

**ПРОГНОЗИРОВАНИЕ МУЛЬТИСЕЗОННЫХ НАГРУЗОЧНЫХ ПРОЦЕССОВ
В ЭЛАСТИЧНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ**

И. Г. МАРТЫНЧУК

*Университет ИТМО, Санкт-Петербург, Россия
mt4.ilja@gmail.com*

Аннотация. Оценивается корректность применения метода мультисезонной сезонно-трендовой декомпозиции на основе локально взвешенного сглаживания диаграммы рассеяния (MSTL) для задач прогнозирования мультисезонных нагрузочных процессов в эластичных системах. Выполнен сравнительный анализ производительности и точности метода MSTL и сезонной интегрированной модели авторегрессионного скользящего среднего (SARIMA). Приведены результаты экспериментов, подтверждающие трудность построения модели SARIMA на данных с высокой степенью дискретизации и значениями периодов, превышающими классические сезонности, такие как 7, 12, 52. При построении модели SARIMA были наложены временные ограничения по подбору параметров вследствие высокого потребления памяти, что приводило к снижению точности прогноза и ограничению возможности построения модели на основе более высоких показателей сезонности. Метод MSTL демонстрирует преимущество по сравнению с моделью SARIMA по времени выполнения прогноза и потреблению памяти, однако на небольшом наборе исходных данных модель SARIMA показывает более высокую точность.

Ключевые слова: эластичные системы, мультисезонные нагрузочные процессы, прогнозирование временных рядов, SARIMA, MSTL

Ссылка для цитирования: Мартычук И. Г. Прогнозирование мультисезонных нагрузочных процессов в эластичных вычислительных системах // Изв. вузов. Приборостроение. 2023. Т. 66, № 11. С. 907—916. DOI: 10.17586/0021-3454-2023-66-11-907-916.

FORECASTING MULTI-SEASONAL LOAD PROCESSES IN ELASTIC COMPUTING SYSTEMS

I. G. Martynchuk

*ITMO University, St. Petersburg, Russia
mt4.ilja@gmail.com*

Abstract. The correctness of using the multi-seasonal season-trend decomposition method based on locally weighted scattergram smoothing for the problems of forecasting multi-seasonal load processes in elastic systems is assessed. A comparative analysis of the performance and accuracy of the above method and the seasonal integrated autoregressive moving average (SARIMA) model is performed. Results of experiments are presented that confirm the difficulty of constructing the SARIMA model based on data with a high degree of discretization and period values exceeding classical seasonality, such as 7, 12, 52. When creating the SARIMA model, time restrictions are imposed on the selection of parameters due to high memory consumption, which lead to a decrease in forecast accuracy and limited the ability to build a model based on higher seasonality indicators. The multi-seasonal season-trend decomposition method demonstrates an advantage over the SARIMA model in terms of forecast execution time and memory consumption, however, with a small set of initial data, the SARIMA model shows higher accuracy.

Keywords: elastic systems, multi-seasonal workload, time series forecasting, SARIMA, MSTL

For citation: Martynchuk I. G. Forecasting multi-seasonal load processes in elastic computing systems. *Journal of Instrument Engineering*. 2023. Vol. 66, N 11. P. 907—916 (in Russian). DOI: 10.17586/0021-3454-2023-66-11-907-916.

Введение. Поток пользовательских запросов к эластичной системе формируется некий нагрузочный процесс, для обработки которого необходимо определенное количество вычислительных ресурсов. Различное поведение пользователей приводит к отличиям в

нагрузочных процессах, что усложняет проектирование и управление системой, так как отсутствует возможность применения единого шаблона нагрузочного процесса [1]. Естественное поведение пользователей, связанное, например, с их суточной активностью, приводит к возрастанию или убыванию количества запросов в единицу времени. Чтобы адаптироваться к таким изменениям в нагрузочном процессе, система должна обладать свойством масштабируемости, т. е. возможностью добавления или исключения вычислительных мощностей в соответствии с изменениями потока пользовательских запросов [2].

