

УДК 004.891.3

ПОПОЛНЕНИЕ БАЗЫ ЗНАНИЙ ИУС С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ АНАЛИЗА ЛОГОВ

Левыкин В.М., Чалая О.В.

Методы process mining (интеллектуального анализа процессов) ориентированы на построение графа последовательности работ (workflow). В качестве исходных данных используются логи событий информационно-управляющей системы (ИУС). Записанные в логе события фиксируют факт выполнения действий лога. Вершины и дуги результирующего графа соответствуют событиям лога и переходам между событиями. Важность действий (событий) процесса при построении модели обычно не учитывается. В ряде подходов учитывается частота повторения действий [1, 2].

Недостаток парадигмы process mining состоит в том, что из полученного графа очень сложно выделить типовые последовательности действий и соответствующие им базовые правила их применения. Это приводит к неадекватности знаний о происходящих на предприятии процессах и решаемых задачах. Как следствие, снижается эффективность управления.

В то же время при решении задач управления необходимо учитывать ограничения на последовательность обработки объектов предметной области, а также на допустимые изменения их состояния [3, 4]. Отметим, что в соответствии с предлагаемым фирмой IBM подходом такие объекты принято именовать артефактами [3]. Указанный подход позволяет выполнить описание контекстных преобразований объектов при решении управленческих задач. Контекст при этом описывается как совокупность атрибутов всех используемых объектов.

Однако ориентированные на артефакты методы рассматривают только взаимодействие отдельных объектов и не учитывают процесс решения задачи в целом. Это не позволяет связать состояние контекста и выбираемые с учетом этого состояния цепочки действий. Иными словами, такой подход не позволяет учесть знания о порядке выбора требуемых последовательностей действий в заданном контексте, что также приводит к снижению адекватности используемых при управлении моделей.

В то же время состояние контекста для каждого выполняемого действия фиксируется в логе информационно-управляющей системы в виде набора значений атрибутов событий [5], что свидетельствует о возможности извлечения этих зависимостей и последующем включении их в базу знаний ИУС [6].

Таким образом, можно сделать вывод о том, что существующие методы process mining используют в качестве входной информации последовательность событий в логе и не выявляют знания о влиянии состояния артефактов контекста на порядок выполнения действий при решении практических задач. В то же время учет влияния состояния артефактов контекста на последовательность операций позволяет выделить знания об условиях запуска и корректировки workflow как алгоритма действий по решению прикладной задачи. Поэтому выявление знаний, отражающих зависимости между состоянием контекста, представленного свойствами используемых объектов, а также действиями по обработке таких артефактов позволяет обеспечить поддержку принятия решений в рамках информационно-управляющей системы и, как следствие, повысить эффективность управления.

Разработанный метод направлен на выявление правил, отражающих выбор отдельных действий либо последовательностей действий по решению практических задач в зависимости от состояния контекста. Метод включает в себя следующие этапы:

филтрация событий логa; формирование повторяющихся последовательностей событий; формирование повторяющихся наборов атрибутов, а также артефактов; построение зависимостей между артефактами и событиями процесса.

Задача первого этапа состоит в том, чтобы удалить избыточные данные из логa, а также создать множество уникальных событий, каждое из которых соответствует одной операции. Необходимость данного этапа связана с тем, что с каждой операций могут быть связаны несколько событий логa (например, при смене исполнителя, приостановке работы и т.п.)

Задача второго этапа состоит в том, чтобы найти повторяющиеся на различных трассах логa не менее двух раз последовательности событий. Выбор такого уровня поддержки определяется тем, что типовое решение должно быть использовано неоднократно. При реализации данного этапа не рассматриваются связи между событиями, а также временная составляющая, поэтому используется адаптированный к особенностям представления информации в лог алгоритм AprioriAll.

Задача третьего этапа заключается в выделении взаимосвязанных атрибутов и артефактов процесса. При реализации данного этапа используется адаптированный алгоритм поиска ассоциативных зависимостей FPG (Frequent Pattern-Growth), применяемый в области в Data Mining. Адаптированный подход отличается тем, что такие зависимости формируются комплексно, для двух уровней иерархии: артефакт; а также атрибут и значение атрибута.

Четвертый этап посвящен формированию зависимостей между артефактами и событиями, отражающими действия. На данном этапе также уточняются атрибуты, которые входят в состав полученной зависимости. Для этого выполняется выявление зависимостей между состоянием артефактов контекста и событиями логa, отражающими действия процесса.

В результате анализа логов ИУС могут быть получены правила двух видов. Зависимости первого вида отражают известные формализованные знания, имеющиеся в БЗ и используемые для поддержки принятия решений. Зависимости второго вида отражают персональные знания неформализованные исполнителей. Формализация таких правил и последующее включение их в БЗ позволяет обосновано выбирать процессы (подпроцессы) решения прикладных задач и тем самым повысить эффективность управления средствами ИУС.

ПЕРЕЧЕНЬ ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Van der Aalst, W. M. P. Process Mining: Discovery, Conformance and Enhancement of Business Processes [Text] / W. M. P. Van der Aalst. – Springer Berlin Heidelberg, 2011. – 352 p. doi:[10.1007/978-3-642-19345-3](https://doi.org/10.1007/978-3-642-19345-3)
2. Van der Aalst, W. M. P. Process Mining in the Large: A Tutorial [Text] / W. M. P. Van der Aalst // Business Intelligence. – Springer Science + Business Media, 2014. – P. 33–76. doi:[10.1007/978-3-319-05461-2_2](https://doi.org/10.1007/978-3-319-05461-2_2)
3. Cohn, D. Business artifacts: A data-centric approach to modeling business operations and processes [Text] / D. Cohn, R. Hull // Bulletin of the IEEE Computer Society Technical Committee on Data Engineering. – 2009. – Vol. 32, № 3. – P. 1–7.
4. Bhattacharya, K. Artifact-centered operational modeling: Lessons from customer engagements [Text] / K. Bhattacharya, N. S. Caswell, S. Kumaran, A. Nigam, F. Y. Wu // IBM Systems Journal. – 2007. – Vol. 46, № 4. – P. 703–721. doi:[10.1147/sj.464.0703](https://doi.org/10.1147/sj.464.0703)
5. Günther C. W. OpenXES. Developer Guide [Text] / C. W. Günther, E. Verbeek. Technische Universiteit Eindhoven University of Technology, 2014.– 38 p.
6. Левыкин В. М. Выделение элементов контекста знание-емких бизнес-процессов на основе анализа логов // В. М. Левыкин, О.В. Чалая // Технологический аудит и резервы производства. – 2016. - № 5/2(31). - С. 65-71.