УДК 681.396.626

Казакова Н.Ф., Международный гуманитарный университет (г. Одесса),
к.т.н., декан ф-та Компьютерных наук
и инновационных технологий
Кушер А.Н., ГУИКТ (г. Киев), аспирант

ПРОБЛЕМЫ РАЗРАБОТКИ И АНАЛИЗА ТЕОРЕТИЧЕСКИХ ПОДХОДОВ К МОДЕЛИРОВАНИЮ ТРАФИКА ЗАЩИЩЕННЫХ ИНФОРМАЦИОННЫХ СЕТЕЙ С ТОЧКИ ЗРЕНИЯ ИХ СООТВЕТСТВИЯ СЕТЯМ НОВОГО ПОКОЛЕНИЯ

Освещены основные положения о проблемах разработки и анализа теоретических подходов к моделированию трафика с точки зрения их соответствия сетям нового поколения.

Substantive provisions are lighted up about the problems of development and analysis of theoretical approaches to the design of traffic from point of their accordance to the networks of new generation.

Современные теоретические методы и подходы описания процессов в информационных системах и сетях разнообразны и требуют научного осмысления для применения в инженерной практике. Все большее значение приобретают теоретико-вероятностные методы исследований, основанные на вероятностной трактовке протекающих в информационных системах процессов. Статистический подход позволяет более полно учесть состояние динамической системы, характер управляющих и возмущающих воздействий, результирующее поведение информационных потоков в многофункциональных сетях и во многих случаях более адекватен для решения различных практических задач.

Круг вытекающих из указанного подхода проблем достаточно широк:

- *описание математических моделей случайных процессов в информационных системах*
- *формирование на базе математических моделей статистических методов проверки гипотез и обнаружения, оценивания и фильтрации*
- *интерполяции (сглаживания) и экстраполяции (прогнозирования)*
- *разработка алгоритмов оптимального управления стохастическими системами.*

Основными критериями при выборе теоретического подхода моделирования сети являются следующие:

- *модель должна быть пригодной для описания современных высокоскоростных защищенных коммерческих сетей;*
- *модель должна быть универсальной, т.е. адаптируемой к различным типам сетей или их фрагментам;*
- *точность соответствия реальным потокам данных должна быть не более 10% при приемлемой вычислительной громоздкости;*

– *модель должна быть пригодной для прогнозирования поведения сети при более интенсивном трафике;*
– *модель должна отображать маршруты трафика и их изменение.*

Для того, чтобы определить, какой из теоретико-вероятностных методов наиболее близок к поведению высокоскоростных защищенных коммерческих сетей необходимо проанализировать существующие. Задача усложняется тем, что большинство авторов осуществляли оценку точности модели при использовании данных локальных сетей, что не гарантирует пригодность такого подхода для защищенных коммерческих сетей. Другими словами, необходима дополнительная проверка предлагаемых подходов на данных измерения трафика защищенных коммерческих сетей. Результаты предлагаемых решений проблемы приводятся в докладе.

Поскольку основной математической моделью процессов, описывающих трафик в информационных сетях, является случайный поток данных, вполне оправданны попытки создания теоретической модели трафика на основании статистической теории. Случайный поток в рассматриваемом практическом приложении обладает следующими основными свойствами:

– независимость вероятностных характеристик от времени (стационарность);

– зависимости вероятностей событий (случайные процессы с памятью);

– бесконечно малая вероятность более одного события за бесконечно малый интервал времени.

Поток как случайный процесс характеризуется своими статистическими свойствами. Чаще всего используются: плотность вероятности поступления данных за период, функция вероятности потока и автокорреляционная функция.

Классической моделью трафика в информационных сетях является Пуассоновский поток. Он характеризуется набором вероятностей $P(k)$ поступления k сообщений за временной интервал t:

$$P(k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}, \qquad (1)$$

где: k=0,1, … – число сообщений; λ – интенсивность потока.

Заметим, что интервал времени измерения количества сообщений t и интенсивность потока λ являются постоянными величинами.