В настоящее время широко применяются два способа масштабирования: вертикальное и горизонтальное [3]. Вертикальное масштабирование основано на наращивании вычислительной мощности конкретного узла системы. Горизонтальное масштабирование заключается в наращивании числа вычислительных узлов при повышении нагрузки. Процесс масштабирования может быть частично или полностью реализован в автоматическом режиме. Такой процесс называется автоматическим масштабированием. Система, обладающая свойством автоматической горизонтальной масштабируемости, называется эластичной системой. В современных эластичных системах с открытым исходным кодом применяется простой метод автоматического масштабирования, основанный на использовании мгновенных данных по нагрузке [4]. Главным недостатком такого подхода является отсутствие возможности заблаговременного реагирования на изменения нагрузочного процесса.

Таким образом, важной задачей в современных эластичных системах является прогнозирование нагрузочных процессов с целью сохранения уровня качества предоставляемых услуг и снижения совокупной стоимости владения вычислительной инфраструктурой.

Обзор исследований. Постановка задачи. Суточная активность реальных пользователей по-разному влияет на характеристики нагрузочных процессов. Например, если обратиться к графикам суммарного транзитного трафика через узлы MSK-IX за неделю и месяц, то можно отметить ярко выраженную суточную периодичность нагрузочного процесса [5]. В теории временных рядов такая периодичность называется сезонностью. Для MSK-IX нагрузочный процесс имеет одну сезонную суточную составляющую.

Нагрузочные процессы в транзитных узлах на сетевом уровне отличаются от нагрузочных процессов, характерных для отдельных эластичных систем на прикладном уровне. На рис. 1 изображен нагрузочный процесс, представляющий собой агрегированное по пятиминутным интервалам число запросов (N) к кластеру моделей машинного обучения, выполняющему прогнозы различных финансовых показателей в крупной коммерческой организации. Период наблюдения — один месяц. Видно, что представленный нагрузочный процесс имеет две сезонные составляющие: суточную и недельную. Такой порядок объясняется пятидневной рабочей неделей, характерной для крупных коммерческих организаций.

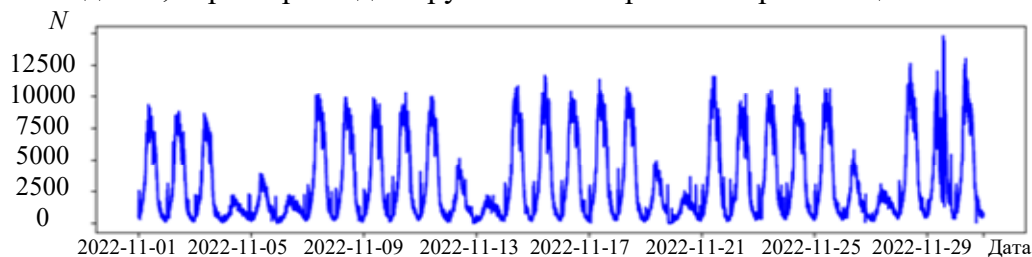


Рис. 1

В открытых источниках существует ряд работ, посвященных прогнозированию нагрузочных процессов в эластичных системах. В работах [6—8] рассматриваются прогнозы, осуществляемые с помощью линейных моделей регрессии, а именно авторегрессии и скользящего среднего (AR, MA, ARMA). В [8] используется также интегрированная модель линейной регрессии (ARIMA), позволяющая осуществить переход от нестационарного временного ряда к стационарному. В [9] применяются скрытые цепи Маркова для прогнозирования загрузки

виртуальных машин и последующего их размещения. Данные подходы не учитывают возможную сезонность в исследуемых нагрузочных процессах.

Для учета сезонности поступления данных используются модели SARIMA и SARIMAX [10]. В [11] показано, что такие модели демонстрируют сравнимые с нейронными сетями показатели точности прогноза. Однако сезонные модели ARIMA используют предположение только об одной сезонной составляющей и исследуемом временном ряде.

В области прогнозирования временных рядов существует ряд моделей и методов для работы с мультисезонностью. В [12] показано применение моделей BATS и TBATS для работы с временными рядами, обладающими несколькими сезонными компонентами. В [13] для тех же целей предложено использовать библиотеку Prophet. Метод MSTL, демонстрирующий уменьшение времени выполнения прогноза по сравнению с моделями BATS и TBATS, а также фреймворком Prophet, предложен в [14].

