



DOI: <https://doi.org/10.15688/NBIT.jvolsu.2022.2.2>

УДК 004.7

ББК 32.972.5



ОЦЕНКА ТОЧНОСТИ МЕТОДОВ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ НАГРУЗКИ КОМПЬЮТЕРНОЙ СЕТИ

Людмила Константиновна Гомазкова

Аспирант, кафедра телекоммуникационных систем,
Волгоградский государственный университет
gomazkova.lk@volsu.ru
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

Иван Дмитриевич Серёженко

Студент, кафедра телекоммуникационных систем,
Волгоградский государственный университет
Itsб-201_884313@volsu.ru
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

Александр Игоревич Трофимов

Студент, кафедра телекоммуникационных систем,
Волгоградский государственный университет
trofimov.ai@volsu.ru
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

Аннотация. В статье представлен обзор и краткое описание некоторых методов краткосрочного прогнозирования. Также проведена оценка точности полученных с помощью рассмотренных методов прогнозов количества пакетов в компьютерной сети.

Ключевые слова: сетевой трафик, нагрузка сети, краткосрочное прогнозирование, экспоненциальное сглаживание, сглаживание по Спенсеру.

В настоящее время наблюдается рост числа устройств с возможностью подключения к компьютерным сетям. В связи с этим также растет количество пользователей. Поэтому все более важным становится анализ сетевого трафика, а также его прогнозирование, так как благодаря результатам анализа и прогнозирования сетевого потока можно распределить нагрузку на сетевую инфраструктуру с целью более эффективного использования предоставляемых ею ресурсов и более рациональной передачи больших объемов дан-

ных по сети. Также результаты, полученные в процессе прогнозирования и анализа сетевого трафика, позволяют более рационально планировать нагрузку на сетевую инфраструктуру в целом, разрабатывать и устанавливать стратегии развития компьютерных сетей, а также реализовывать системы адаптивного управления трафиком.

Краткосрочное прогнозирование можно соотнести с одним из существующих подходов к управлению сетевым трафиком, который основан на использовании оперативных

методов (online-методов [2]) управления сетью. Так, в ситуациях, где должна быть предусмотрена адаптация планов маршрутизации к изменяющимся условиям сети, методов краткосрочного прогнозирования будет вполне достаточно для получения необходимых результатов.

Краткосрочным прогнозом считается прогноз, для которого период упреждения (то есть временной интервал, для которого разрабатывается прогноз) составляет от нескольких секунд до десятков минут [3]. При этом для периода основания прогноза, то есть промежутка времени, на основе информации о котором строится прогноз, подобных ограничений нет.

В настоящее время характер пользовательской нагрузки на компьютерные сети является непредсказуемым, при этом степень непредсказуемости имеет тенденцию к возрастанию. Также существуют условия постоянного изменения архитектуры компьютерных сетей из-за увеличения числа беспроводных сетей по сравнению с кабельной инфраструктурой. Поэтому наиболее подходящими и перспективными методами для прогнозирования поведения сетевого трафика видится категория именно online-методов, так как данная группа позволяет с высокой скоростью реагировать на происходящие изменения в архитектуре сети и на рост или спад пользовательской нагрузки на компьютерную сеть.

Все множество online-методов, или методов краткосрочного прогнозирования, которые можно отнести к категории научно-исследовательских методов, можно разделить на две крупных подкатегории. Первой из них является подкатегория методов искусственного интеллекта. Применение таких методов имеет большие преимущества, но в то же время эти методы отличаются высокой сложностью ре-

ализации. Вторая подкатегория включает в себя статистические методы. В отличие от методов искусственного интеллекта, статистические методы более просты и обладают преимуществом достаточно точного предсказания суточного графика нагрузки сети при условии анализа сети в обычный будний день. Таким образом, основным недостатком статистических методов является невозможность их применения для анализа нагрузки сети в праздничные или другие выходные дни в связи с отсутствием гибкости их структуры.

Метод экспоненциального сглаживания относится к группе статистических методов краткосрочного прогнозирования. При использовании данного метода прогнозируемое значение складывается из двух составляющих, первой из которых является взвешенные средние величины текущего периода, а второй – сглаженные значения предшествующего периода.

В случае математической записи метод экспоненциального сглаживания представляет собой простую экспоненциальную модель для временных рядов с устойчивым трендом и малыми (независимыми) периодическими колебаниями. Однако для многих временных рядов наблюдается очевидная картина периодичности и случайности, поэтому простая экспоненциальная модель может расширяться включением в нее двух последних компонент [1].