Для семейства Пуассоновских распределений (1) большее значение λ соответствует более широкому и симметричному графику плотности и большему объему информационных потоков. Зная вероятность поступления данных за период, можно получить распределение интервала τ между соседними событиями: $P(\tau) = \lambda e^{-\lambda t}$.

Основным свойством пуассоновского потока, обусловливающим его широкое применение при моделировании, является аддитивность: результирую-

щий поток суммы пуассоновских потоков тоже является пуассоновским с суммарной интенсивностью $\lambda_\Sigma = \sum_{n=1}^{N} \lambda_n$.

Однако, применение статистики (1) с учетом многофункциональности современных сетей возможно исключительно для описания очередей пакетов. Динамические процессы, происходящие в современных сетях, имеют сложную природу и относятся к стохастическим процессам. Такие свойства трафика возникают из-за недетерминированности системы в целом. Другими словами, долгосрочное прогнозирование действий, осуществляемых обрабатывающими трафик алгоритмами, не возможно предсказать, как впрочем, и массу других воздействующих на трафик факторов. Как отмечалось ранее, трафик обрабатывается алгоритмами, используемыми в различных реализациях протоколов семейства TCP/IP: генерация трафика протоколами транспортного уровня, управление трафиком на промежуточных сетевых устройствах, динамическая маршрутизация и т.д. В результате, процессы в компьютерных сетях находятся под постоянным влиянием регулирующих и возбуждающих стохастических воздействий, обуславливающих сложные флюктуации исследуемых процессов. Другими словами, требуется модель, являющаяся случайным процессом, управляемым другим случайным процессом.

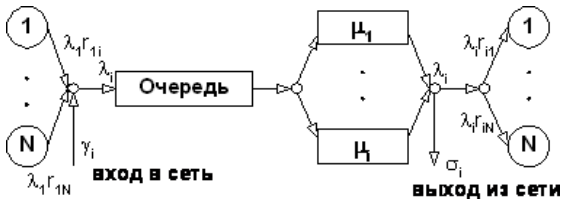

Рис. 1 – Структура моделирования очередей трафика: r_{ij} – вероятность входа/выхода; λ_i – полная интенсивность потока; γ_i – параметр управления.

В научных исследованиях последних лет отдается предпочтение описанию статистических свойств сети на основе применения Марковских скрытых цепей управляемых пуассоновской статистикой для описания очередей потоков (рис. 1).

Дискретный случайный Марковский поток характеризуется следующими свойствами: $P\left(X_{n+1}=j \mid X_n=i_n, X_{n-1}=i_{n-1},\ldots\right)=P\left(X_{n+1}=j \mid X_n=i_n\right)$, $P\left(X_{m+1}=j \mid X_n=i\right)=p_{ij}$, где: p_{ij} – вероятность перехода из состояния i в состояние j, $p_{ij} \geq 0$ для $\forall i, j$ и $\sum_j p_{ij}$=1 äëÿ $\forall i$. Процесс перехода некоторого дискретного случайного процесса из одного состояния в другое описывается матрицей переходов:

$$P=\begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots & \cdots \\ \vdots & \vdots & \vdots & & \\ p_{i0} & p_{i1} & p_{i2} & \cdots & \cdots \\ \vdots & \vdots & \vdots & & \end{pmatrix}.$$

Наиболее адекватно отобразить стохастические переходы сети из одного состояния в другое позволяет модель скрытых цепей Маркова, СМЦ (рис. 2).

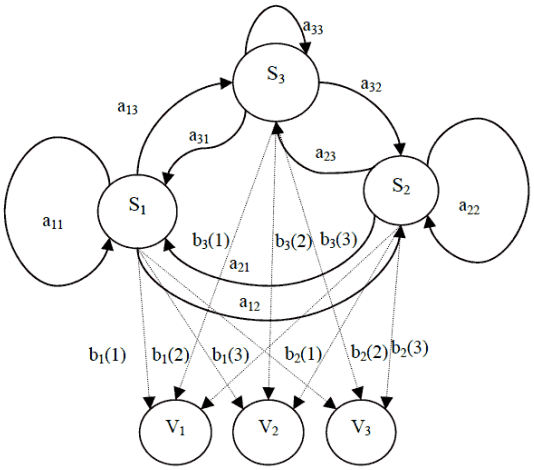

Рис. 2 – Граф скрытой цепи Маркова

Для определения модели на основе СМЦ необходимо задать число возможных состояний N, объем символов используемого алфавита, M. Для некоторого конечного алфавита можно определить конечный набор вероятностей перехода $\Lambda = \left(a_{ij}\right)$: $a_{ij} = P\left(s_{t+1} = j \mid s_t = i\right)$, $1 \le i, j \le N$, где: s_t описывает текущее состояние.

