

**ИНФОРМАЦИОННЫЕ И ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ
В ТЕХНИКЕ И ОБРАЗОВАНИИ**

УДК 004.413

DOI 10.52348/2712-8873_MMTT_2021_8_79

**РАЗРАБОТКА ИНСТРУМЕНТА ДЛЯ ИЗВЛЕЧЕНИЯ И АНАЛИЗА МОДЕЛЕЙ
ПРОЦЕССА РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ****А.Е. Кущенко¹, А.В. Самочадин²***Санкт-Петербургский политехнический университет Петра Великого,
Россия, Санкт-Петербург,
kutshenko.ae@edu.spbstu.ru, ¹ samochadin_av@spbstu.ru ²*

Аннотация. Методы глубинного анализа процессов (Process Mining) позволяют извлекать знания из журналов событий, которые формируют информационные системы, поддерживающие деятельность различных организаций. Полученные в результате модели используются для анализа процессов и их усовершенствования. Методы Process Mining применяются в разных областях: продажи, финансы, медицина, логистика, производство, в IT и т. д. Одной из основных областей деятельности, в которой применяются методы Process Mining, является разработка программного обеспечения. Качество разрабатываемого ПО существенно зависит от эффективности процессов разработки, используемых разработчиками, и от соблюдения организационных требований, принятых в компании. Это определяет актуальность решения задачи отслеживания процессов разработки и создания их моделей для последующего анализа и улучшения. В работе описывается комплекс инструментальных средств, ориентированный на извлечение процессов из журналов событий, используемых в процессе разработки ПО информационных систем и построения их моделей.

Ключевые слова: интеллектуальный анализ процессов, разработка программного обеспечения, обнаружение модели процесса, проверка соответствия модели процесса, улучшение модели процесса, сеть Петри.

**DEVELOPMENT OF A TOOL FOR EXTRACTING AND ANALYZING
SOFTWARE DEVELOPMENT PROCESS MODELS****A. Kushchenko¹, A. Samochadin²***Peter the Great St. Petersburg Polytechnic University
Russia, St. Petersburg**kutshenko.ae@edu.spbstu.ru, ¹ samochadin_av@spbstu.ru ²*

Abstract. The methods of deep process analysis (Process Mining) allow to extract knowledge from the event logs generated by information systems that support the activities of various organizations [1]. The resulting process models are used to analyze processes and improve them. Process Mining methods are used in many areas: sales, finance, medicine, logistics, manufacturing, IT, etc. One of the main areas of activity in which Process Mining methods are used is a software development. The quality of the software depends on the effectiveness of the development processes used by developers, and on compliance with the organizational requirements adopted in the company. This determines the relevance of the solving problem of tracking development processes and creating their models for analysis and improvement. The paper describes a set of tools aimed at extracting processes from event logs used in the process of developing information systems software and building their models.

Keywords: process mining, software development, discovery, conformance checking, enhancement, Petri net.

Методы глубинного анализа процессов (Process Mining) позволяют извлекать знания из журналов событий, которые формируют информационные системы, поддерживающие деятельность различных организаций [1]. С использованием извлеченных моделей процессов организации могут определять показатели производительности, выявлять узкие места в бизнес-процессах и предпринимать шаги по улучшению процессов. Методы Process Mining позволяют сократить время, затраченное на понимание процессов и построения их моделей.

Методы Process Mining применяются в разных областях: продажи, финансы, медицина, логистика, производство, в IT и т.д. [2] В каждой из этих сфер присутствуют

свои варианты использования технологии. Компании могут проверить, соответствуют реальные рабочие процессы документированным в спецификации, выявить отклонения и проанализировать причины их возникновения. В результате могут быть приняты меры для уменьшения этих отклонений и обеспечения стандартизации процессов. Моделирование процессов является не менее актуальной задачей. Анализируя модели процессов, полученных из журналов событий, можно проводить их улучшения, оптимизацию, делать прогнозы.