Таким образом, существует три модели метода экспоненциального сглаживания [3]:

1. Простое экспоненциальное сглаживание (модель N-N).
2. Экспоненциальное сглаживание с аддитивным трендом (модель A-N).
3. Экспоненциальное сглаживание с затухающим аддитивным трендом (модель DA-N).

Соответствующие выражения приведены в таблице 1.

Таблица 1

Модели экспоненциального сглаживания

Модель N-N	Модель A-N	Модель DA-N
$S_t = aX_t + (1 - a)S_{t-1}$	$S_t = aX_t + (1 - a)(S_{t-1} + T_{t-1})$	$S_t = aX_t + (1 - a)(S_{t-1} + \varphi T_{t-1})$
$\widehat{X}_t(m) = S_t$	$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}$	$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)\varphi T_{t-1}$
	$\widehat{X}_t(m) = S_t + mT_t$	$\widehat{X}_t(m) = S_t + \sum_{i=1}^m \varphi^i T_t$

Широко известным методом выделения полиномиального тренда временного ряда является использование сглаживающих формул Спенсера [4].

Для задачи краткосрочного прогнозирования удобно использовать процедуры сглаживания с помощью кубического полинома по 5 и 7 точкам (элементам), относящиеся к классу взвешенного скользящего среднего. Оригинальные формулы Спенсера служат для выделения тренда в центральном элементе последовательности, то есть для третьего элемента (точки) при использовании формулы для пяти точек и для четвертого – для семи точек (элементов).

Формулы (1) и (2) представляют собой формулы Спенсера для пяти и семи элементов соответственно [2]:

$$x_k = \frac{1}{35} (-28x_{k-5} + 77x_{k-4} - 28x_{k-3} - 98x_{k-2} + 112x_{k-1}) \quad (1)$$

$$x_k = \frac{1}{21} (-12x_{k-7} + 18x_{k-6} + 12x_{k-5} - 9x_{k-4} - 24x_{k-3} - 12x_{k-2} + 48x_{k-1}) \quad (2)$$

где x_k – прогнозируемый элемент временного ряда; $x_{k-1}, x_{k-2}, x_{k-3}, x_{k-4}, x_{k-5}, x_{k-6}, x_{k-7}$ – элементы временного ряда, предшествующие прогнозируемому.

В качестве исходных данных использовался временной ряд (массив данных), элементы которого представляли собой количество пакетов в сети в секунду времени. Временной ряд был получен в результате обработки дампа сетевого трафика, собранного с помощью программы-анализатора компьютерных сетей «Wireshark» (рис. 1) [5].

Затем при помощи описанных выше методов краткосрочного прогнозирования на основе данного массива были получены еще три массива данных, содержащие уже сглаженные, или предсказанные, значения.

На основе полученных данных были построены следующие графики, где по оси ординат отображалось количество пакетов в секунду, а по оси абсцисс – секунды (рис. 2–4).

Также для более наглядного восприятия был построен сводный график полученных в ходе исследования данных (рис. 5).

На основании полученных графиков можно сделать предположение, что наиболее точным методом краткосрочного прогнозирования среди рассмотренных является метод простого экспоненциального сглаживания.

Для проверки была найдена средняя ошибка прогнозирования путем суммирования ошибки предсказания для каждого элемента исходного массива и последующего деления на общее количество элементов.

