

АНАЛИЗ ЭФФЕКТИВНОСТИ СРЕДСТВ И МЕТОДОВ ОБРАБОТКИ ВРЕМЕННЫХ РЯДОВ ПРИ РАЗРАБОТКЕ РАСШИРЯЕМЫХ ПРОГРАММНЫХ СИСТЕМ

Райла М. (Университет ИТМО, Санкт-Петербург)

Научный руководитель – к.т.н, доцент, Перл И.А.

(Университет ИТМО, Санкт-Петербург)

В данном исследовании проводится обзор методов обработки временных рядов применительно к разработке расширяемых программных систем в области системной динамики и облачного моделирования. Проводится сравнение классических подходов к обработке, предлагаются рекомендации по повышению эффективности обработки временных рядов в условиях ограниченности вычислительных ресурсов.

Введение. В настоящее время существует ряд формальных математических методов применительно к обработке больших объемов временных рядов. Такой формат данных активно используется на фондовых биржах, в областях математического моделирования, системной динамики и системах, где требуется наблюдение за несистематическими элементами и рисками.

Данное исследование проводится в рамках международного проекта с открытым исходным кодом `sdLitica`. В основе проекта лежит программно-информационная система для обработки и анализа временных рядов. Основная задача – сократить выбросы в данных, привести к общему виду, провести классификацию, выявить формальные зависимости и превратить большие объемы данных в полезную информацию для дальнейшего использования.

Приняты выделять статистические и динамические методы обработки временных рядов. При рассмотрении статистических методов, решают задачи идентификации (нахождения исходных коэффициентов в линейных моделях) путем присвоения весовых коэффициентов элементам временного ряда. Основной проблемой при схожих подходах является наличие выбросов в данных, которое невозможно спрогнозировать ввиду стохастичности анализируемых процессов.

Для динамических методов основополагающей задачей является восстановление изначальной системы, порождающей временной ряд на основе векторов задержек в последовательные временные промежутки. На основе полученных результатов необходимо выстроить прогноз дальнейшей динамики последовательности.

Основная часть. В ходе исследований был проведен обзор методов обработки временных рядов. Основной особенностью применяемых методов является возможность решения задач корреляции, выявления аномалий в данных и предсказания дальнейшего поведения, что является актуальной научно-технической задачей.

Одной из поставленных целей исследования было сократить расходы на вычислительные мощности сервера, а также оптимизировать по времени процесс обработки временных зависимостей. Было предложено провести классификацию рассматриваемых временных рядов, что позволяет рассматривать и анализировать исторические временные данные сразу, не прибегая к дополнительным фильтрациям.

К разрабатываемому решению также был применен функциональный метод определения размерности вложения, что позволило визуализировать исходные данные с финансовых бирж и упростить задачи поиска корреляций.

Для оценки эффективности методов обработки был проведен ряд интеграционных и модульных тестов, позволяющих установить необходимость дальнейшего исследования доступных практик из областей системной динамики и облачного моделирования.

Выводы. Результаты, полученные в ходе данного исследования позволяют оценить эффективность использования методик при обработки временных рядов на основе разрабатываемой программно-информационной системы, а также позволяют заранее планировать архитектуру схожих платформи-независимых облачных решений, тем самым снижая оптимизационные затраты к разрабатываемым системам.

Райла М. (автор)

Подпись

Перл И.А. (научный руководитель)

Подпись