Одной из основных областей деятельности, в которой применяются методы Process Mining, является разработка программного обеспечения. Качество разрабатываемого ПО существенно зависит от эффективности процессов разработки, используемых разработчиками, и от соблюдения организационных требований, принятых в компании. Многие компании сталкиваются с большими трудностями, связанными с тем, что большинство подходов к организации процесса разработки, по-прежнему, основаны на личном опыте и не используют эффективных методов. Возникают задачи отслеживания деятельности разработчиков для оптимизации и улучшения процесса разработки. Гибкие методологии разработки (Agile methodology) предоставляют возможность внедрения технологии Process Mining в процессе выполнения проекта разработки ПО. Многие из этих методологий предполагают разделение процесса разработки на итерации или «спринты» (используя терминологию Scrum), в промежутках между которыми процесс разработки можно формально проанализировать, чтобы выявить узкие места и предпринять шаги к его улучшению.

Целью инструментальных средств, описанию которых посвящена эта работа, является извлечение и анализ моделей процесса разработки программного обеспечения для последующего использования полученных моделей для улучшения процесса разработки.

Существует довольно большое число работ, посвященных применению технологии Process Mining в разработке ПО. Так, например, в работе [3] авторы рассматривают реальный опыт внедрения Process Mining в IT-отделе крупной автомобильной компании. Авторы перечислили список информационных систем, но не указали, какие события извлекались и каким образом. Для обнаружения модели процесса использовался алгоритм Heuristic Miner. Они описали конкретный список подходов к анализу и улучшению модели, однако, в рамках этой работы не применили их на практике к существующей модели. В другой публикации [4] внимание уделяется событиями, собранным из двух программных систем компании Jira Software. Разработка велась в соответствии с методологией Scrum. Авторы основное внимание уделили выявлению отклонений от следования требованиям Scrum.

В работах, посвященных использованию Process Mining в разработке ПО, большое внимание уделяется обнаружению процесса и предпринимаются шаги по его улучшению. Однако в них не учитывается один из важнейших этапов, предшествующих Process Mining, а именно, подготовке журнала событий. Возможно, информационные системы имеют в открытом доступе журнал в необходимом формате, однако это случается достаточно редко. В основном обработка журнала проводится на основе специально разрабатываемого инструмента для бизнес-системы, который отвечает за сбор данных и формирование журнала.

В настоящей работе важное место занимает разработка инструмента, настраиваемого на различные форматы журналов. В каждой организации определён перечень информационных бизнес-систем и возникает необходимость в разработке расширяемого инструмента для сбора исходных данных, их последующей фильтрации и формирования готового журнала.

Для выявления моделей процесса разработаны собственные и применены существующие инструментальные средства, базирующиеся на различных подходах. В

основе работы содержится технология Process Mining. Ранее модели процессов обычно создавались вручную без использования данных о событиях. Однако действия, выполняемые людьми, аппаратным и программным обеспечением, оставляют след в журналах событий. Каждое событие в таком журнале связано с деятельностью, т.е. с четко определенным шагом в процессе, и связан с конкретным case, т.е. экземпляром процесса. События, относящиеся к case, могут рассматриваться как один «прогон» процесса. Журналы событий могут хранить дополнительную информацию о событиях, такую как ресурс, т.е. лицо или устройство, инициировавшее действие, делают отметку времени события или прочие данные, которые записываются вместе с событием, например, размер заказа. В методах Process Mining журналы событий используются для решения трёх основных задач [5]: обнаружение модели процесса, проверка ее соответствия журналу событий и улучшение, направленное на изменение или расширение модели.

В подавляющем большинстве случаев обнаружение модели процесса сводится к построению сети Петри [6]. Сеть Петри определяется как двудольный ориентированный мультиграф, состоящий из вершин двух типов: позиций и переходов, соединённых дугами. В позициях могут размещаться метки (маркеры), способные перемещаться по сети. Событием называют срабатывание перехода, при котором метки из входных позиций этого перехода перемещаются в выходные позиции. Представление закрепилось в качестве стандарта в области Process Mining, так как наглядно отображает параллелизм, выбор и причинно-следственные связи между событиями. Кроме сетей Петри для представления модели используется известное расширение сети Петри, называемое workflow net [7]. Workflow nets в отличие от сетей Петри имеют один узел-источник и один узел-приемник. Следовательно, все узлы и переходы представляют путь от источника к приемнику.