В открытых источниках не найдено работ, посвященных использованию метода мультисезонной сезонно-трендовой декомпозиции на основе локально взвешенного сглаживания диаграммы рассеяния (MSTL — Multiple Trend Decomposition Using LOWESS) для работы с временными рядами, которые формируются системами мониторинга эластичных систем. В связи с этим представляется актуальным исследование возможности использования метода MSTL и сравнительный анализ его производительности и сезонной интегрированной модели авторегрессионного скользящего среднего (SARIMA — Seasonal Autoregressive Integrated Moving Average) для решения задач исследования и прогнозирования мультисезонных нагрузочных процессов в эластичных вычислительных системах.

Сезонная регрессионная модель случайных процессов. В задачах прогнозирования временных параметров в эластичных системах случайный процесс — это интенсивность поступления пользовательских запросов, а также события, возникающие в системе после поступления очередного запроса. Такие процессы фиксируются с помощью систем мониторинга. На выходе случайный процесс представляет собой временной ряд. Поскольку при обработке случайных процессов используются их реализации в виде временных рядов, понятия случайного процесса и временного ряда будем считать взаимозаменяемыми.

Для описания стационарных случайных процессов применяются линейные модели регрессии. Случайный процесс называется стационарным, если выполняется следующий набор условий [15]:

- математическое ожидание E не зависит от времени; временной ряд не имеет тенденции к росту или убыванию, а остается на одном уровне в течение всего периода наблюдений;
- дисперсия σ не зависит от времени; амплитуды колебаний временного ряда одинаковые вне зависимости от сдвига во времени;
- сила линейной связи между двумя величинами зависит только от расстояния между ними, т.е. автоковариационная и автокорреляционные функции зависят только от величины сдвига.

Модель скользящего среднего порядка q определяет состояние случайного процесса на основе взвешенного среднего q предыдущих значений белого шума:

$$MA(X_t, q) = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q X_{t-q},$$

где X_t — состояние случайного процесса в момент времени t ; q — порядок модели; ε_t — белый шум с нулевым математическим ожиданием; θ_i — весовой коэффициент.

Порядок модели определяется с помощью автокорреляционной функции, которая показывает степень линейной статистической зависимости между исходным рядом и его копией, сдвинутой на заданное число интервалов. Для модели MA(q) автокорреляционная функция затухает экспоненциально при сдвиге, превышающем порядок модели q .

Авторегрессионная модель порядка p определяет состояние случайного процесса на основе линейной комбинации p его предшествующих значений и белого шума:

$$\text{AR}(X_t, p) = \varepsilon_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p},$$

где p — порядок модели, α_i — коэффициенты модели.

Порядок модели определяется функцией частной автокорреляции, позволяющей, как и в случае обычной автокорреляции, оценить характеристику взаимосвязей между предыдущими и текущими значениями временного ряда. Влияние малой задержки при этом устранено, так как частная автокорреляция направлена на исключение косвенной корреляции путем удаления линейной зависимости между сдвинутыми рядами. Для модели $\text{AR}(p)$ частная автокорреляционная функция затухает экспоненциально при сдвиге, превышающим порядок модели p .

Модель авторегрессионного скользящего среднего состоит из комбинации моделей процессов $\text{AR}(p)$ и $\text{MA}(q)$:

$$\text{ARMA}(X_t, p, q) = \varepsilon_t + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}.$$

В случае процесса $\text{ARMA}(p, q)$ автокорреляционная и частная автокорреляционная функции, начиная с некоторого сдвига, представляют собой сумму затухающих экспонент, если ряд стационарен. С помощью величины данного сдвига определяются порядки моделей $\text{AR}(p)$ и $\text{MA}(q)$. Согласно теореме Вольда любой стационарный временной ряд может быть представлен с помощью модели $\text{ARMA}(p, q)$ с любой точностью [16].

Во временных рядах обычно выделяют важные для понятия стационарности компоненты: тренд и сезонность. Трендом называется плавное долгосрочное изменение уровня ряда. Это приводит к изменению математического ожидания временного ряда, что делает ряд нестационарным. Сезонность представляет собой циклические изменения среднего значения ряда с постоянным периодом. Такие изменения могут затрагивать как математическое ожидание, так и дисперсию ряда, делая его также нестационарным.

