

УДК 004.891.3

С.Ф. ЧАЛЫЙ, И.В. ЛЕВЫКИН

Харьковский национальный университет радиоэлектроники

### РАЗРАБОТКА МЕТОДА ФОРМИРОВАНИЯ ПРОЦЕССНОЙ МОДЕЛИ РЕШЕНИЯ ЗАДАЧИ В СОСТАВЕ ПРЕЦЕДЕНТА

*В данной работе решается задача разработки метода построения процессной модели решения задачи, как составляющей прецедента. Особенность построения такой модели состоит в интеграции зависимостей, содержащихся как в логах процесса, получаемых при фиксации его выполнения в виде последовательности событий, так и в дискретных моделях процессов, получаемых методами process mining. Результатом использования данного метода является модель процесса решения задачи с возможностью адаптации с учетом ограничений предметной области и ожидаемых результатов. Оценка ограничений выполняется на базе интервального представления времени. Данная оценка позволяет выяснить возможность распараллеливания действий процесса. Адаптация выполняется путем удаления из модели трасс процесса, которые не удовлетворяют временным ограничениям, а также смещения фрагментов трасс на временной линии выполнения процесса.*

*Ключевые слова: анализ процессов (process mining), рассуждения на основе прецедентов (case-based reasoning), лог (журнал событий).*

С.Ф. ЧАЛИЙ, І.В. ЛЕВИКІН

Харківський національний університет радіоелектроніки

### РОЗРОБКА МЕТОДУ ФОРМУВАННЯ ПРОЦЕСНОЇ МОДЕЛІ РІШЕННЯ ЗАДАЧІ У СКЛАДІ ПРЕЦЕДЕНТА

*У даній роботі вирішується завдання розробки методу побудови процесної моделі рішення задачі як складової прецеденту. Особливість побудови такої моделі полягає в інтеграції залежностей, що містяться як в логах процесу, одержуваних при фіксації його виконання у вигляді послідовності подій, так і в дискретних моделях процесів, одержуваних методами process mining. Результатом використання даного методу є модель процесу рішення задачі з можливістю адаптації з урахуванням обмежень предметної області і очікуваних результатів. Оцінка обмежень виконується на базі інтервального представлення часу. Дана оцінка дозволяє з'ясувати можливість розпаралелювання дій процесу. Адаптація виконується шляхом видалення з моделі трас процесу, які не задовольняють тимчасовим обмеженням, а також зміщення фрагментів трас на тимчасовій лінії виконання процесу.*

*Ключові слова: аналіз процесів (process mining), розміркування на основі прецедентів (case-based reasoning), лог (журнал подій)*

S.F. CHALYI, I.V. LEVYKIN

Kharkiv National University of Radio Electronics

### DEVELOPMENT OF METHOD FORMING PROCESS MODEL SOLUTION TASKS OF A PART CASE-BASED REASONING

*In this paper, we solve the problem of developing a method of constructing a process model for solving tasks of a part case-based reasoning. The peculiarity of the construction of such a model is to integrate dependencies contained in the logs of the process, obtained by fixation of its implementation in the form of a sequence of events and processes in discrete models, derived methods process mining. The result of using this method is the model of the process of solving the problem with the ability to adapt to the constraints of the subject area and the expected results. Evaluation limits is performed on the basis of representation of interval time. This assessment allows us to determine the ability to parallelize the process of action. Adaptation is performed by removing the process from the model runs, which do not meet the time constraints, as well as the displacement tracks the tracks on the timeline of a process.*

*Keywords: process mining, case-based reasoning, event log*

#### Постановка проблемы

Рассуждения на основе прецедентов (CBR – Case-Based Reasoning) [1,2] основаны на использовании накопленного опыта решения задач при работе с новыми заданиями [3], поэтому ключевым элементом прецедента-аналога является процесс решения задачи. Цикл рассуждений

включает в себя построение (извлечение), использование, адаптацию и сохранение прецедента. Построение процесса решения задачи в составе прецедента определяется спецификой предметной области, что и обуславливает применимость методов process mining [4] для его формирования. Методы process mining используют формируемые информационной системой логи выполняющихся процессов для построения моделей таких процессов с дискретным представлением времени. В то же время при практическом использовании прецедента-аналога, в особенности при параллельном решении нескольких однотипных задач с использованием одного прецедента возникает проблема оценки продолжительности решения задачи с тем, чтобы она удовлетворяла ограничениям предметной области. Для решения этой проблемы необходимо перейти от событийного описания процесса в виде лога, а также моделей процесса с дискретным временем к моделям с интервальным представлением времени, содержащим оценку времени выполнения отдельных операций процесса. Изложенное определяет актуальность темы данной работы.

