

Черниговский Александр Валерьевич,
аспирант, Ангарский государственный технический университет,
e-mail: chernigovsky.alex@gmail.com

Кривов Максим Викторович,
к.т.н., доцент, зав. кафедрой, Ангарский государственный технический университет,
e-mail: vmk@angtu.ru

ПОДБОР МАТЕМАТИЧЕСКОЙ МОДЕЛИ СЕТЕВОГО ТРАФИКА
Chernigovskiy A.V., Krivov M.V.
SELECTION OF THE MATHEMATICAL MODEL OF NETWORK TRAFFIC

Аннотация. Проведен анализ сетевого трафика, выполнено сравнение полученных результатов с существующими математическими моделями. Выявлено, что наилучшее математическое описание трафика получено с помощью модели Парето.

Ключевые слова: модели сетевого трафика, распределение Пуассона, распределение Парето, самоподобие.

Abstract. Network traffic was analyzed. Obtained results were compared with the existing mathematical models. It was found that the best mathematical description of network traffic was obtained using the Pareto model.

Keywords: network traffic models, Poisson distribution, Pareto distribution, self-similarity.

Общемировая практика насчитывает множество попыток определения математической модели сетевого трафика. Среди них наибольшее распространение получили модели на основе распределения Пуассона (классическая модель Пуассона, модель Пуассона, модулированная Марковским процессом, модель цепочки сообщений), а также ряд моделей, относящихся к самоподобным процессам – модель фрактального броуновского движения, модели Парето и Вейбулла, и другие модели, характеризующиеся медленно убывающей зависимостью [1]. Однако единой модели трафика не существует, а выбор применяемой модели зачастую определяется потребностями исследования.

С учетом того, что современный трафик представляет собой комбинацию данных различных типов (текстовые, видео, графические, и т.д.), актуальным вопросом является определение характеристик сетевых потоков и выбор математической модели, способной наилучшим образом описать трафик с комбинированной структурой.

Нами были проведены исследования мультимедийного, пирингового, потокового трафика (видеоконференций, вебинаров), а также трафика протоколов прикладного уровня (http, https). Данные для анализа трафика были получены как для отдельных ПК, так и для всей сети.

На основе полученных результатов были рассчитаны распределения скоростей передачи данных, определены их характеристики. Проведенные исследования показали, что распределение скоростей практически не зависит от типа передаваемых данных, от количества пользователей сети, и сглаживается с увеличением уровня агрегации. Внешний вид кривых плотности распределения (рисунок 1) свидетельствует о том, что распределение скоростей является

несимметричным и смещено в область меньших значений скорости. Однако остаточная часть данных, приходящаяся на область больших значений, является статистически значимой и не может быть исключена из эксперимента. Это позволяет делать вывод, что полученное распределение имеет «тяжелый хвост».

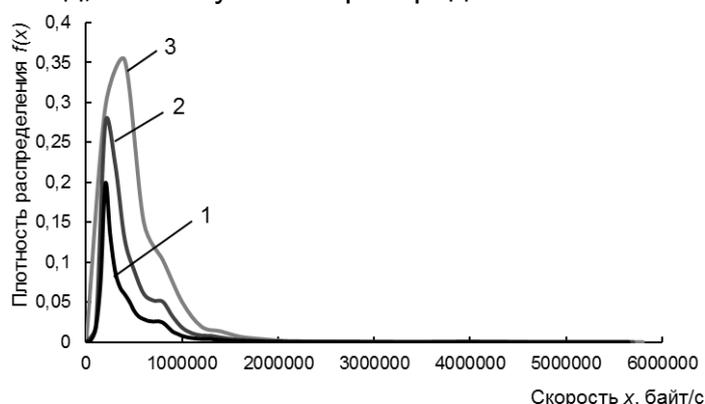


Рисунок 1 – Функция плотности распределения скоростей передачи данных при уровне агрегации: 1 – 50 000, 2 – 100 000, 3 – 200 000

При сравнении распределений экспериментальных данных с результатами расчета с применением математических моделей на основе распределения Пуассона было выявлено, что данные модели характеризуются низкой адекватностью (коэффициент детерминации $R^2 \leq 0,1$), а, следовательно, не могут быть использованы для описания трафика. В свою очередь, модели распределений с «тяжелыми хвостами» показали высокую надежность, при этом наилучшее математическое описание было получено при использовании модели Парето (коэффициент детерминации $R^2 = 0,96$).

Расчет автокорреляционной функции показал наличие долговременной зависимости между пакетами данных, передаваемыми в различные промежутки времени. Это позволяет сделать вывод о самоподобии сетевого трафика.

Данный вывод также подтверждается значениями показателя Херста, рассчитанными для различных видов трафика, а также для потока данных комбинированной структуры. Считается, что трафик обладает свойством самоподобия при значениях параметра Херста $0,5 \leq H \leq 1$. Согласно расчетам, для любых видов трафика значение H составляет не менее 0,9, что свидетельствует о самоподобности полученного трафика.

Таким образом, можно сделать вывод о том, что независимо от вида передаваемых данных сетевой трафик представляет собой самоподобный процесс. При этом наилучшее математическое описание достигается при использовании модели Парето.

ЛИТЕРАТУРА

1. Черниговский, А.В. Основные модели сетевого трафика / А.В. Черниговский, М.В. Кривов // Вестник АНГТУ. – 2017. – № 11. – С. 137-143.