Интегрированная модель авторегрессионного скользящего среднего ($\text{ARIMA}(p, d, q)$) представляет собой расширение модели $\text{ARMA}(p, q)$ и содержит интегрированную составляющую:

$$\text{ARIMA}(X_t, p, d, q) = \varepsilon_t + \sum_{i=1}^p \alpha_i \Delta^d X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i},$$

где Δ^d — оператор разности временного порядка d .

Данная составляющая позволяет перейти от нестационарного ряда к стационарному путем его дифференцирования d раз. Дифференцирование представляет собой переход от исходного временного ряда к ряду попарных разностей.

Сезонная интегрированная модель авторегрессионного скользящего среднего (SARIMA) является расширением модели ARIMA . Учет сезонности во временном ряду достигается путем добавления набора компонентов авторегрессии и скользящего среднего:

$$\text{SARIMA}(X_t, p, d, q, P, D, Q)_s = \varepsilon_t + \sum_{i=1}^p \alpha_i \Delta^d X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^P \beta_i \Delta^D X_{t-si} + \sum_{i=1}^Q \vartheta_i \varepsilon_{t-si},$$

где P — порядок сезонной авторегрессионной составляющей, D — порядок интегрирования сезонной составляющей, Q — порядок сезонной составляющей скользящего среднего, s — показатель сезонности; β_i , ϑ_i — коэффициенты сезонной составляющей.

Мультисезонная сезонно-трендовая декомпозиция временного ряда.

Локально взвешенное сглаживание диаграммы рассеяния (LOWESS) представляет собой механизм сглаживания на основе взвешенного скользящего среднего и взвешенной линейной или полиномиальной регрессии. Метод состоит из следующих операций.

1) Расчет расстояния между попарно распределенными точками исходного набора данных:

$$r(x_i, x_j) = |x_i - x_j|;$$

для каждой точки x_i определяются m смежных точек x_j , которые находятся ближе всего к x_i ; параметр m называется размахом или размером локального подынтервала.

2) Для каждой из m точек x_j в локальном подынтервале вычисляется масштабированная дистанция, где знаменатель содержит расстояние от точки x_i до наиболее удаленной точки x_j в локальном подынтервале:

$$r^*(x_j) = \frac{r(x_i, x_j)}{\max r(x_i, x_j)}.$$

3) Для каждой точки в локальном подынтервале рассчитывается вес следующим образом:

$$w(x_j) = \left(1 - [r^*(x_j)]^3\right)^3.$$

4) На основе точек в локальном подынтервале и полученных весов выполняется взвешенная линейная регрессия.

Результатом сглаживания в точке x_i является значение регрессии $y = kx_i + b$, где k , b — параметры регрессии.

Сезонно-трендовая декомпозиция временного ряда на основе LOWESS (STL) используется для определения тренда и сезонной составляющей исследуемого временного ряда и состоит из внешнего и внутреннего циклов. Основные операции метода реализуются во внутреннем цикле, отвечающем за сглаживание сезонной и трендовой составляющих временного ряда. Внутренний цикл состоит из следующих операций.

1. Извлечение тренда из временного ряда; на первой итерации цикла этот шаг не выполняется.

2. Сглаживание сезонных подсерий; временной ряд разбивается на сезонные подсерии, каждая из которых сглаживается с помощью метода LOWESS.

3. Низкочастотная фильтрация сглаженных сезонных подсерий; сезонные подсерии суммируются и сглаживаются с помощью фильтров на основе скользящего среднего и LOWESS.

4. Извлечение трендовой составляющей из сезонных подсерий.

5. Извлечение из исходного временного ряда сезонной составляющей, полученной на предыдущих шагах.

6. Получение тренда путем сглаживания полученного на шаге 5 временного ряда с помощью LOWESS.

Внешний цикл отвечает за минимизацию влияния выбросов путем замены обычного метода LOWESS на его более устойчивую к большим отклонениям версию.

Мультисезонная сезонно-трендовая декомпозиция временного ряда на основе LOWESS (MSTL) является расширением STL, позволяющим определять несколько сезонных компонентов во временном ряду. Метод состоит из следующих операций.

1. Извлечение каждой сезонной составляющей с помощью STL; сезонные составляющие сортируются по возрастанию периода для предотвращения возможного поглощения коротких сезонных циклов более длинными.

2. Определение тренда наиболее длинной сезонной составляющей (составляющей, извлеченной на заключительной итерации шага 1).

