

РАЗРАБОТКА ВЕБ-ПЛАТФОРМЫ ДЛЯ АНАЛИЗА КЛИЕНТСКИХ ПУТЕЙ МЕТОДАМИ PROCESS MINING

Мухтаров Р. Э., Грачева А. Р.

Научный руководитель – канд. техн. наук Меншиков А. А.

Университет ИТМО

annagracheva14@yandex.ru

Введение

Событийные данные цифровых сервисов позволяют рассматривать пользовательское поведение как наблюдаемый процесс и применять процессную аналитику (process mining) для реконструкции клиентских путей по журналам событий [1, 2]. Существующие подходы ориентированы на выявление фактических траекторий с возвратами и ветвлениями, их структурирование и поиск «узких мест» [3]. Барьером практического использования process mining остаётся доступ к данным: интеграция часто строится через подключение к базам данных и настройку коннекторов, что увеличивает стоимость внедрения и может привести к нарушению конфиденциальности данных [4, 5]. Такой подход доминирует и в отечественных решениях, например, VK Process Mining и Sber Process Mining. При этом process mining платформы зачастую ориентированы на бизнес-процессы корпоративных ИС, а не на поведение пользователей. Веб-аналитика (например, Яндекс.Метрики) в основном даёт метрики без процессной модели event log. Таким образом, типовые решения либо ограничиваются веб-метриками, либо требуют сложного подключения к внутренним данным организации.

Основная часть

Предлагаемое решение представляет собой веб-платформу customer journey mining, которая объединяет удобство интеграции веб-аналитики и аналитическую силу process mining. Платформа собирает события с сайта через SDK, формируя унифицированный журнал событий с идентификаторами сессии, временными метками и семантическими шагами сценария. Это позволяет запускать процессную аналитику по пользовательским путям без подключения внешних компонентов к внутренним базам данных организации и без необходимости встраиваться в корпоративные контуры данных на этапе первичного внедрения.

Функционально платформа обеспечивает сбор и потоковую обработку событий, формирование аналитического хранилища и построение визуальных представлений пользовательских путей. Пользователь платформы может построить граф journey для выбранного сценария (например, путь до покупки или регистрации), определить наиболее распространённые варианты траекторий, выявить нетипичные ветвления, циклы и повторные попытки, а также определить этапы, на которых пользователи чаще всего прекращают взаимодействие.

Оригинальность подхода заключается в переносе process mining-представления на веб-поведение. Экономичность достигается за счёт разделения хранения событий и витрин, фоновое обновления агрегатов и отказа от регулярных полных пересчётов. В качестве исследовательской части предлагается серия экспериментов: сравнение сегментов и метрик до и после изменений интерфейса целевого сайта. Анализ результатов показал, что предложенное решение позволяет не только измерять итоговые показатели, но и объяснять, как именно пользователи проходят сценарии и где возникают причины потерь.

Выводы

Предложенная платформа обеспечивает построение процессной модели пользовательских путей по веб-событиям и поддерживает анализ вариантов прохождения, отклонений и «узких мест» без сложной интеграции с внутренними базами данных организации. Результаты исследования подтверждают, что применение process mining к данным, собранным через SDK, позволяет перейти от описания поведения набором метрик к интерпретируемой структуре траекторий, выявлять эффекты внесенных изменений на уровне формы путей и локализовать этапы, на которых формируются потери пользователей. Полученные выводы могут быть использованы для обоснования и приоритизации продуктовых изменений и для построения практик мониторинга пользовательских сценариев на основе событийных данных.

Литература

1. Topaloglu B., Oztaysi B., Dogan O. Cart-State-Aware Discovery of E-Commerce Visitor Journeys with Process Mining // *Journal of Theoretical and Applied Electronic Commerce Research*. 2024. Vol. 19, no. 4. P. 2851–2879. DOI: 10.3390/jtaer19040138.
2. Suriadi S., Andrews R., ter Hofstede A. H. M., Wynn M. T. Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs // *Information Systems*. 2017. Vol. 64. P. 132–150. DOI: 10.1016/j.is.2016.07.011.
3. Pal G., Atkinson K., Li G. Real-time user clickstream behavior analysis based on apache storm streaming // *Electronic Commerce Research*. 2023. Vol. 23. P. 1829–1859. DOI: 10.1007/s10660-021-09518-4.
4. Benaroch M. Third-party induced cyber incidents—much ado about nothing? // *Journal of Cybersecurity*. 2021. Vol. 7, no. 1. Art. tyab020. DOI: 10.1093/cybsec/tyab020.
5. González López de Murillas E., Reijers H. A., van der Aalst W. M. P. Connecting databases with process mining: a meta model and toolset // *Software and Systems Modeling*. 2019. Vol. 20, no. 4. P. 1381–1398. DOI: 10.1007/s10270-018-0664-7.