

М.Г. Дубровин, И.Н. Глухих

---

ПРИМЕНЕНИЕ МОДЕЛИ ХОЛЬТА – ВИНТЕРСА  
ДЛЯ ПРОГНОЗИРОВАНИЯ РАБОТОСПОСОБНОСТИ СЕРВЕРНЫХ  
СИСТЕМ

---

Анализируется применение модели Хольта – Винтерса для прогнозирования работоспособности серверных систем. Отмечается, что для этой цели необходимо собрать статистическую информацию о параметрах объектов, представляющих собой временные ряды. Описывается специфика временных рядов серверных систем для выбора подходящей модели прогнозирования. Проводится сравнение моделей авторегрессии, нейронных сетей и экспоненциального сглаживания для применения к поставленной математической задаче. Выдвинуто утверждение, что модель Хольта – Винтерса обладает преимуществом при анализе рядов группы серверов, содержащих десятки/сотни параметров. Также производятся экспериментальные исследования для оценки точности модели Хольта – Винтерса по отношению к временным рядам, описывающим динамику изменения параметров серверов. Делается вывод, что модель тройного экспоненциального сглаживания демонстрирует хорошие результаты и может быть использована для решения практических задач.

*Ключевые слова:* прогнозирование временных рядов, модель Хольта – Винтерса, работоспособность сервера.

M.G. Dubrovin, I.N. Glukhikh

---

USING HOLT – WINTERS MODELS FOR FORECASTING  
THE PERFORMANCE OF SERVER SYSTEMS

---

The application of the Holt – Winters model for predicting the performance of server systems is analyzed. It is noted that for this purpose it is necessary to collect statistical information about the parameters of objects representing time series. The specifics of the time series of server systems to select the appropriate forecasting model are described. Autoregression models, neural networks, and exponential smoothing are compared for application to the mathematical problem posed. It has been argued that the Holt – Winters model has the advantage of analyzing the series of a group of servers containing tens/hundreds of parameters. Experimental studies are also carried out to assess the accuracy of the Holt – Winters model with respect to time series describing the dynamics of changes in server parameters. It is concluded that the triple exponential smoothing model shows good results and can be used to solve practical problems.

*Keywords:* time series forecasting, Holt – Winters model, server performance.

*Введение*

Увеличение стоимости серверных систем, существенная зависимость организаций от эффективности работы серверов, повышение требований к бесперебойности делают актуальной задачу обеспечения работоспособного состояния корпоративных серверов. Под сервером понимается программно-аппаратный комплекс, предоставляющий клиентам доступ к определенным ресурсам или службам [11]. Для обеспечения раннего пред-

упреждения специалиста, ответственного за работоспособность серверной инфраструктуры, о возможных нарушениях в функционировании серверов необходимо применять методы анализа состояния объектов, ориентированные как на текущие, так и на прогнозные тенденции изменения параметров серверов.

Для прогнозирования работоспособности серверных систем следует собрать статистическую информацию о работе объектов. Собранные статистические данные о каждом параметре сервера представляют собой временные ряды. В общем виде модель временного ряда выглядит следующим образом:  $X(t) = T + S + E$  для аддитивной модели или  $X(t) = TSE$  для мультипликативной модели временного ряда, где  $t = \overline{1, n}$ ;  $T$  – тренд;  $S$  – сезонная компонента;  $E$  – случайная компонента [3].

Пусть  $X_1, X_2, \dots, X_N$  –  $N$  технических параметров сервера, характеризующих его работоспособность. Тогда состояние сервера  $S(t)$  в момент времени  $t$  описывается перечнем значений его параметров:

$$S(t) = (x_1(t), x_2(t), \dots, x_N(t)). \quad (1)$$

Значения каждого параметра сервера  $x_i, i = \overline{1, N}$ , известные в дискретные моменты времени  $t = \overline{1, T}$ , представляют собой временной ряд  $X_i(t) = (X_i(1), X_i(2), \dots, X_i(T))$ . Для решения задачи прогнозирования параметров сервера необходимо найти функциональную зависимость между временным рядом  $X_i(t)$  и будущим прогнозируемым значением  $\hat{X}_i(T+k)$  [2]:

$$\hat{X}_i(T+k) = F(X_i(T), X_i(T-1), \dots, X_i(T-p)), \quad (2)$$

где  $k$  – интервал прогноза;  $p$  – глубина погружения временного ряда.

