

## АВТОМАТИЗИРОВАННОЕ УПРАВЛЕНИЕ СТРУКТУРОЙ ЭЛАСТИЧНЫХ СИСТЕМ И ИХ КОМПОНЕНТАМИ

Мартынчук И.Г. (Университет ИТМО)

Научный руководитель – д.т.н., профессор Алиев Т.И.  
(Университет ИТМО)

Современные подходы к прогнозированию рабочей нагрузки основаны на использовании строго одного метода или модели. Однако нагрузочные процессы в современных системах имеет различный характер. Таким образом, можно поставить под сомнение эффективность использования только одного определенного метода прогнозирования. Кроме того, зачастую сложно оценить, какое количество вычислительных ресурсов необходимо для успешного функционирования системы на старте. В работе предлагается метод автоматизированного управления структурой эластичных систем и их компонентами, который позволяет не только осуществлять прогноз нагрузочных процессов на основе исторических данных, но и выполнять оценку необходимого количества вычислительных ресурсов для функционирования компонентов систем в различных условиях.

**Введение.** В современных информационно-вычислительных системах для своевременного выполнения операций масштабирования недостаточно знать величину текущей рабочей нагрузки, так как такие операции сопровождаются определенными накладными расходами. Прогнозирование нагрузочных процессов и автоматизация процесса масштабирования является важной составляющей в задаче управления компонентами информационно-вычислительных систем. Современные подходы к прогнозированию рабочей нагрузки основаны на использовании строго одного метода или модели. Однако нагрузочные процессы в современных системах отличаются друг от друга и, более того, создают неоднородную нагрузку на компоненты системы. Таким образом, можно поставить под сомнение эффективность использования только одного определенного метода прогнозирования. Кроме того, большинство алгоритмов рассматривают систему как черный ящик и ориентируются только на загруженность вычислительных ресурсов, что не позволяет подобрать наиболее эффективный подход к масштабированию того или иного компонента системы.

**Основная часть.** Целью работы является обеспечение возможности прогнозирования характера влияния неоднородных нагрузочных процессов на характеристики функционирования эластичных систем. Для достижения поставленной цели определен следующий ряд задач:

1. Выявить факторы, влияющие на эластичность информационно-вычислительных систем.
2. Предложить подход к прогнозированию характеристик функционирования эластичных систем на основе математического аппарата теории массового обслуживания.
3. Разработать алгоритм обработки временных рядов для работы с нагрузочными процессами, протекающими в эластичных системах.
4. Разработать метод автоматизированного управления структурой эластичных систем и их компонентами.

В ходе выполнения работы проанализированы подходы к построению и организации информационно-вычислительных систем различного назначения. По результатам анализа выявлены факторы, влияющие на эластичность систем. Стоит отметить, что важным этапом при определении масштабируемости системы является разделение её на составные компоненты, так как на данный момент определение эластичности системы в целом является трудновыполнимой задачей из-за зачастую различной структурно-функциональной организации её компонентов и требований к ним. Например, база данных и веб-сервер

нуждаются в различных подходах к масштабированию, так как к ним предъявляются различные требования с точки зрения обработки данных.

В рамках работы решаются две задачи управления компонентами эластичных систем: это прогнозирование как необходимого количества инфраструктурных ресурсов для первоначального развертывания конкретной системы и характеристик её функционирования, так и непосредственное прогнозирование нагрузочных процессов на систему в период её функционирования. Вторая задача представляет из себя ничто иное как обработка и прогнозирование временных рядов.

Полученные аналитические модели позволяют оперативно оценить характер нагрузочного процесса в конкретный момент времени в условиях нескольких потоков пользовательских запросов. Также стоит отметить, что в случае использования аналитического моделирования и появления сложных зависимостей параметра распределения от времени возникает проблема непосредственного получения функции изменения этого параметра, что затрудняет решение задачи непосредственного прогнозирования нагрузочных процессов.

Для решения данной проблемы было решено провести анализ моделей и численных методов прогнозирования процессов в сфере облачных технологий. По результатам анализа найденных открытых источников выбраны наиболее популярные подходы из области статистического анализа временных рядов, а именно линейные статистические модели. В рамках работы формализован алгоритм обработки временных рядов для работы с нагрузочными процессами в сфере облачных систем, включающий в себя разметку ряда, обработку недостающих данных и выбор статистического метода прогноза.

Однако линейные статистические методы показывают низкую точность при обработке больших наборов накопленных исторических данных. В таком случае предпочтительнее использовать более сложные модели машинного обучения. Наиболее популярным подходом здесь оказалось использование сетей долгой краткосрочной памяти – особой разновидности рекуррентных нейронных сетей, способных к обучению долговременным зависимостям.

В результате разработан метод управления вычислительными ресурсами эластичных систем с помощью комбинации аналитического моделирования и методов статистики и машинного обучения. В предложенном подходе происходит разделение системы на структурные компоненты и выбор типа масштабирования для каждого из них. Кроме того, осуществляется регулярное перестроение прогнозных моделей и проверка их корреляции с исторической нагрузкой на систему для выбора наиболее подходящей модели для рассматриваемой системы. Само автоматическое масштабирования может выполняться как на основе прогноза, сформированного выбранной на данный момент моделью, так и в онлайн-режиме в случае неожиданных изменений в нагрузочных процессах (например, при резком скачке популярности какой-то функции системы).

**Выводы.** Основные результаты работы:

1. Выявлены факторы, влияющие на эластичность информационно-вычислительных систем.
2. Предложен подход к прогнозированию нагрузочных процессов на основе математического аппарата теории массового обслуживания, позволяющий провести первоначальный приблизительный расчет характеристик развертываемых систем.
3. Разработан алгоритм обработки временных рядов для работы с нагрузочными процессами, протекающими в эластичных системах.
4. Разработан метод автоматизированного управления структурой эластичных систем и их компонентами.