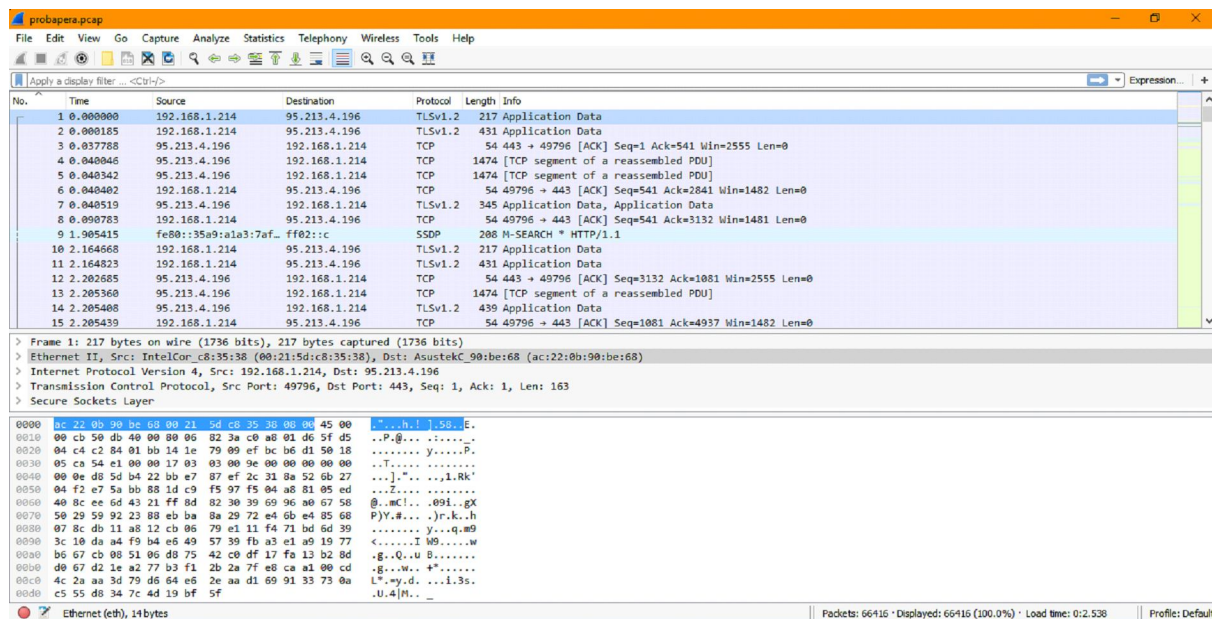


Рис. 1. Собранный дамп трафика

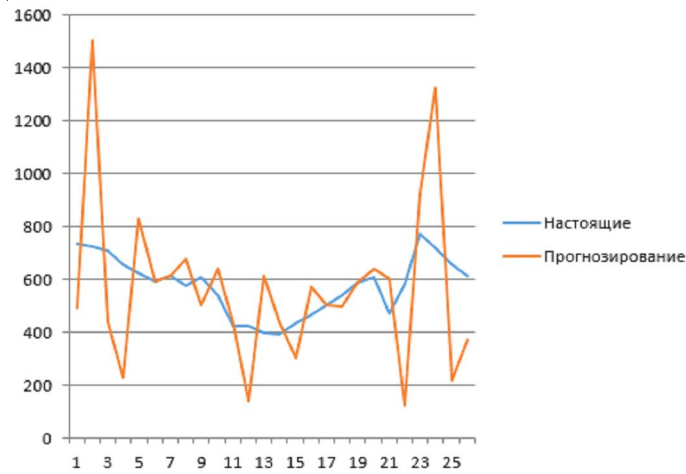


Рис. 2. Графики исходных данных и данных, полученных с помощью формулы Спенсера для 5 элементов

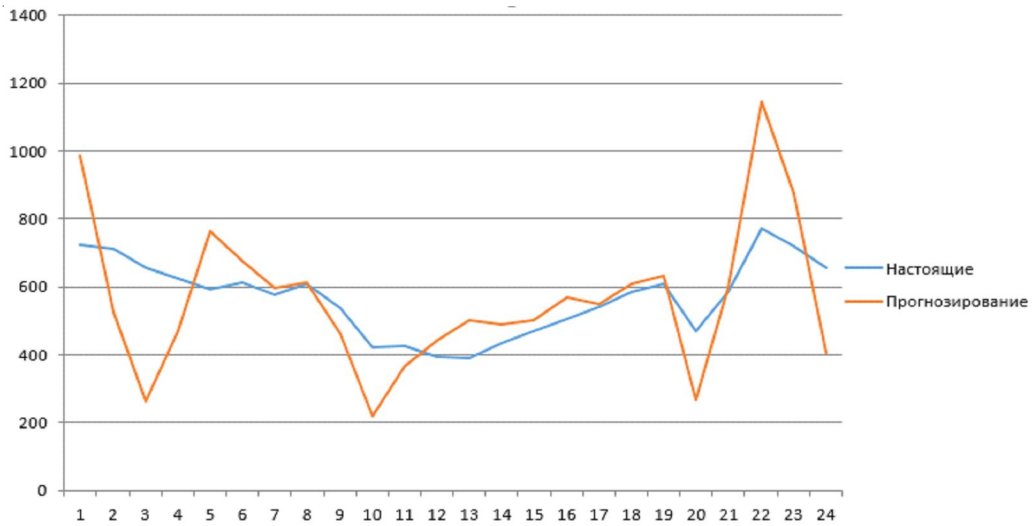


Рис. 3. Графики исходных данных и данных, полученных с помощью формулы Спенсера для 7 элементов