Для решения задач Process Mining предлагается использовать наиболее известный инструмент с открытым исходным кодом ProM, который предоставляет широкие возможности. В ProM реализованы практически все основные инструменты Process Mining в виде плагинов, которых насчитывается более 200.

Разработка инструмента для извлечения исходных данных о событиях и построения готового журнала ведётся на языке Java с использованием фреймворка Spring. Большинство информационных бизнес-систем имеют открытый REST API, что позволяет средством HTTP протокола запрашивать сущности, содержащие информацию об интересующих событиях. Для взаимодействия с API используется библиотека Webflux, включающая класс WebClient, методы которого позволяют формировать HTTP запросы.

На рис. 1 схематично изображена архитектура комплекса инструментальных средств. Имеется инструмент (Event log extractor) для извлечения данных о событиях и построения журнала событий в формате привычном для методов Process Mining. Методы Process Mining используются для решения задач обнаружения модели процесса, проверки соответствия этой модели и её последующего улучшения. Далее на основе полученных результатов, аналитик бизнес-процессов может предпринять шаги по оптимизации процесса разработки.

Разрабатываемый инструмент (Event log extractor) состоит из двух взаимосвязанных компонентов.

Первый компонент предназначен для извлечения исходных данных о событиях из предопределённого списка информационных бизнес-систем. Он представляет собой приложение, которое работает в фоновом режиме и периодически опрашивает API различных бизнес-систем для получения и сохранения исходных данных о событиях. Предусмотрен удобный механизм расширения источников событий, в основе которого содержится шаблон проектирования “Фабрика”, позволяющий удобным способом

добавлять новые классы в проект. Данный шаблон определяет интерфейс для создания экземпляров одного из нескольких возможных классов. В этом интерфейсе определен метод, отвечающий за создание объектов, и его принято называть фабричным методом. Класс (конкретная фабрика), наследующий этот интерфейс, определяет экземпляр какого класса требуется инициализировать. Для каждого отдельного ресурса API, добавляется метод для его опроса и указываются поля, которые необходимо сохранить для дальнейшего формирования журнала.