#### *Материалы и методы*

На данный момент существует более двадцати различных моделей и соответствующих им методов прогнозирования временных рядов [4]. Выбор модели должен базироваться на особенностях анализируемых временных рядов и специфике решаемой задачи.

Для лица, ответственного за работоспособность сервера, больший интерес представляет краткосрочный прогноз (минуты или часы) для оперативного детектирования поведения рядов, не соответствующих прошлой истории. Долгосрочные прогнозы (недели или месяцы) представляют меньший интерес, поскольку можно ожидать, что временные ряды будут развиваться в динамичной среде.

Сервер состоит из множества подсистем, параметры каждой из которых необходимо фиксировать и прогнозировать для получения полноценной информации о работоспособности сервера.

Временные ряды, характеризующие параметры сервера, могут быть описаны изменяющимися трендами при планомерном увеличении или уменьшении потребностей в ресурсах сервера. Также возможно присутствие сезонной составляющей при регулярном и повторяющемся обращении к ресурсам сервера. Таким образом, временные ряды серверных систем не являются стационарными, а модель должна учитывать тренд и сезонную составляющую ряда.

Распространенными моделями прогнозирования временных рядов являются: авторегрессионные модели; модели экспоненциального сглаживания; нейронные сети [Там же].

Использование модификации авторегрессионной модели, учитывающей сезонную составляющую ряда (SARIMA), показывает хорошие результаты в прогнозировании нагрузки серверных систем [7]. Согласно методологии Бокса – Дженкинса [5], если исходный ряд не является стационарным, то его необходимо привести к стационарному виду для использования данной модели. Для распределенной ИТ-инфраструктуры предприятия, состоящей из нескольких серверов, для каждого из которых необходимо анализировать десятки/сотни временных рядов, использование авторегрессионных моделей достаточно трудозатратно.

Различные модели нейронных сетей весьма распространены для прогнозирования компьютерных систем [8] и позволяют проанализировать нелинейные закономерности в наборах данных. С другой стороны, использование нейронных сетей также требует дополнительной предобработки данных, включающей нормализацию значений рядов и приведение к стационарному виду. К тому же нейронные сети работают по принципу черного ящика, что затрудняет интерпретацию модели.

Модель Хольта – Винтерса относится к классу моделей экспоненциального сглаживания. Модель успешно применяется в разных областях, в ИТ-системах находятся примеры использования для прогнозирования сетевого трафика и дальнейшего обнаружения аномалий [6; 9]. Данная модель позволяет распознавать тренд и сезонную компоненту и требует меньшей подготовки данных, чем рассмотренные выше модели. Мультипликативная модель описывается следующей системой уравнений [10]:

$$R_t = \alpha \frac{Y_t}{S_{t-L}} + (1 - \alpha)(R_{t-1} + T_{t-1}); \tag{3}$$

$$T_t = \beta(R_t - R_{t-1}) + (1 - \beta)T_{t-1}; \tag{4}$$

$$S_t = \gamma \frac{Y_t}{S_t} + (1 - \gamma)S_{t-L}; \tag{5}$$

$$Y_{t+m} = (S_t + mT_t)S_{t-L+m}, \tag{6}$$

где  $R_t$  – экспоненциально-сглаженный ряд;  $T_t$  – значение тренда;  $S_t$  – сезонная составляющая ряда;  $Y_{t+m}$  – функция прогноза на  $m$  шагов;  $\alpha, \beta, \gamma$  – коэффициенты сглаживания ряда, тренда и сезонности;  $L$  – период сезонности.