3. Определение остаточной составляющей путем извлечения тренда из полученного на шаге 2 временного ряда.

Если временной ряд не является сезонным, то для определения тренда используется алгоритм „Supersmoother“, разработанный Д. Фридманом и представляющий собой сглаживание двумерной регрессии на основе локальной линейной регрессии с адаптивной пропускной способностью [14].

Эксперимент. Для исследуемого временного ряда были выбраны данные, полученные системой мониторинга кластера моделей машинного обучения (см. рис. 1), обсуждаемые в постановке задачи. Взят отрезок наблюдения с июня по декабрь 2022 г. Пропуски в данных заполнены ближайшими следующими по порядку существующими отсчетами для корректной работы программных реализаций методов. Шаг наблюдений в исследуемом наборе данных равен 5 мин. Таким образом, дневной цикл равен 288 отсчетам, а недельный — 2016. Данные значения сезонности использовались в сравниваемых моделях.

В качестве модели SARIMA использована реализация *auto_arima* библиотеки *pmdarima* языка Python. Для определения порядка дифференцирования регрессионной составляющей внутри модели использовался тест Квятковского — Филлипса — Шмидта — Шина. Определение порядка сезонности в автоматическом режиме осуществлялось на основе комбинации тестов OCSB и Кановы — Хансена. Показатель сезонности принимался и в качестве входного параметра. Для автоматического определения наилучшей модели использовался информационный критерий Акаике (AIC).

Для метода MSTL использована реализация библиотеки *statsforecast* языка Python, входной параметр — показатель сезонности.

Эксперименты проводились на тестовом стенде с установленным процессором AMD Ryzen 9 5950X и емкостью оперативной памяти 64 Гб. Запуск реализаций осуществлялся в окружении WSL2 под управлением ОС Windows 10.

Оценка качества прогноза выполнялась путем расчета следующих показателей:

— среднеквадратической ошибки

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \hat{X}_i)^2},$$

где n — количество наблюдений в исследуемом наборе данных, X_i — фактическое значение, а \hat{X}_i — спрогнозированное значение временного ряда (набора данных); показатель RMSE совпадает с размерностью исходного набора данных;

— коэффициента детерминации:

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_i - \hat{X}_i)^2}{\sum_{i=1}^n (X_i - \bar{X}_i)^2},$$

где \bar{X}_i — среднее значение временного ряда (набора данных); показатель представляет собой нормированную среднеквадратическую ошибку, чем ближе данный коэффициент к единице, тем выше качество прогноза.

Результаты. Рассчитанные показатели качества прогноза на один день вперед при построении моделей на недельных и месячных данных представлены в табл. 1.

Таблица 1

Показатель	SARIMA(0,1,0)(0,0,0)[288]		MSTL(288,2016)	
	Набор данных		Набор данных	
	недельный	месячный	недельный	месячный
RMSE	474,94404	—	2524,79448	711,03479
R^2	0,97192	—	0,206541	0,93902

Показатели производительности методов по результатам экспериментов представлены в табл. 2. Под временем выполнения понимается суммарное время, затраченное на построение модели и выполнение прогноза на день вперед.

Таблица 2

Метод	Недельный набор данных		Месячный набор данных	
	Время выполнения, с	Потребление памяти, Гб	Время выполнения, с	Потребление памяти, Гб
SARIMA	1080,92 ± 1,2	48,447 ОП 6,4 ± 0,4 swap	—	—
MSTL	6,48 ± 0,03	0,375 ± 0,025	15,23 ± 0,08	0,375 ± 0,025

Результаты прогноза на день вперед представлены на рис. 2: *а, б* — SARIMA и MSTL соответственно — недельные данные для построения модели; *в* — MSTL — заключительная неделя наблюдений для построения модели. В ходе экспериментов подтверждена высокая вычислительная сложность построения модели SARIMA на данных с высокой степенью дискретизации и значениями периодов, превышающими классические сезонности, такие как 7, 12 и 52. Для выбора подходящей модели SARIMA в автоматическом режиме было наложено мягкое ограничение на время построения модели, равное 5 мин. В противном случае ресурсов тестового стенда не хватало. При таких ограничениях алгоритмом по критерию AIC была выбрана модель SARIMA(0,1,0)(0,0,0)[288]. Прогноз выполнялся циклически на один пяти-минутный интервал вперед с одновременным обновлением модели фактическими данными для повышения точности прогноза.