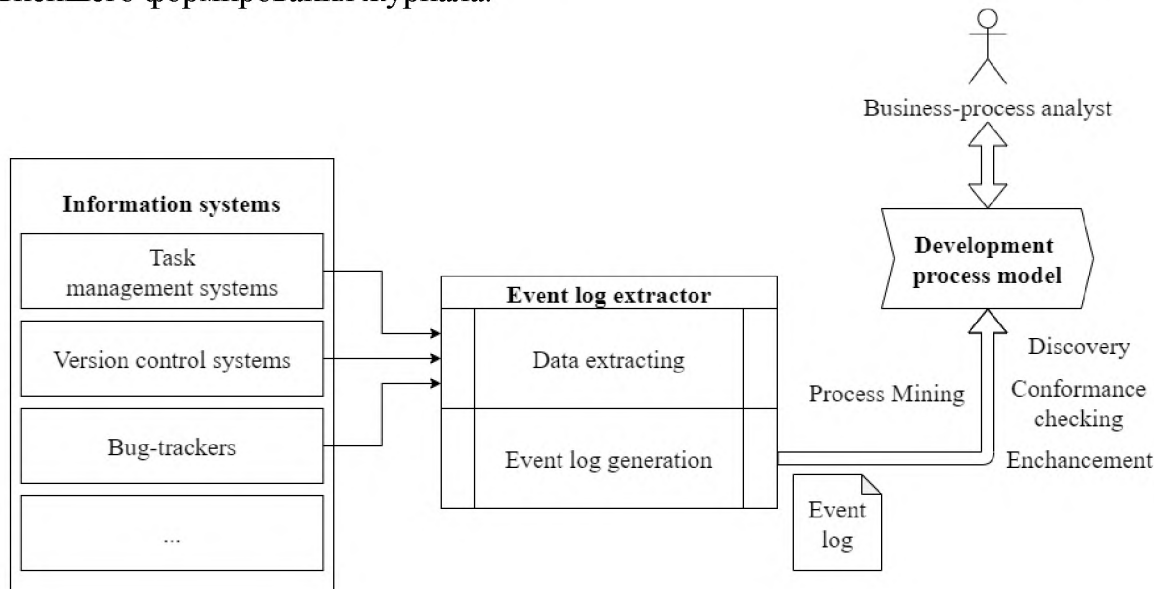


Рис. 1. Process Mining в процессе разработки ПО

Второй компонент использует извлечённые данные для формирования журнала событий, который в дальнейшем попадает на вход методов Process Mining. Как известно из определения, журналы должны быть разбиты на cases (экземпляры процесса). В проекте предусмотрен конфигурационный файл, в котором задаются параметры генерации журнала. В параметре “sources” перечисляются информационные бизнес-системы, события из которых будут участвовать в формировании журнала. В параметре “events” перечисляются типы событий, который попадут в журнал. Таким образом можно построить журнал определённого подпроцесса, например, указав события, относящиеся к тестированию или настройке непрерывной интеграции.

Перед тем, как применять методы Process Mining к процессу разработки программного обеспечения, следует определить источники данных с событиями.

В ходе работы разработчики взаимодействуют с рядом информационных систем, которые могут сохранять информацию, в определенной степени характеризующую процесс разработки. К системам, используемым в процессе разработки, можно отнести:

- Системы контроля версий (напр., GitLab, BitBucket).
- Системы управления задачами и баг-трекеры (напр., Youtrack, Jira)
- Интернет-браузеры (напр., Google Chrome, Edge)

Список может быть расширен любыми системами, функционирующими в организации, которые могут содержать информацию, отражающую процесс разработки.

Следующие сущности могут выступать в качестве исходных данных для построения журнала событий:

- Branches – ветки (событие по созданию, удалению, слиянию ветки).
- Commits – коммиты (события, связанные с фиксацией изменений любых файлов, входящих в репозиторий)
- Repositories – репозитории (события по инициализации репозитория).
- Wiki – вики (события по созданию или редактированию вики-страницы).

- Tasks – задачи (информация об изменении статуса задач в корпоративной системе управления задачами).
- Timetrack – затраченное время (добавление затраченного времени к задаче).
- Attachments – вложения (комментарии и вложения к задаче, которые играют роль обратной связи от сотрудника или руководства).
- Visit history – история посещений сайтов через web-браузер Chrome.

В результате применения разработанного комплекса к информационным системам GtiLab и YouTrack собраны экспериментальные данные за период в 7 дней и построен журнал событий. На основе извлечённых данных построена модель (см. рис. 2) в виде сети Петри. В качестве алгоритма для обнаружения модели использовался известный α -miner.