Для реализации модели Хольта – Винтерса и решения системы уравнений необходимо определить период сезонности  $L$  и оптимальные значения параметров  $\alpha, \beta, \gamma$ . Следует провести экспериментальные исследования для анализа применения данной модели. Оценка точности модели осуществляется с помощью средней абсолютной ошибки в процентах (MAPE):

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|Z(t) - \hat{Z}(t)|}{Z(t)} \cdot 100. \tag{7}$$

#### Экспериментальные исследования

Для проведения экспериментальных исследований была произведена подготовка необходимых данных, представленных в виде временных рядов. В течение недели с дискретностью в 10 минут с двух корпоративных серверов производственного предприятия

в Тюмени фиксировались значения трех параметров. Достаточность данных параметров для оценки работоспособности серверов в рамках данной статьи не рассматривается. Зафиксированы значения следующих параметров:

1. CPU load (загрузка центрального процессора).
2. Memory usage (использование физической памяти).
3. Outgoing network traffic (исходящий сетевой трафик).

Все операции с временными рядами и их графиками производились в специализированном программном обеспечении Statistica. Рассмотрим графики временных рядов использования физической памяти (рис. 1) и загрузки центрального процессора (рис. 2) первого сервера.

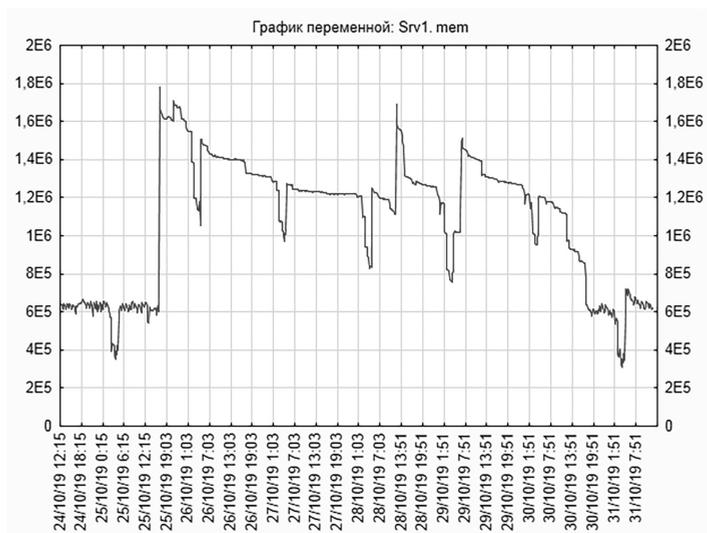


Рис. 1. Сервер 1. Физическая память

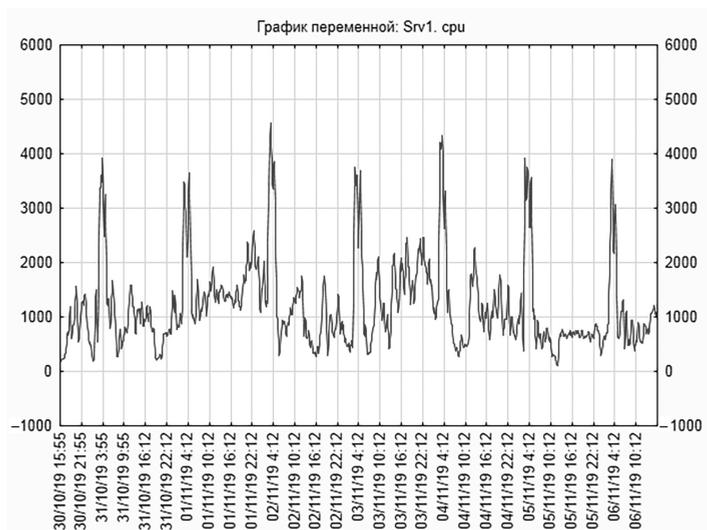


Рис. 2. Сервер 1. Загрузка процессора

В связи с присутствием шумов для лучшего отображения динамики рядов временные ряды предварительно сглажены скользящим средним с шагом в две единицы. Визуальный анализ показывает, что данные графики отражают присутствие тренда и мультипликативной сезонности периодом в сутки (144 измерения).