В случае недельного набора данных при прогнозе на день вперед модель SARIMA демонстрирует меньшее значение показателя RMSE и большее значение R^2 , что свидетельствует о меньшей среднеквадратической ошибке прогноза. На месячных данных не удалось осуществить построение модели SARIMA, так как доступных ресурсов на тестовом стенде оказалось недостаточно.

Анализ табл. 2 показывает, что метод MSTL значительно превосходит SARIMA по скорости выполнения прогноза, а также по потреблению оперативной памяти и страниц swap.

Для метода MSTL также определена зависимость показателя R^2 от величины исходного набора данных (количество недель — M) при различных показателях сезонности (рис. 3). По результатам экспериментов можно сделать вывод, что точность прогноза при недельной сезонности выше, чем при мультисезонном варианте. Это объясняется значительными отличиями в нагрузке в выходные дни. Точность прогноза при небольшом количестве данных в исходном наборе приближается к точности прогноза при использовании модели SARIMA.

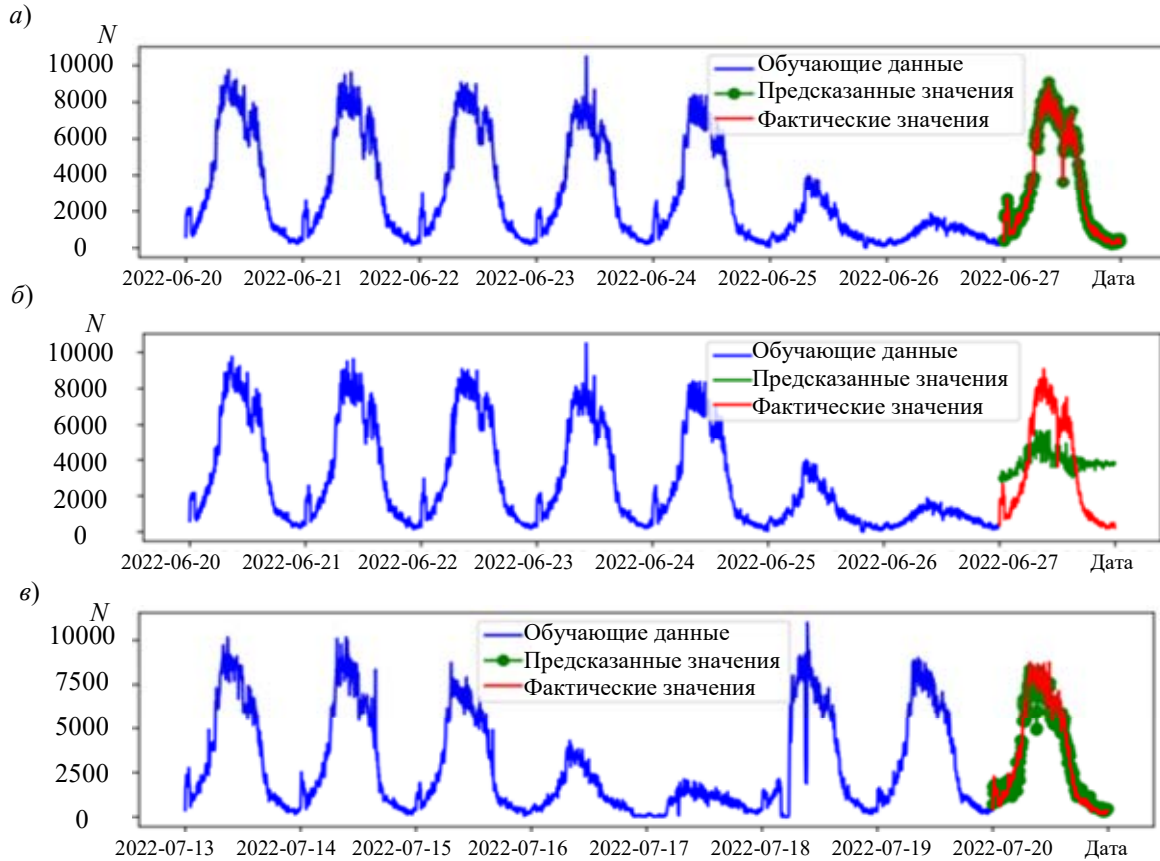


Рис. 2

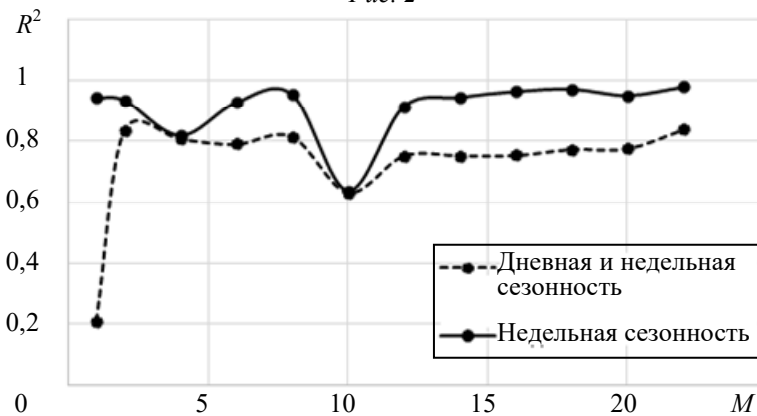


Рис. 3

Основной целью процесса автоматизированного масштабирования является уменьшение затрат на вычислительную инфраструктуру. В контексте эластичных систем данная цель достигается путем снижения количества функционирующих вычислительных узлов в течение некоторого периода наблюдения. Согласно результатам экспериментов для построения модели SARIMA требуется значительное количество вычислительных ресурсов и времени в случае работы с временными рядами высокой степени дискретизации. Для данной модели необходим отдельный вычислительный узел со значительным количеством ресурсов, что снижает эффективность автоматического масштабирования, особенно в условиях частого перестроения моделей.

Заключение. По результатам исследования и прогнозирования мультисезонных нагрузочных процессов, характерных для эластичных вычислительных систем, метод MSTL продемонстрировал преимущество по сравнению с моделью SARIMA по таким показателям, как время выполнения прогноза и потребление памяти.

На небольшом наборе исходных данных модель SARIMA демонстрирует более высокую точность по сравнению с методом MSTL в мультисезонной конфигурации. При использовании же одной сезонной составляющей метод MSTL демонстрирует близкую к модели SARIMA точность на таком же наборе данных. Однако для построения модели SARIMA приходится накладывать ограничения по времени подбора параметров вследствие высокого потребления памяти, что ведет к снижению точности прогноза и ограничивает возможность построения модели на основе более высоких показателей сезонности.

СПИСОК ЛИТЕРАТУРЫ

