

## **Метод предварительной обработки журналов регистрации событий в задачах интеллектуального анализа процессов**

*В статье рассмотрена проблема подготовки данных к построению формальных моделей бизнес-процессов средствами processmining на основе анализа регистрационных журналов. Такие журналы содержат записи последовательностей событий, которые отображают порядок выполнения неформализованных бизнес-процессов. Для решения данной проблемы предложен метод предварительной обработки журналов регистрации событий, обеспечивающий преобразование регистрационных файлов в формат, пригодный для использования методов processmining.*

**Ключевые слова:** бизнес-процессы, процессный подход к управлению, шаблоны, журнал событий

### **Постановка проблемы**

В настоящее время широкое применение находят процессно-ориентированные системы, основанные на представлении деятельности организации, предприятия, рабочей группы в виде совокупности взаимодействующих процессов. В основе таких систем лежат процессно-ориентированные модели, отражающие последовательность деятельности организации по производству товаров и предоставлению услуг с учетом знаний, представленных в корпоративной культуре[1]. Это и обуславливает актуальность построения формальных процессно-ориентированных моделей, детально отражающих все действия по достижению конечного результата процесса с учетом ограничений, накладываемых корпоративной культурой.

Указанные модели обычно разрабатываются на основе существующей технической документации, а также результатов интервьюирования исполнителей. Проблема разработки формальных моделей бизнес-процессов связана с противоречием между «идеальным» представлением процесса, отраженным в документации, и реальным выполняемым процессом. Последний обычно содержит значительное количество деталей реализации, не отраженных в документации, но оказывающих существенное влияние на успешное выполнение бизнес-процесса. Указанные детали реализации связаны с особенностями корпоративной культуры, представленной в виде бизнес-правил.

Следовательно, решение проблемы построения формальных моделей бизнес-процессов связано с уточнением существующей модели (или построением новой) на основе фактической информации о реально выполняющихся процессах.

Источником такой информации являются регистрационные файлы и базы данных информационных систем в организациях, где выполняются бизнес-процессы.

В настоящее время существуют два основных вида регистрационных данных о процессах:

– структурированные журнальные файлы процессно-ориентированных информационных систем, логически отражающие последовательность действий бизнес-процесса во времени;

– журнальные файлы систем коллективной работы, регистрирующие данные обо всех происходящих в системе действиях с привязкой последних ко времени.

Построение моделей бизнес-процессов на основе извлечения и обработки данных из файлов первого типа осуществляется с помощью методов интеллектуального анализа процессов - processmining [1,2]. Применение методов processmining позволяет:

– разработать «черновую» модель бизнес-процесса;

– выявить узкие места, а также бизнес-правила.

Основное ограничение для существующих методов интеллектуального анализа процессов состоит в том, что исходные данные для построения модели должны быть логически структурированы таким образом, чтобы отображать фактически выполнившиеся последовательности событий в рамках процесса. Каждое из таких событий отражает выполнение той или иной процедуры бизнес-процесса. Формирование указанных структурированных журнальных файлов осуществляется только в BPMS системах в автоматическом режиме.

В то же время, существует значительное количество организаций, где используются информационные системы, не поддерживающие

бизнес-процессы. Текущие действия в таких системах протоколируются с метками времени. Иными словами, регистрационные файлы в таких системах содержат информацию, необходимую для построения процессной модели, отражающей реально выполняющиеся бизнес-процессы. Однако указанная информация не является упорядоченной в соответствии с существующими требованиями интеллектуального анализа процессов. Это и обуславливает актуальность темы статьи.

Вышеизложенное определяет актуальность проблемы автоматизированного преобразования регистрационных файлов в формат, пригодный для формирования моделей бизнес-процессов методами processmining.

### Постановка задачи

Задача предварительного преобразования журнала регистрации событий в форму, пригодную для проведения интеллектуального анализа процессов, заключается в следующем. Регистрационные файлы (регистрационные базы данных) систем коллективной работы необходимо преобразовать в структурированные журнальные файлы, аналогичные журналам регистрации событий процессно-ориентированных информационных систем.

Исходными данными задачи являются:

- собственно регистрационный файл, содержащий в общем случае неупорядоченное описание событий с указанием метки времени;
- перечень обрабатываемых объектов предметной области и их атрибутов.

Необходимо разработать метод преобразования структуры регистрационного файла системы коллективной работы в файл регистрации событий процессно-ориентированной информационной системы, пригодный к использованию в качестве входных данных для решения задач интеллектуального анализа бизнес-процессов.

