаимоотношений в поступающих данных. При применении методов, основывающихся на использовании результатов суперпозиции шаблонов, становится возможным равномерное распределение нагрузки в сети.

Сложность выбора алгоритмов и программных средств обуславливается необходимостью проведения анализа в реальном масштабе времени. Применение фрактальных методов на всех этапах анализа трафика и принятия решения о дальнейшей маршрутизации позволяет унифицированно решить все задачи. Этот подход позволяет уменьшить скорость разработки программного обеспечения за счет использования небольшого количества алгоритмов. Также следует отметить значительное увеличение скорости принятия решения за счет использования потоковых инструкций и распараллеливания в многопроцессорных системах.

Список литературы: 1. 3GPP TS 23.140: «Multimedia Messaging Service; Functional Description; Stage 2». 2. Федер Е. Фракталы. М.: Мир, 1991. 254 с. 3. Barnsley M., Hurd L. Fractal Image Compression. San-Diego: Academic Press, 1988. 4. Дюк В., Самойленко А. Data mining: Учебный курс. СПб.: Питер, 2001. 368 с. 5. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Наука, 1976. 544 с.

*Харьковский национальный
университет радиоэлектроники*

Поступила в редколлегию 02.04.2004