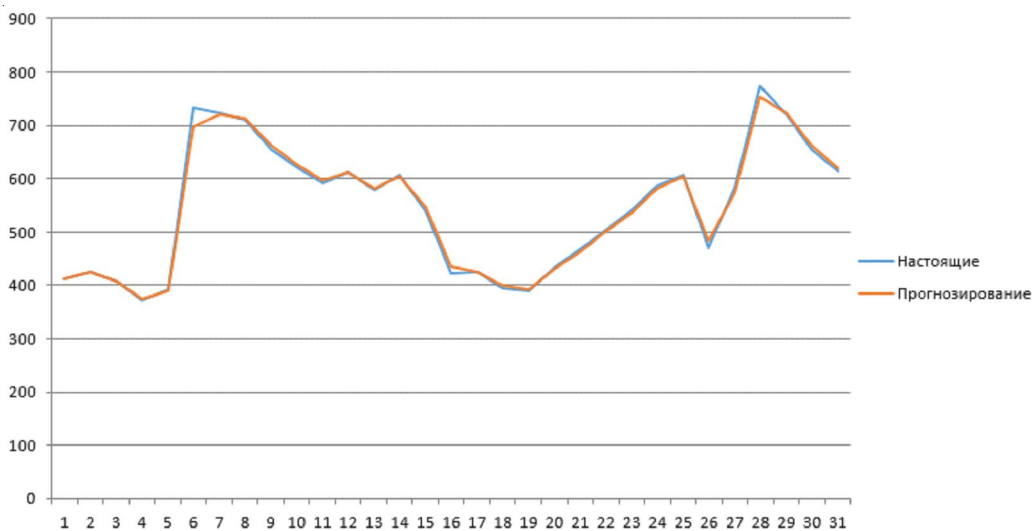


Рис. 4. Графики исходных данных и данных, полученных с помощью метода простого экспоненциального сглаживания

Полученные значения приведены в таблице 2.

Как видно из таблицы, сделанное предположение оказалось верным: наиболее высокая точность прогнозирования наблюдается у временного ряда, полученного с помощью метода простого экспоненциального сглаживания, средняя ошибка прогноза при использовании которого составляет всего 5,5 пакетов/с. При этом менее точным является метод сглаживания с помощью формулы Спенсера для 7 точек (154 пакета/с) и наименее точным – метод сглаживания с помощью формулы Спенсера для 5 точек (203,4 пакета/с).

Таким образом, среди рассмотренных методов краткосрочного прогнозирования была проведена оценка точности полученных с их помощью прогнозов, на основе полученных значений средней ошибки предсказания был выявлен наиболее точный метод краткосрочного прогнозирования сетевого трафика – метод простого экспоненциального сглаживания.

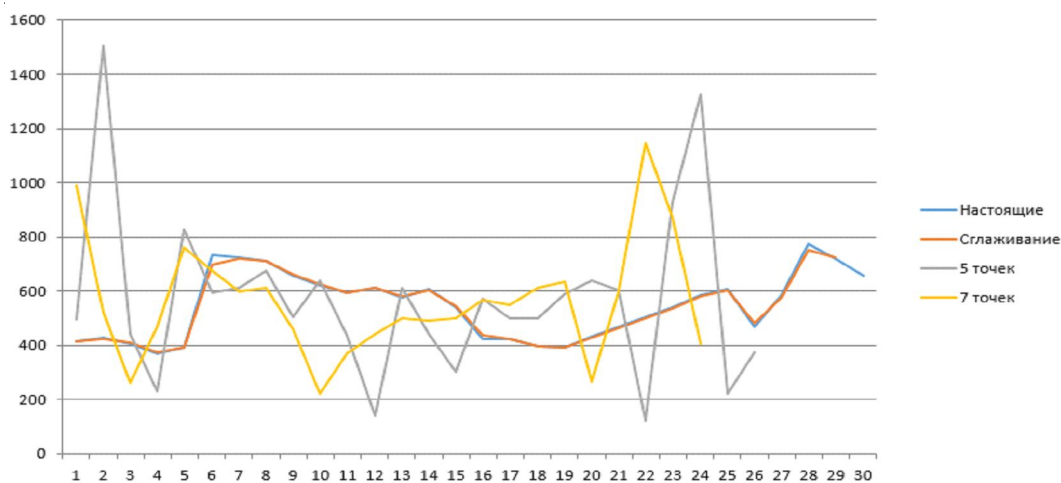


Рис. 5. Графики исходных данных и данных, полученных с помощью методов краткосрочного прогнозирования