Основные шаги решения поставленной задачи включают в себя:

- анализ структур журналов регистрации событий обоих типов;
- разработка метода предварительной обработки журнала регистрации событий, обеспечивающего преобразование журнала первого типа в журнал второго, пригодный для обработки методами processmining;
- разработка моделей представления указанных журналов.

В данной статье рассматриваются первый и второй шаги решения поставленной задачи

### Жизненный-цикл бизнес-процессов.

На сегодняшний день множество предприятий функционирует в соответствии с процессно-ориентированным подходом.

Данный подход подразумевает ориентированность предприятия и функционирующей в нем информационной системы на бизнес – процессы (БП). Последние определяют последовательность действий по достижению заданного результата (товара, услуги).

Жизненный цикл (ЖЦ) процессно-ориентированных информационных систем таких информационных систем основан на ЖЦ БП, который в свою очередь включает следующие этапы (рис.1):

- анализ требований;
- перепроектирование
- реализация;
- внедрение и мониторинг;
- согласование.



Рисунок 1 – Жизненный цикл БП

К процессно-ориентированным системам относят системы, имеющие следующие свойства:

- в основе таких систем используются средства управления бизнес-процессами, зачастую в качестве такого средства выступает WFM, хотя последнее время все больше организаций переводят свои системы на BPM;
- организация применяет бизнес-правила для управления БП;
- организация использует высококачественное промежуточное программное обеспечение для связки функционирующих в рамках данной организации систем;
- в рамках организации используются средства мониторинга БП;
- информация о протекании БП хранится в общем хранилище данных;

– в рамках организации налажена система обратной связи (связи потребителя с производителем);

– в рамках предприятия осуществляется сбор и хранения регистрационной информации.

Среди процессно-ориентированных систем можно выделить следующие типы

- классические WFM системы;
- BPM системы;
- системы планирования ресурсов предприятия ERP;
- системы управление данными о продукте PDM;
- системы управления связями с клиентами CRM.

Все вышеперечисленные типы систем поддерживают функцию регистрации событий. Регистрируемая информация храниться в базе данных в виде журналов регистрации событий (ЖРС). Хранение файлов регистрации событий чаще всего осуществляется в виде текстового файла или таблиц в базе данных. ЖРС каждой из систем имеют ряд отличий, что связано со спецификой каждой из систем. Данные отличия заключаются в разных по смыслу и по количеству атрибутах. Так, классические WFM системы содержат подробную информацию о структуре БП, системы планирования ресурсов предприятия ERP – о ресурсах и их потребностях в рамках технологического процесса, системы управления связями с клиентами CRM – о клиентах и коэффициентах описывающих интересы клиентов и т.д.

Технология Process mining предполагает извлечение информации, знаний о БП и дальнейшее построение его модели посредством анализа ЖРС. Технология Process mining базируется на технологии интеллектуального анализа данных (Datamining).

Process mining применяется в рамках BPM систем. Данная технология позволяет отобразить процесс выполнения БП, выявить скрытые связи между объектами БП, значительно повысить актуальность и качество модели БП.

На рис.2 представлен ЖЦБП и место интеллектуального анализа бизнес-процессов в рамках жизненного цикла [3]. Таким образом очевидно, что интеллектуальный анализ процессов применим на всех этапах ЖЦ БП:

– на этапе анализа требований позволяет обеспечить поддержку автоматизированного формирования моделей существующих БП (AS-IS);

– на этапе перепроектирования обеспечивается возможность верификации построенной улучшенной модели, согласование модели с построенными деревьями целей и критериев;

– на этапе реализации обеспечивается возможность поддержки спецификации на проектируемую модель;

– на этапе внедрения и мониторинга осуществляется предоставление информации по последовательности реализации процедур и обработки объектов функционирующего бизнес-процесса.

