

ВМАР-ПОТОК КАК МОДЕЛЬ ТРАФИКА РЕАЛЬНОЙ СЕТИ

Г. В. Царенков

Белорусский государственный университет

Минск, Беларусь

g.tsarenkov@sam-solutions.net

Целью данной работы является разработка подхода, позволяющего на основании данных о трафике реальной телекоммуникационной сети перейти к модели группового марковского входного потока (Batch Markovian Arrival Process – ВМАР), чтобы в дальнейшем использовать его для анализа характеристик производительности телекоммуникационных сетей с использованием аппарата теории массового обслуживания.

Ключевые слова: входной поток, ВМАР, моделирование.

1. ВВЕДЕНИЕ

В последнее время большое количество работ посвящено исследованию систем массового обслуживания (СМО), в качестве входного потока в которых используется групповой марковский входной поток (Batch Markovian Arrival Process – ВМАР). Этот поток ввел в рассмотрение Д. Лукантони в 1991 году в работе [1] как развитие идеи разносторонних потоков М. Ньютса [2]. Однако опубликованные работы по аналитическому исследованию с применением аппарата ВМАР-потоков в большинстве случаев либо не содержат численных примеров, либо входной поток системы конструируется по заданным характеристикам, например средняя интенсивность, коэффициенты корреляции или вариации, но никак не на основании измеренного трафика реальной сети.

В первую очередь это вызвано большим количеством параметров, которые необходимо определить для конструктивного построения модели ВМАР-потока. В отличие от пуассоновского потока, для которого достаточно проверить гипотезу об экспоненциальном распределении длительности интервалов времени между моментами поступления сетевых пакетов, а затем при помощи известных методов вычислить значение интенсивности входного потока, для ВМАР-потока возможны несколько уровней значения интенсивности входного потока. Совместно с возможностью поступления групп требований это усложняет определение переходных вероятностей управляющего процесса и значений интенсивности поступления пакетов в каждом из режимов. В то же время имеется ряд работ, в которых высказываются рекомендации по переходу от измеренного трафика к моделям СМО с различными входными потоками.

2. ПОСТАНОВКА ЗАДАЧИ И ОПУБЛИКОВАННЫЕ РЕЗУЛЬТАТЫ

Современные сети небольших компаний в основном базируются на протоколе Ethernet 10М или 100М, предоставляющем оптимальное значение параметра по цене оборудования, монтажа и обслуживания к скорости передачи информации. Размер пакета в таких может варьироваться от 60 до 1500 байт и изменяться динамически (хотя доминирующими значениями все же являются размеры пакетов в 60, 576, 1500 байт).

Широкое распространение также имеют АТМ-сети. В этих сетях передаваемый пакет имеет строго ограниченную длину, равную 60 байтам, 5 из которых предназначены для служебной информации.

Анализ трафика и переход к моделям теории массового обслуживания в мульти-сервисных IP-сетях с применением техники, учитывающей информацию об оконечных пользователях (например, такую как скорость коммутируемого доступа каждого пользователя совместно с интервалами их подключения к сети), можно найти в работе [3]. Результаты исследований IP-трафика у поставщика услуг Интернет университета Дортмунда опубликованы в работе [4]. Эти результаты также эксплуатируют информацию о пользователях, такую как скорость подключения, тарифные планы, а также асимметричность доступа для входящего и исходящего трафиков пользователя. В работе [5] проводилось смешанное исследование IP-трафика на уровне TCP и на уровне HTTP, что позволило авторам построить модель, в которой запросы на Web-сервер подразделяются на основные (main) и уточняющие (inline). В работе [7] исследуется трафик в фрагменте сети на предмет выявления его взрывного характера (bursty traffic) и предлагается модель для аппроксимации такого трафика, при этом используется аппарат корреляционного анализа и фильтров на основе скользящих средних. В работе [8] рассматривается процесс оценки пропускной способности сети на основе экспоненциальных скользящих средних.

В данной работе предлагается подход, который позволит осуществить переход от реально измеренного трафика в сети switched Ethernet 100 МБит/с к модели СМО на основе *ВМАР*-потока. Для данного эксперимента проводилось наблюдение в течение недели за трафиком сервера централизованного хранилища исходных кодов и управления проектами. Результаты этих измерений представляют собой ряд значений переданного/принятого количества информации в течение последовательных интервалов длительностью 2 секунды. Основное рабочее время приходится на сервер в период с 10.00 до 11.00 ежедневно. Съём информации о трафике происходил при помощи стандартной службы "Quality of Service" операционной системы Microsoft Windows 2000.

3. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ВМАР-ПОТОКА

Для определения ВМАР-потока рассмотрим марковскую цепь $v_t, t \geq 0$, с непрерывным временем и пространством состояний $\{0, \dots, W\}$. Следуя Д. Лукантони, введем матричную производящую функцию

$$D(z) = \sum_{k=0}^{\infty} D_k z^k,$$

где $D_k = [d_{v,v'}^{(k)}]$, $k > 0$, – квадратные матрицы размерности $(W + 1) \times (W + 1)$ характеризуют интенсивности переходов цепи v_t из состояния v в состояние v' , сопровождающиеся поступлением группы требований размера k ,

$$D_0 = \left[\begin{array}{l} d_{v,v}^{(0)} = -\lambda_v, v = \overline{0, W}, \\ d_{v,v'}^{(0)}, v, v' = \overline{0, W}, v \neq v' \end{array} \right],$$

где λ_v – средняя интенсивности переходов цепи v_t в любое состояние (возможно, в то же самое), когда она находится в состоянии v , а $d_{v,v'}^{(0)}$ характеризует интенсивности переходов цепи v_t из состояния v в состояние v' , которое не сопровождается приходом ни одного требования (естественно, полагается, что переход в то же самое состояние без поступления требования не возможен).

4. ПЕРЕХОД ОТ ИЗМЕРЕННЫХ ДАННЫХ К МАТЕМАТИЧЕСКОЙ МОДЕЛИ

Данные реальных измерений объема трафика представлены на рисунке. Очевидно, что трафик не носит характер простейшего, причем интенсивности значительно различаются в зависимости от момента времени.

После применения фильтра на скользящих средних [7] можно выделить пять следующих уровней интенсивности поступления требований (см. рис.).

- 1) Самый низкий уровень трафика соответствует обмену служебными пакетами между сервером, контроллером домена, WINS-сервером и др. При построении ВМАР-потока считаем, что запросы в эти интервалы не поступают.
- 2) Данный уровень соответствует обращениям к базе данных, призванным определить, были ли в ней какие-либо обновления. При построении ВМАР-потока этот уровень представляется интенсивностью $\lambda_0 = 1,5 \times 10^2$.
- 3) Данный уровень соответствует обслуживанию запросов по записи небольших порций информации. В ВМАР-потоке интенсивность запросов этого уровня λ_1 полагается $4,5 \times 10^2$.
- 4) Данный уровень соответствует обслуживанию запросов по записи небольших порций информации. В ВМАР-потоке интенсивность запросов этого уровня λ_2 полагается 16×10^2 .
- 5) Данный уровень соответствует обслуживанию запросов по чтению небольшого количества записей в базе данных. В ВМАР-потоке интенсивность запросов этого уровня λ_3 полагается 31×10^2 .

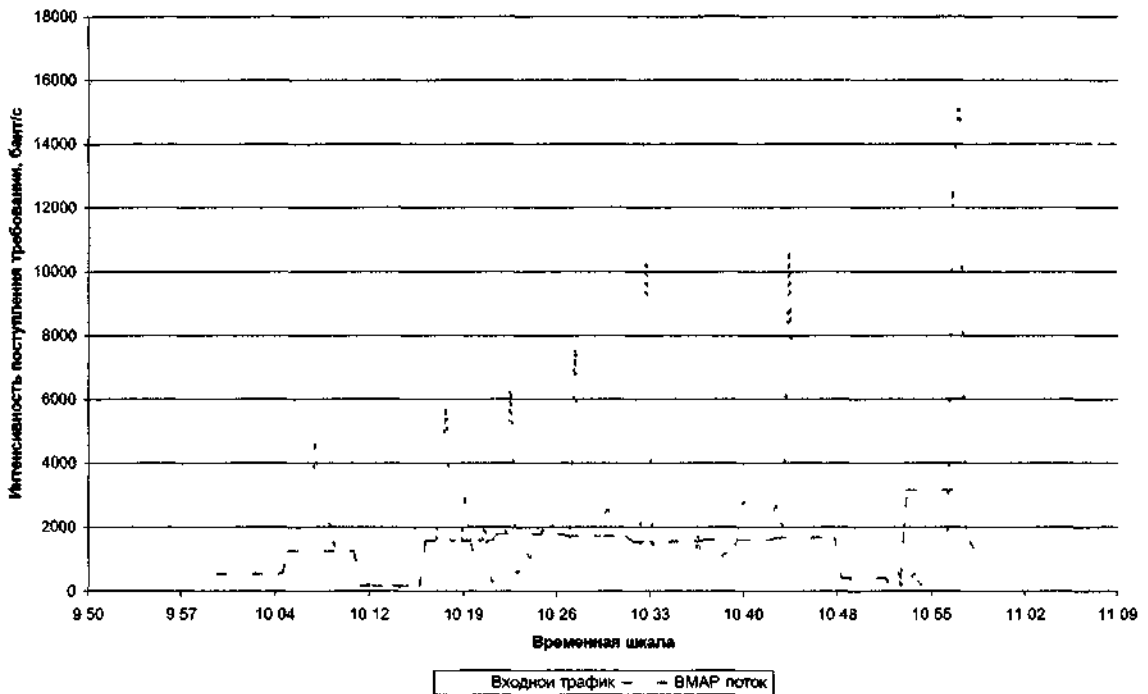


Рис. Входной трафик сервера базы данных $10^{00}-11^{00}$ и его аппроксимация ВМАР-поток