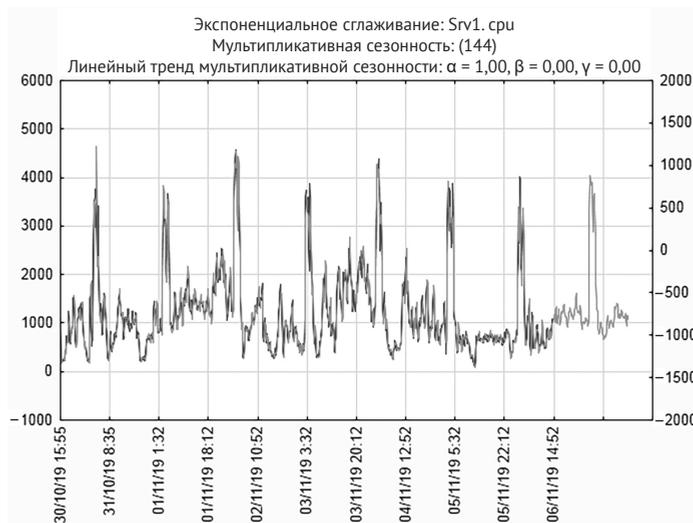
Для каждого исходного ряда построена мультипликативная модель Хольта – Винтерса с линейным трендом. Лаг сезонной компоненты  $L$  выбран в качестве 144 измерений. Выбор коэффициентов  $\alpha$ ,  $\beta$ ,  $\gamma$  осуществлялся с помощью встроенного в Statistica алгоритма автоматического поиска.

Результаты экспериментов на обучающей выборке представлены в таблице. Среднее значение оценки MAPE составляет 10,2%. Считается, что модель показывает высокую точность при MAPE меньше 10% и хорошую – при MAPE от 10 до 20% [1]. Таким образом, можно сделать заключение, что модель Хольта – Винтерса демонстрирует достаточно хорошие результаты для прогнозирования параметров серверов. Для сравнения: эксперименты по анализу модели SARIMA для прогнозирования часовой загрузки двух серверов показали значения точности в 11,46 и 11,67% соответственно [7]. В ходе эксперимента также выявлено, что модель показывает менее точные результаты на рядах с повышенным уровнем шума.

**Результаты экспериментов**

№ п/п	Сервер	Параметр	$L$	$\alpha$	$\beta$	$\gamma$	MAPE, %
1	1	CPU load	144	1	0	0	14,8
2	1	Memory usage	144	0,9	0	0	2,3
3	1	Outgoing network traffic	144	0,4	0,1	0	11,3
4	2	CPU load	144	0,5	0	0	13,7
5	2	Memory usage	144	0,9	0,1	0	6,4
6	2	Outgoing network traffic	144	0,3	0,1	0	12,7

Далее продемонстрирован прогноз загрузки центрального процессора первого сервера на сутки вперед с помощью модели Хольта – Винтерса (рис. 3).



**Рис. 3.** Прогноз загрузки CPU

Темной линией на графике обозначен временной ряд статистических данных. Светлой линией отображается временной ряд прогнозных значений. На графике видно, что при формировании прогнозных значений учитывается тренд-сезонная динамика ряда.

#### Выводы

В процессе сравнительного анализа моделей прогнозирования сделано следующее заключение. Для использования модели Хольта – Винтерса не требуется трудозатратных операций по предварительной обработке значений временных рядов, что является преимуществом при анализе распределенной ИТ-инфраструктуры предприятия, состоящей из нескольких серверов, для каждого из которых необходимо анализировать десятки/сотни временных рядов.

Проведены экспериментальные исследования по оценке использования модели Хольта – Винтерса для прогнозирования параметров серверных систем. Полученные значения оценки точности модели MAPE лежат в диапазоне от 2,3 до 14,8%, среднее значение MAPE составляет 10,2%. Можно сделать заключение, что модель Хольта – Винтерса показывает достаточно хорошие результаты и может быть использована для решения практических задач.

