О ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ САМОПОДОБНЫХ СЕТЕВЫХ ПРОЦЕССОВ

Соловьев А.Ю.

Старооскольский технологический институт

Старый Оскол, Россия

До недавнего времени в задачах проектирования и прогнозирования вычислительных сетей связи, теоретической основой служил аппарат теории массового обслуживания. При этом моделируемые сетевые потоки рассматривались как простейшие, т.е. потоки обладающие свойством стационарности, ординарности и отсутствием последействия. В 1993 году западными исследователями впервые были опубликованы результаты работы, которая изменила представление о процессах, происходящих в телекоммуникационных сетях. Оказалось, что сетевой трафик является самоподобным (self-similar) или фрактальным (fractal) по своей природе, т.е. в нем присутствуют так называемые вспышки или пачки (burst) пакетов, наблюдаемые в различных временных интервалах (от миллисекунд до минут или даже часов) и корреляция между пакетами.

На практике случайные процессы сохраняют свойство самоподобия только до определенного предела. Этот предел или мера статистической устойчивости процесса при многократном масштабировании определяется так называемым ***параметром Херста H*** или параметром самоподобия. Дадим определение самоподобного процесса. Пусть дан стационарный случайный процесс дискретного аргумента (времени) $X=(X_1, X_2, ..., X_t)$. Обозначим через $r(k)$ коэффициент корреляции. Обозначим через $X^{(m)}=(X_1^{(m)}, X_2^{(m)}, ..., X_t^{(m)})$ усредненный по блокам длины m процесс X, компоненты которого определяются равенством:

$$X_t^{(m)}=\frac{1}{m}(X_{tm-m+1}+...+X_{tm}),\quad m,t\in N$$

В дальнейшем будем называть такой ряд ***агрегированным***. Соответственно $r_m(k)$ коэффициент корреляции процесса $X^{(m)}$. Процесс X является ***строго самоподобным в широком смысле слова,*** с коэффициентом Херста *H,* если выполняется условие:

$$r_m(k)=r(k), m\in\{2,3,...\}$$

То есть самоподобный процесс не меняет свой коэффициент корреляции после усреднения по блокам. Другими словами процесс $X^{(m)}$ не отличим от процесса *X,* как минимум в отношении статистических характеристик второго порядка.

Так же существуют ***асимптотически самоподобные процессы***, главное свойство этих процессов заключается в том, что при $m \to \infty$, процесс сходится к строго самоподобному:

$$r_m(k) \to r(k), m \to \infty$$

Значение коэффициента $H=0,5$ указывают на отсутствие долгосрочной зависимости. Корреляция между событиями отсутствует. Ряд является случайным, а не фрактальным. Чем ближе значение H к 1, тем выше степень устойчивости долгосрочной зависимости. При $0<H\leq 0,5$ временной ряд является трендонеустойчивым. Он более изменчив, чем случайный ряд, поскольку состоит из частых реверсов. При $0,5<H\leq 1$ ряд обладает длительной памятью. Тенденция его изменения может быть спрогнозирована.

Современный сетевой трафик, включающий в себя помимо обычного трафика передачи данных, трафик VoIP и видео-трафик, так же обладает свойством самоподобия. При этом как уже говорилось выше, для таких процессов характерны слабо убывающая автокорреляционная функция (АКФ), частые всплески и спады активности, циклическая составляющая. На Рис.1 приведена реализация реального суточного трафика проходящего через канал связи интернет-провайдера ООО «Радиотелеком» г. Старый Оскол. На Рис.2 приведен график АКФ полученного временного ряда.

Рис.1. Реализация суточного трафика в Мбит/сек.

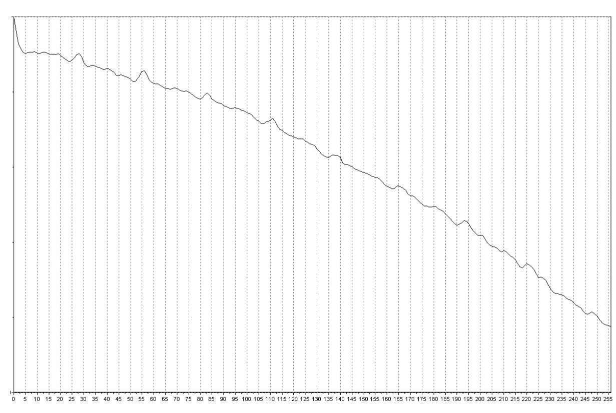

Рис.2. АКФ реализации суточного трафика

Несмотря на достаточно долгое исследование самоподобных процессов в телекоммуникациях, остается ряд нерешенных задач:

- Фактически отсутствует строгая база, которая смогла бы заменить классическую теорию массового обслуживания;
- Нет общепризнанной модели самоподобного трафика;
- Нет единого метода для прогнозирования поведения сетевого трафика.

Задача прогнозирования поведения сетевого трафика имеет особую актуальность на сегодняшний день. Подбор правильно модели для оценивания временного ряда и его прогнозирования позволит решить ряд важных задач, а именно:

- Получать прогнозы о доступности полосы пропускания и соответственно модифицировать стратегии предотвращения перегрузок на уровне транспортного протокола TCP;
- Обеспечить оптимальные временные характеристики, например временную задержку;
- Отследить тенденцию загрузки определенных узлов и сделать прогноз на длительное время (часы, сутки, неделя).

Следует отметить, что именно свойство долгой памяти сетевых процессов дает предпосылки к их прогнозированию и оправдывает применение AR (авторегрессионных) моделей вида:

$$x_t = j_0 + \sum_{k=1}^{p} j_k x_{t-k} + e_t ,$$

где j_k - константы, e_t - белый шум. Для моделирования процессов с длительной памятью широкое распространение получили авторегрессионные модели ARIMA (интегральный процесс скользящего среднего) и ARFIMA (дробовый процесс скользящего среднего). Так же в последнее время широкую популярность приобретает свободный от модели метод сингулярно-спектрального анализа (метод Гусеница).

В дальнейшем направлением исследований является изучение и сравнение моделей и методов для анализа и прогнозов сетевых процессов в частности моделей типа ARFIMA и метода Гусеница.

Список литературы

1. Петров В.В. Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия. Автореферат диссертации. Москва. – 2005. - 20 с.
2. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях: Учеб. пособие. /СПб.: Изд-во СПбГТУ, 2000.- 102 с.