Таким образом, управляющий процесс ВМАР-потока имеет пространство состояний $\{0, \dots, 4\}$.

Попутно отметим, что, например, в маршрутизаторах CISCO – одного из ведущих мировых производителей, используется алгоритм анализа трафика на основе 5-минутного скользящего среднего.

На интервалах, на которых сохраняются значения интенсивностей, можно вычислить вероятности поступления групп определенного размера. Для этого делим количество поступившей информации на значение интенсивности в на 2-секундном интервале и округляем до целого числа. Таким образом определяется размер групп.

В нашем случае наибольший размер групп составляет 7. Для определения значений элементов $d_{v,v'}^{(k)}$, $k = \overline{0,7}$, $v, v' = \overline{0,4}$, воспользуемся прямым подсчетом количества переходов между состояниями, сопровождающимися приходом требований. Для перехода к вероятностям делим их на общее количество переходов и рассчитываем элементы матриц D_k , $k \geq 0$, как

$$d_{v,v'}^{(k)} = \begin{cases} -\lambda_v, & k = 0, v, v' = \overline{0,4}, v = v' \\ \lambda_v p_{v,v'}^{(0)}, & k = 0, v, v' = \overline{0,4}, v \neq v' \\ \lambda_v p^{(k)}, & k = \overline{1,4}, v, v' = \overline{0,4} \end{cases}$$

Таким образом можно получить параметры ВМАР-потока. В частности, для приведенного выше примера они будут выглядеть, как приведено в таблице со следующими

характеристиками: интенсивность $\lambda = 3.65154$, интенсивность группы $\lambda_g = 2.3295$, коэффициент корреляции $c_{cor} = -0.0190585$, коэффициент вариации $c_{var} = 0.341751$.

Таблица

Матрица D_0				Матрица D_1			
-1.5308	0	0	0	0.3645	0	1.0935	0
0	-4.5668	0	2.7401	0	0	0	0.9133
1.1267	1.1267	-15.9873	5.6336	0.6395	0.6395	3.1975	0
23.016	0	0	-31.647	2.877	0	0	0
Матрица D_2				Матрица D_3			
0.0182	0	0.0546	0	0	0	0	0
0	0	0	0	0	0	0	0.9133
0.3349	0.3349	1.6748	0	0.0304	0.0304	0.1522	0
0	0	0	0	0	0	0	0
Матрица D_4				Матрица D_5			
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0.0609	0.0609	0.3045	0	0.0304	0.0304	0.1522	0
2.877	0	0	0	2.877	0	0	0
Матрица D_6				Матрица D_7			
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0.0609	0.0609	0.3045	0
0	0	0	0	0	0	0	0

5. ЗАКЛЮЧЕНИЕ

В заключение хотелось бы отметить, что для практического применения имеет большое значение возможность перехода от реально измеренных данных к теоретическим моделям. Это, во-первых, может быть использовано для проверки адекватности применения моделей на практике, а во-вторых, для оптимизации работы действующих сетей в реальном времени. Данная работа является продолжением ряда работ по проведению оценки параметров потока в реальных сетях. Дальнейшим развитием данной работы является разработка метода, который позволит автоматизировать построение моделей ВМАР-потока на основе реальных данных о трафике и проверке качества построенных моделей.

ЛИТЕРАТУРА

1. Lucantoni D. M. New results on the single server queue with a batch Markovian arrival process // Communications in Statistics - Stochastic Models. 1991. V. 7. P. 1–46.
2. Neuts M. F. A versatile Markovian point process // J. Appl. Prob. 1979. V. 16. P. 764–770.
3. Anderlind E., Zander J. A Traffic Model for Non Real-Time Data Users in a Wireless Radio Network // IEEE Communication Letters. 1997. V. 1. P. 37–39.
4. Kilpi J., Norros I. Call Level Traffic Analysis of a Large ISP // Proc. 13th ITC Specialist Seminar on Measurement and Modeling of IP Traffic. Monterey CA. 2000. 6.1–6.9.