#### Литература

1. Васильев А.А. Критерии селекции моделей прогноза (обзор) // Вестник Тверского государственного университета. 2012. № 2. С. 133–148.
2. Дубровин М.Г. Обобщенный метод проактивного мониторинга сервера базы данных // Передовые инновационные разработки. Перспективы и опыт использования, проблемы внедрения в производство: сборник статей. Казань, 2019. С. 36–38.
3. Садовникова Н.А., Шмойлова Р.А. Анализ временных рядов и прогнозирование. М., 2016. 152 с.
4. Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобия: дис. ... канд. техн. наук. М., 2012.
5. Box G.E.P., Jenkins G.M. Time Series Analysis: Forecasting and Control. San Francisco, 1970.
6. Brutlag J.D. Aberrant Behavior Detection in Time Series for Network Monitoring // LISA. 2000. Vol. 14. P. 139–146.
7. Debusschere V. et al. Hourly Server Workload Forecasting Up to 168 Hours Ahead Using Seasonal ARIMA Model // 2012 IEEE International Conference on Industrial Technology (ICIT). [S. l.]: IEEE, 2012. P. 1127–1131.
8. Debusschere V. et al. Neural Networks for Web Server Workload Forecasting // 2013 IEEE International Conference on Industrial Technology (ICIT). [S. l.]: IEEE, 2013. P. 1152–1156.
9. Jašek R., Szmit A., Szmit M. Usage of Modern Exponential-Smoothing Models in Network Traffic Modelling // Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems. Heidelberg: Springer, 2013. P. 435–444.
10. Kalekar P.S. Time Series Forecasting Using Holt – Winters Exponential Smoothing // Kanwal Rekhi School of Information Technology. 2004. T. 4329008, № 13.
11. Oluwatosin H.S. Client-Server Model // IOSR Journal of Computer Engineering. 2014. Vol. 16, № 1.

## Literatura

1. *Vasil'ev A.A.* Kriterii seleksii modelej prognoza (obzor) // Vestnik Tverskogo gosudarstvennogo universiteta. 2012. № 2. S. 133–148.
2. *Dubrovin M.G.* Obobshchennyj metod proaktivnogo monitoringa servera bazy dannykh // Peredovye innovatsionnye razrabotki. Perspektivy i opyt ispol'zovaniya, problemy vnedreniya v proizvodstvo: sbornik statej. Kazan', 2019. S. 36–38.
3. *Sadovnikova N.A., Shmojlova R.A.* Analiz vremennykh ryadov i prognozirovanie. M., 2016. 152 s.
4. *Chuchueva I.A.* Model' prognozirovaniya vremennykh ryadov po vyborke maksimal'nogo podobiya: dis. ... kand. tekhn. nauk. M., 2012.
5. *Box G.E.P., Jenkins G.M.* Time Series Analysis: Forecasting and Control. San Francisco, 1970.
6. *Brutlag J.D.* Aberrant Behavior Detection in Time Series for Network Monitoring // LISA. 2000. Vol. 14. P. 139–146.
7. *Debusschere V. et al.* Hourly Server Workload Forecasting Up to 168 Hours Ahead Using Seasonal ARIMA Model // 2012 IEEE International Conference on Industrial Technology (ICIT). [S. l.]: IEEE, 2012. P. 1127–1131.
8. *Debusschere V. et al.* Neural Networks for Web Server Workload Forecasting // 2013 IEEE International Conference on Industrial Technology (ICIT). [S. l.]: IEEE, 2013. P. 1152–1156.
9. *Jašek R., Szmit A., Szmit M.* Usage of Modern Exponential-Smoothing Models in Network Traffic Modelling // Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems. Heidelberg: Springer, 2013. P. 435–444.
10. *Kalekar P.S.* Time Series Forecasting Using Holt – Winters Exponential Smoothing // Kanwal Rekhi School of Information Technology. 2004. T. 4329008, № 13.
11. *Oluwatosin H.S.* Client-Server Model // IOSR Journal of Computer Engineering. 2014. Vol. 16, № 1.