Для вероятностей перехода $\sum_{j=\&}^{N} a_{ij} = 1$, $a_{ij} \ge 0$, $1 \le i, j \le N$. Функция вероятности возможных состояний $B = \{b_j(k)\}$:

$b_j(k) = P(o_t = \nu_k / s_j = j)$, $1 \le j \le N$, $1 \le k \le M$;

$b_j\left(k\right) = P\left(o_t = v_k \mid q_t = j\right)$, $1 \le j \le N$, $1 \le k \le M$,

где v_k определяет k^{th} наблюдаемый символ алфавита и o_t отображает текущее состояние. Для описанной стохастической модели: $b_j\left(k\right) \ge 0$; $1 \le j \le N$, $1 \le k \le M$, $\sum_{k=1}^{M} b_j\left(k\right) = 1$, $1 \le j \le N$.

Если на основе, например, пуассоновской статистики определяется случайный процесс инициализации, $\pi = (\pi_i)$, где $\pi_i = P\left(s_1 = i\right)$, $1 \le i \le N$, то для определения СМЦ необходимо задание $\lambda = \left(\Lambda, B, \pi\right)$.

Таким образом, для рассматриваемого практического приложения, при заданном λ и наблюдениях $Y = y_1, y_2, \ldots, y_T$, модель позволяет оценить условную вероятность $P\{Y|\lambda\}$. Проблема моделирования будет заключаться в нахождении таких параметров модели $\{\Lambda, B, \pi\}$, которые позволят максимизировать вероятность $P\{Y|\lambda\}$.

Однако, несмотря на вычислительную сложность реализации данного подхода при моделировании даже локальных сетей проверка точности модели на данных реальных измерений высокоскоростных сетей дает большие расхождения модели реальных оценок трафика. Исследования последнего десятилетия показали несостоятельность классических методов оценки вероятностно-временных характеристик сетей пакетной коммутации для защищенных коммерческих сетей [1…4]. Усложнение модели посредством дополнения итерационных процедур для коррекции оценки параметров позволяет улучшить точность. Но открытыми остаются вопросы инициализации алгоритмов.

Стохастическая динамика реальных сетевых процессов (рис. 3) обычно не принимается во внимание при разработке информационных приложений в локальных сетевых сегментах, где вариации параметров могут быть строго ограничены. Однако при создании глобальной инфраструктуры компьютерных телекоммуникаций методы прогнозирования состояния виртуальных соединений становятся весьма актуальными. Широкие перспективы применения адаптив-

ных транспортных протоколов для мультимедиа приложений стимулируют разработку новой технологии моделирования трафика современных сетей.