1. Aliev T. I., Rebezova M. I., Russ A. A. Statistical Methods for Monitoring Travel Agencies in the Settlement System // Automatic Control and Computer Sciences. 2015. Vol. 49, N 6. P. 321—327.
2. Bogatyrev V. A., Bogatyrev S. V., Bogatyrev A. V. Control of Multipath Transmissions in the Nodes of Switching Segments of Reserved Paths // Intern. Conf. on Information, Control, and Communication Technologies (ICCT), IEEE. 2022. P. 1—5.
3. Богатырев В. А., Богатырев С. В. Резервированная передача данных через агрегированные каналы в сети реального времени // Изв. вузов. Приборостроение. 2016. Т. 59, № 9. С. 735—740.
4. Portnoy M. Virtualization Essentials. John Wiley & Sons, 2012. Vol. 19.
5. АО „Центр взаимодействия компьютерных сетей «МСК-IX»“. Суммарный трафик участников, передаваемый через MSK-IX. [Электронный ресурс]: <https://www.msk-ix.ru/traffic/>, 15.07.2023.
6. Roy N., Dubey A., Gokhale A. Efficient autoscaling in the cloud using predictive models for workload forecasting // IEEE 4th Intern. Conf. on Cloud Computing. 2011. P. 500—507.
7. Tirado J. M. et al. Predictive data grouping and placement for cloud-based elastic server infrastructures // 11th IEEE/ACM Intern. Symp. on Cluster, Cloud and Grid Computing. 2011. P. 285—294.
8. Liao S. et al. Adaptive resource prediction in the cloud using linear stacking model // 5th Intern. Conf. on Advanced Cloud And Big Data (CBD). IEEE. 2017. P. 33—38.
9. Melhem S. B. et al. Markov prediction model for host load detection and VM placement in live migration // IEEE Access. 2017. Vol. 6. P. 7190—7205.
10. Yazdani P., Sharifian S. Cloud workload prediction using convnet and stacked lstm // 4th Iranian Conf. on Signal Processing and Intelligent Systems (ICSPIS). IEEE. 2018. P. 83—87.
11. Vagropoulos S. I. et al. Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting // IEEE Intern. Energy Conf. (ENERGYCON). 2016. P. 1—6.
12. Naim I., Mahara T., Idrisi A. R. Effective short-term forecasting for daily time series with complex seasonal patterns // Procedia Computer Science. 2018. Vol. 132. P. 1832—1841.
13. Xie T., Ding J. Forecasting with multiple seasonality // IEEE Intern. Conf. on Big Data (Big Data). 2020. P. 240—245.
14. Bandara K., Hyndman R. J., Bergmeir C. MSTL: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns // arXiv preprint arXiv:2107.13462. 2021.
15. Fryzlewicz P., Van Belleghem S., Von Sachs R. Forecasting non-stationary time series by wavelet process modelling // Annals of the Institute of Statistical Mathematics. 2003. Vol. 55, N 4. P. 737—764.
16. Antoni J., Randall R. B. Unsupervised noise cancellation for vibration signals: part I—evaluation of adaptive algorithms // Mechanical Systems and Signal Processing. 2004. Vol. 18, N 1. P. 89—101.