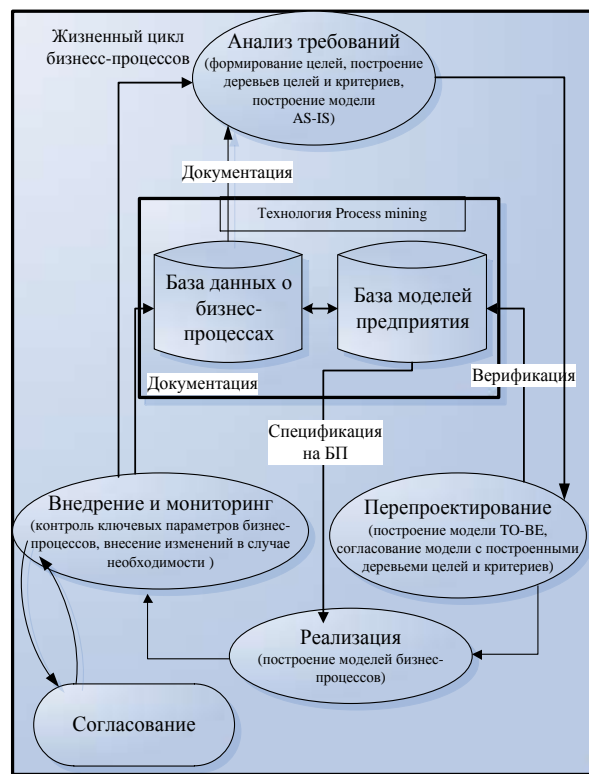


Рисунок 2 – место технологии Process mining в ЖЦ БП

### Структуризация журналов регистрации событий

Выполненный анализ регистрационных файлов как систем коллективной работы, так и процессно-ориентированных систем показал, что для файлов обоих типов характерны следующие особенности:

– зарегистрированные события отражают элементарные действия в информационной системе;

– зарегистрированные события имеют атрибут времени;

– зарегистрированные события, как правило, связаны с теми или иными ресурсами.

В то же время проведение интеллектуального анализа процессов требует журнального файла, обладающего следующей структурой:

- идентификатор последовательности выполнения процедуры;
- временной параметр;
- действие (событие);
- ресурс (пользователь).

Перечисленные выше атрибуты являются обязательными, но также журнальные файлы

могут содержать ряд дополнительных атрибутов: стоимость, затраченные ресурсы и др.

Анализ регистрационных файлов систем коллективной работы показал, что большинство журналов имеют отличную от представленной выше структуру, состоящую из следующих обязательных атрибутов:

- временной параметр;
- действие (событие);
- ресурс (пользователь);
- документ (объект) т.е. результат совершенного действия;

Таким образом, при сравнительном анализе видно, что главное отличие между двумя видами журналов регистрации событий заключается в отсутствии атрибута «идентификатор последовательности выполнения процедуры», который отображает последовательность событий при реализации бизнес-процесса. Такое отличие ограничивает возможность использования методов интеллектуального анализа процессов для многих информационных систем, ориентированных на коллективную работу.

### Метод предварительной обработки журналов регистрации событий в задачах интеллектуального анализа процессов

В основе предлагаемого метода лежит общая схема предварительного преобразования структуры журнала регистрации событий, представленная на рис.3.



Рисунок 3 – Общая схема предварительного преобразования структуры журнала регистрации событий

В соответствии с представленной схемой, атрибут «идентификатор последовательности выполнения процедуры» формируется на основе атрибутов событий имеющегося ЖРС, а также таблицы, описывающей принадлежность входов БП к действиям и объектам, обрабатываемым в рамках бизнес-процесса.

Метод преобразования журналов регистрации событий систем коллективной работы структуры, пригодной для решения задач интеллектуального анализа процессов, включает следующие пять этапов:

- отбор подмножеств событий, происшедших в заданном интервале времени, и связанных с этими событиями объектов;
- формирование пар объект – действие для каждого отобранного на первом этапе подмножества;
- выявление начальных и конечных событий будущих последовательностей для каждого отобранного подмножества;
- формирование наборов частично упорядоченных последовательностей событий;
- формирование журнала регистрации событий с заданной структурой.

Детальнее рассмотрим каждый из этапов.

На первом этапе происходит формирование перечня объектов, событий, процедур, отраженных в журнале регистрации событий за заданный отрезок времени. Отобранное подмножество в дальнейшем используется для формирования таблицы с описанием входной и выходной информации событий, а также отбора подмножеств последовательностей событий, отражающих различные траектории реализации одного и того же процесса либо подпроцесса.

На втором этапе происходит формирование пар объект-действие, где выбранные на первом этапе объекты позволяют описать входы и выходы соответствующих событий. Пары представляются в табличном виде с введением дополнительного атрибута (параметра) «тип», который описывает место объекта по отношению к рассматриваемому участку БП или всему бизнес-процессу. Пример структуризации входной информации для событий отображен в таблице 1.

Таблица 1 – Структуризация входной информации отобранных множеств событий БП

| Входная информация | Тип входной информации | Событие   |
|--------------------|------------------------|-----------|
| Информация_1       | объект                 | Событие_2 |
| Информация_2       | внешние данные         | Событие_1 |
| Информация_3       | объект                 | Событие_3 |
| ...                | ...                    | ...       |
| Информация_n       | объект/ внешние данные | Событие_n |

На третьем этапе осуществляется выявление начальных конечных событий для отобранных на первом этапе подмножеств.

Критерием отбора начальных событий является наличие внешних данных (объектов), связанных с отобранными событиями.

Критерием отбора конечных событий является наличие внешних выходных данных. Иными словами, конечными будут те события, выходная информация которых не связана ни с одним из событий отобранного множества.