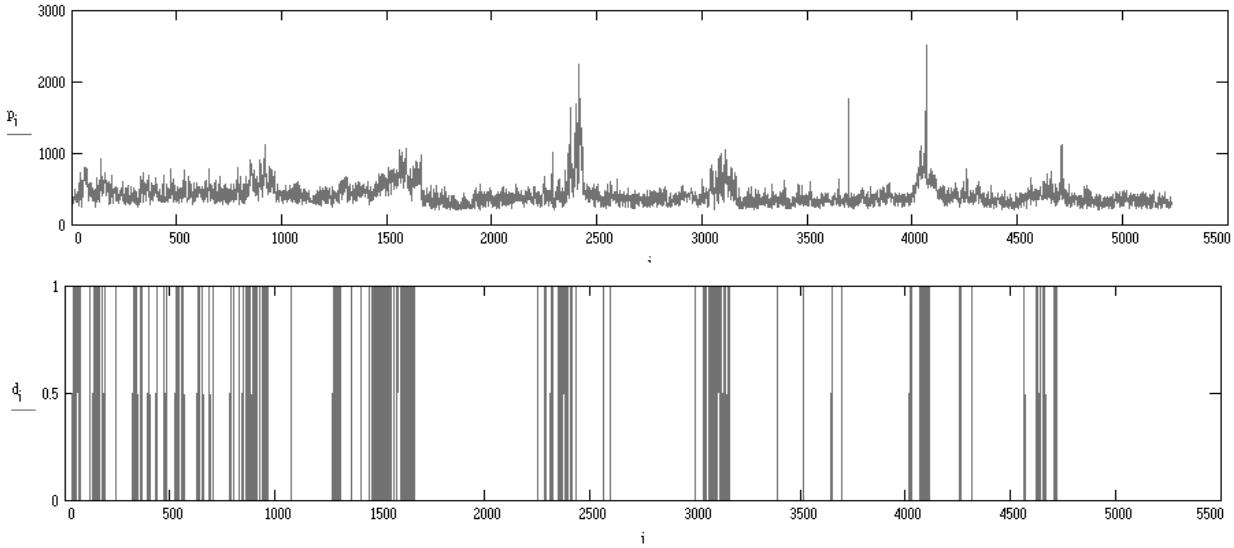

Рис. 3 – Характер изменения задержек при передаче данных: верхний рисунок – абсолютные значения интервалов времени между отправкой пакета и получением подтверждения, мсек; нижний рисунок – структура потока, превышающего пороговое значение 600 мсек

С другой стороны, если протоколы маршрутизации используют только локальную информацию о связанности узлов пакетной коммутации, то для них индикатором разрыва соединения является увеличение объема выходной очереди и ее возможное переполнение, если время перекоммутации виртуального соединения превышает определенное пороговое значение. В силу случайного характера рассматриваемых событий их вероятностное моделирование должно отражать возможность появления подобных резких отклонений.

Важным выводом из проведенного анализа является необходимость применения комбинированного подхода для статистического описания трафика современных сетей.

Вероятностное распределение анализируемых сигналов, например, времени подтверждения, может иметь вид аддитивной смеси двух или более распределений с различными наборами параметров:

$$P(x) = \sum_{j=1}^{m} p(x \mid j)\, p(j), \qquad (2)$$

где: $p(x|j)$ – плотность вероятности j-ой компоненты; $p(j)$ – весовой коэффициент учета j-ой компоненты в общей суперпозиции.

Для подхода (2) выполняется: $0 \le p(j) \le 1, \quad \sum_{j=1}^{M} p(j) = 1, \quad \int p(x \mid j)dx = 1$.

Открытым остается вопрос о том, какое семейство плотности вероятности нужно применять в качестве составляющих (2) для достижения наилучшей точности модели. Очевидно, что ответ на этот вопрос можно получить только на основании изучения статистических свойств реального трафика защищенных коммерческих сетей.

Литература

1. Маракова И.И., Скопа А.А., Сыропятов А.А. Комплексная защита информации в беспроводных системах связи // Матер. IV наук.-конф. Департамента спец. телеком. систем та захисту інформ. та Служби безпеки «Правове, нормативне та метрологічне забезпечення систем захисту інформації в Україні». – К.: НДЦ «Тезис» НТУУ «КПІ». – 2007. – С.73-75.

2. Кашанов И.В., Шамин П.Ю. Симуляция сети с переменной топологией с использованием параллельных вычислений // Успехи современного естествознания: Российская Академия Естествознания. – №8. – 2008.

3. Потапов М.В., Сиропятов А.О., Оценка эффективности информационной защиты комплексных систем связи // Управління проектами та розвиток виробництва: Вісник СНУ ім. В. Даля. – Луганськ: СНУ ім. В. Даля. – 2006. – 7 стор.

4. Хорошко В.А., Шелест М.Е., Маракова И.И., Сыропятов А.А. Защита информации в беспроводных системах связи // Захист інформації. – К.: ДУІКТ. – 2005. – №3 (25) – С. 83-91.