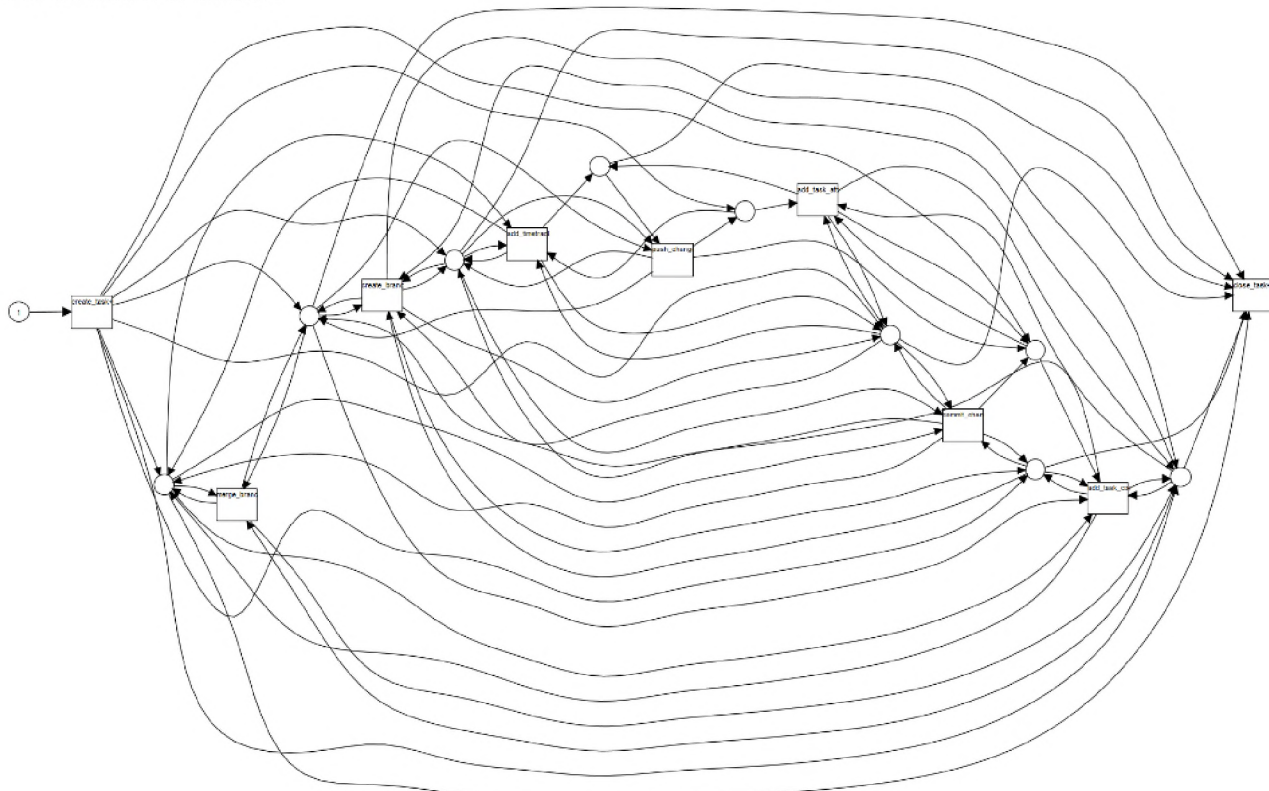


Рис. 2. Модель процесса, обнаруженная с использованием α -miner

Улучшение построенной модели и соответствующего процесса разработки является основным направлением будущей работы. На основе извлечённой модели планируется решить следующие задачи, направленные на оптимизацию и улучшение процесса разработки:

- Корректировка требований к регламенту разработки.
- Перераспределение задач.
- Обнаружение сотрудников, которые по каким-либо причинам не следуют требованиям к организации процесса разработки.
- Выявление слабых мест в процессе.

В настоящей работе предложен подход для интеллектуального анализа процесса разработки ПО, основанный на информации, доступной в информационных бизнес-системах, с которыми взаимодействуют разработчики ПО в процессе деятельности. Для извлечения данных о событиях из нескольких информационных систем и дальнейшего построения готового журнала разработан расширяемый инструмент с настраиваемой фильтрацией и сортировкой событий в журнале.

В работе также рассмотрены инструменты и подходы, направленные на построение модели процесса разработки ПО и её улучшения.

Библиографический список

1. Process Mining Manifesto / W. Aalst, A. Adriansyah, A. Medeiros, F. Arcieri, T. Baier, T. Blickle, P. Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. de Leoni, M. Wynn // *Lecture Notes in Business Information Processing*. 2011. № 99. pp. 169-194.
2. Thamizhara R., Appavoo K. A Review on Software Process Mining Using Petri Nets // *Asian Journal of Applied Sciences*. 2016. № 9. pp. 131-142.
3. Dilmegani C. 33 Use Cases / Applications of Process Mining in 2021 [Электронный ресурс] // AIMultiple Research. URL: <https://research.aimultiple.com/process-mining-use-cases/> (дата обращения: 01.05.21).
- Sebu M., Ciocarlie H., Applied process mining in software development // 2014 IEEE 9th International Symposium on
4. Applied Computational Intelligence and Informatics (SACI). 2014. pp. 55-60.
- Marques R., Mira da Silva M., Ferreira D. Assessing Agile Software Development Processes with Process Mining: A Case Study // 2018 IEEE 20th Conference on Business Informatics (CBI). 2018. pp. 109-118.
5. Aalst W. Process Mining: Discovery, Conformance and Enhancement of Business Processes // Springer. 2011. № 136.
6. Aalst W. The Application of Petri Nets to Workflow Management // *Journal of Circuits, Systems, and Computers*. 1998. № 8. pp. 21-66.