На четвертом этапе происходит формирование подмножеств частично упорядоченных последовательностей событий по времени и входным - выходным данным. Данный этап позволяет получить наборы событий, описывающие различные варианты реализации одного и того же подпроцесса, что обеспечивает в дальнейшем возможность построения модели БП путем объединения различных путей его реализации.

Каждый вариант реализации процесса на данном этапе получается путем связывания событий через общие входные и выходные данные, полученные на предыдущих этапах.

Критерием, характеризующим подмножества событий, отражающих различные траектории реализации одного и того же процесса либо подпроцесса, является наличие в них общих элементов (событий).

Отметим, что каждая из полученных на данном этапе последовательностей соответствует элементу CASE в структуре журнала процессно-ориентированной системы.

На последнем этапе осуществляется интеграция полученных подмножеств в журнал регистрации событий, пригодный для решения задач интеллектуального анализа процессов.

### Выводы

Проведенный в статье анализ показал, что современные подходы к process mining основаны на использовании специально структурированных журнальных файлов процессно-ориентированных информационных систем. Такие файлы структурно отражают последовательность действий бизнес-процесса во времени. В то же время для организаций, в которых процессно-ориентированные системы не используются, исходные данные о бизнес-процессах могут быть получены лишь на основе обработки текущих неструктурированных регистрационных файлов (либо баз данных). Это и обуславливает важность разработки метода предварительной обработки журнала регистрации событий.

В статье предложен метод предварительной обработки журнала (файла) регистрации событий, который заключается в построении упорядоченных во времени множеств последовательностей событий бизнес-процесса.

Каждое из указанных множеств отражает обработку связанной последовательности объектов на заданном интервале времени и соответствует одному из подпроцессов общего бизнес-процесса.

Предлагаемый метод позволяет, во-первых, преобразовать исходные неструктурированные с позиций процессного подхода данные к виду структурированных журнальных файлов, для которых применимы существующие методы интеллектуального анализа бизнес-процессов.

Во-вторых, метод обеспечивает возможность построения моделей бизнес-процессов с изменяемой структурой на основе сочетания преимуществ workflow и сервис-ориентированной архитектуры.

### Список литературы

1. W.M.P. van der Aalst Process-Aware Information Systems: Lessons to be Learned from Process Mining / W.M.P. van der Aalst // Department of Mathematics and Computer Science Eindhoven University of Technology. – Eindhoven, 2000. – 26 p.
2. W.M.P. van der Aalst Business Process Management Demystified: A Tutorial on Models, Systems and Standards for Workflow Management / W.M.P. van der Aalst // Department of Mathematics and Computer Science Eindhoven University of Technology. – Eindhoven, 2004. – 68 p.
3. Process Mining: Discovery, Conformance and Enhancement of Business Processes / W.M.P. van der Aalst . – N.-Y.: Springer Verlag. – 2011 . – 370.

### Авторы:

**В.М. Левыкин**, доктор технических наук, заведующий кафедрой информационных управляющих систем Харьковского национального университета радиоэлектроники

**С. Ф. Чалий**, доктор технических наук, профессор кафедры информационных управляющих систем Харьковского национального университета радиоэлектроники

**Е. О. Богатов**, аспирант кафедры информационных управляющих систем Харьковского национального университета радиоэлектроники

**М. И. Дикусар**, магистрант кафедры информационных управляющих систем Харьковского национального университета радиоэлектроники

### Метод попередньої обробки журналів реєстрації подій в задачах інтелектуального аналізу процесів

В.М. Левикін, С.Ф. Чалий, Є.О. Богатов, М.І. Дікусар

У статті розглянута проблема підготовки даних до побудови формальних моделей бізнес-процесів засобами інтелектуального аналізу процесівна основі обробки реєстраційних журналів. Останні містять записи послідовностей подій, які відображають порядок виконання неформалізованих бізнес-процесів. Для вирішення цієї проблеми запропонований метод попередньої обробки журналів реєстрації подій, що забезпечує перетворення реєстраційних файлів у формат, придатний для використання методів processmining.

**Ключові слова:** *бізнес-процеси, інтелектуальний аналіз процесів, виявлення бізнес-процесів, журнал реєстрації подій*

### **The method of logs preprocessing in tasks of process mining**

S. F. Chaliy, E. O. Bogatov, M. I. Dikumar

The analysis of files of registration of events of the process-oriented information systems is conducted in the article. It is certain on the basis of analysis, that technology of process mining is not applicable within the framework of row of non-process-oriented information systems. For the decision of this problem the method preprocessing of log files is offered.

**Keywords:** *business processes, process mining, identifying business processes, event logs*