Таблица 2

Точность краткосрочного прогнозирования сетевого трафика при помощи различных методов

Метод краткосрочного прогнозирования	Величина ошибки, пакеты/с
Метод сглаживания с помощью формулы Спенсера для 5 точек	203,4
Метод сглаживания с помощью формулы Спенсера для 7 точек	154,0
Метод простого экспоненциального сглаживания	5,5

СПИСОК ЛИТЕРАТУРЫ

1. Анализ временных рядов // StatSoft : Электронный учебник по статистике. – Режим доступа: <http://statsoft.ru/home/textbook/modules/sttimser.html/> (дата обращения: 28.03.2022).
2. Дорт-Гольц, А. А. Анализ эффективности методов краткосрочного прогнозирования сетевого трафика / А. А. Дорт-Гольц // Информационные технологии и телекоммуникации. – 2014. – № 2. – С. 12–19.
3. Дорт-Гольц, А. А. Экспериментальная оценка методов краткосрочного прогнозирования для задачи предсказания интенсивности сетевого трафика / А. А. Дорт-Гольц // Информационные технологии и телекоммуникации. – 2014. – № 3. – С. 97–110.
4. Kendall, M. G. The advanced theory of statistics. Vols. II and III / M. G. Kendall, A. Stuart. – Charles Griffin Publishers, 1961.
5. Wireshark. – 2022. – Mode of access: <https://www.wireshark.org/> (date of access: 28.03.2022).

REFERENCES

1. Analiz vremennykh riadov [Time Series Analysis]. StatSoft: Elektronnyy uchebnyy po

statistike [StatSoft. Electronic Statistical Textbook]. URL: <http://statsoft.ru/home/textbook/modules/sttimser.html> (accessed 28 March 2022).

2. Dort-Golts A.A. Analiz effektivnosti metodov kratkosrochnogo prognozirovaniya setevogo trafika [Analysis of the Effectiveness of Methods for Short-Term Forecasting of Network Traffic]. *Informatsyonnye tekhnologii i telekommunikatsyi* [Information Technologies and Telecommunications], 2014, no. 2, pp. 12-19.

3. Dort-Golts A.A. Eksperimentalnaya otsenka metodov kratkosrochnogo prognozirovaniya dlya

zadachi predskazaniya intensivnosti setevogo trafika [Experimental Evaluation of Short-Term Forecasting Methods for the Problem of Predicting the Intensity of Network Traffic]. *Informatsyonnye tekhnologii i telekommunikatsyi* [Information Technologies and Telecommunications], 2014, no. 3, pp. 97-110.

4. Kendall M.G., Stuart A. *The Advanced Theory of Statistics. Vols. 2 and 3*. S. 1., Charles Griffin Publishers, 1961.

5. *Wireshark*. URL: <https://www.wireshark.org> (accessed 28 March 2022).

ESTIMATION OF THE COMPUTER NETWORK LOAD SHORT-TERM FORECASTING METHODS ACCURACY

Lyudmila K. Gomazkova

Postgraduate Student, Department of Telecommunication Systems,
Volgograd State University
gomazkova.lk@volsu.ru
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

Ivan D. Seryozhenko

Student, Department of Telecommunication Systems,
Volgograd State University
Itsb-201_884313@volsu.ru
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

Alexander I. Trofimov

Student, Department of Telecommunication Systems,
Volgograd State University
trofimov.ai@volsu.ru
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

Abstract. Nowadays, the nature of the load of the users on computer networks is unpredictable, and the degree of such unpredictability tends to increase. There are also conditions for the constant changes in the architecture of the computer networks due to the increase in the number of the different wireless networks compared to the cable infrastructure in whole. Therefore, the category of the methods of the short-term forecastation seems to be the most suitable and promising methods for predicting the behavior of the network traffic, since this group of methods allows you to do a quickly respond to different ongoing changes in the architecture of the network and to an increase or decrease in user's load on a computer network. The article provides a review and a brief description of some methods of short-term forecasting. The estimation of the accuracy of the prediction of the considered methods is also given in this paper. This estimation is based on a time series (data array), whose elements were the number of packets in the network per second of time. The data array was obtained as a result of processing a network traffic dump, which collected using the Wireshark computer network analyzer program.

Key words: network traffic, network load, short-term forecasting, exponential smoothing, Spencer's smoothing.