Сведения об авторе

Илья Геннадьевич Мартынюк

— Университет ИТМО, кафедра вычислительной техники; преподаватель; E-mail: mt4.ilja@gmail.com

Поступила в редакцию 19.06.2023; одобрена после рецензирования 28.06.2023; принята к публикации 27.09.2023.

REFERENCES

1. Aliev T.I., Rebezova M.I., Russ A.A. *Automatic Control and Computer Sciences*, 2015, no. 6(49), pp. 321–327.
2. Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. *2022 International Conference on Information, Control, and Communication Technologies (ICCT)*, 2022, pp. 1–5.
3. Bogatyrev V.A., Bogatyrev S.V. *Journal of Instrument Engineering*, 2016, no. 9(59), pp. 735–740. (in Russ.)
4. Portnoy M. *Virtualization essentials*, John Wiley & Sons, 2012, 336 p.
5. <https://www.msk-ix.ru/traffic/>. (in Russ.)
6. Roy N., Dubey A., Gokhale A. *2011 IEEE 4th International Conference on Cloud Computing*, 2011, pp. 500–507.
7. Tirado J.M. et al. *11th IEEE/ACM Intern. Symp. on Cluster, Cloud and Grid Computing*, 2011, pp. 285–294.
8. Liao S. et al. *5th Intern. Conf. on advanced cloud and big data (CBD)*, 2017, pp. 33–38.
9. Melhem S. B. et al. *IEEE Access*, 2017, vol. 6, pp. 7190–7205.
10. Yazdaniyan P., Sharifian S. *2018 4th Iranian Conf. on Signal Processing and Intelligent Systems (ICSPIS)*, 2018, pp. 83–87.
11. Vagropoulos S.I. et al. *2016 IEEE Intern. Energy Conf. (ENERGYCON)*, 2016, pp. 1–6.
12. Naim I., Mahara T., Idrisi A.R. *Procedia Computer Science*, 2018, vol. 132, pp. 1832–1841.
13. Xie T., Ding J. *2020 IEEE Intern. Conf. on Big Data (Big Data)*, 2020, pp. 240–245.
14. Bandara K., Hyndman R.J., Bergmeir C. *arXiv preprint arXiv:2107.13462*, 2021.
15. Fryzlewicz P., Van Belleghem S., Von Sachs R. *Annals of the Institute of Statistical Mathematics*, 2003, no. 4(55), pp. 737–764.
16. Antoni J., Randall R.B. *Mechanical Systems and Signal Processing*, 2004, no. 1(18), pp. 89–101.

Data on author**Ilya G. Martynchuk**— ITMO University, Department of Computing Technique; Lecturer;
E-mail: mt4.ilja@gmail.com

Received 19.06.2023; approved after reviewing 28.06.2023; accepted for publication 27.09.2023.