#### **Анализ последних достижений и публикаций**

Представленные в работе [5] общая модель прецедента, а также метод выбора, сравнения и корректировки прецедента аналога требуют дальнейшей формализации элементов модели текущего прецедента.

Элементы прецедента определяют условия его использования, а также последовательность решения задачи. В качестве рамочного метода к построению моделей процесса решения задачи в работе используется методология интеллектуального анализа процессов (process mining) [4]. Ключевая особенность методологии состоит в том, что модель процесса решения задачи формируется на основе интеграции трасс лога. Каждая из трасс фиксирует однократное решение задачи в виде последовательности событий, возникающих в ходе процесса. Каждое событие обычно содержит временную метку. События отражают выполнение действий процесса. Каждому действию процесса может соответствовать несколько событий лога. Общее обоснование использования данной методологии при построении элементов прецедента было приведено в работах [6-9].

Полученная модель отражает практически реализованные с учетом различных внешних воздействий варианты решения задачи. Указанные варианты представляются в виде логически связанных последовательностей событий, включая ветвление [10]. Однако полученная методами process mining модель не учитывает продолжительность действий процесса решения задачи и потому не позволяет выполнить оценку продолжительности его выполнения.

#### **Формулирование цели исследования**

Целью работы является задачи разработки метода построения процессной модели, объединяющей дискретное и интервальное описание процесса решения задачи. Особенность построения такой модели состоит в интеграции зависимостей, содержащихся в логах процесса, получаемых при фиксации его выполнения в виде последовательности событий, а также в дискретных моделях процессов, получаемых методами process mining.

#### **Изложение основного материала исследования**

При выборе подходящего решения задачи (прецедента-аналога), необходимо учитывать не только описание проблемы  $Z$ , но и ограничения на процесс решения задачи. Одним из существенных ограничений является ограничение по времени выполнения процесса, которое зависит от времени выполнения отдельных операций процесса решения задачи.

Продолжительность выполнения операций зависит от ряда факторов, в частности:

– варианта ограничений предметной области  $Πo_m \equiv \bigwedge_1 Πo_{m,l}$ ;

– непредвиденных воздействий на ход течения процесса, что потребует в дальнейшем изменения последовательности действий и приведет к получению альтернативного варианта решения задачи  $ΠP_s$ .

Примером первого фактора является набор ограничений по материальным ресурсам, в частности допустимой загрузке имеющегося оборудования. Поломка оборудования является примером второго фактора.

Поэтому при построении модели процесса решения задачи в составе прецедента необходимо перейти от событий к действиям процесса с учетом их временных параметров (рис.1).

Каждое действие обладает продолжительностью во времени, которую можно оценить на основе сравнения временных меток события, свидетельствующего о начале действия и о его завершении. В силу рассмотренных выше ограничений продолжительность одного и того же действия для различных трасс может отличаться.

В общем случае каждое действие процесса может быть отражено в логе в виде нескольких событий. Каждое событие отражает изменение в состоянии процесса. Иными словами, событие

представляет собой один из результатов действия, которое переводит процесс из старого состояния в новое.

При построении модели процесса необходимо учитывать слияние различных трас лога, а также ошибки и неточности, возникающие при формировании лога информационной системой. Поэтому временной интервал выполнения операции в общем случае определяется на основе временных меток нескольких событий.

Предложенный в данном подразделе метод основан на дополнении полученной средствами process mining модели процесса интервальным представлением времени с учетом особенностей записи временных меток в реальных информационных системах.

В соответствии с данным методом процесс решения отображается на следующих уровнях (рис. 1):

- набор последовательностей событий, отражающих выполнение процесса решения задачи;
- качественное описание решения задачи;
- количественное описание решения задачи.

Первый уровень содержит множество последовательностей событий, каждая из которых отражает однократное выполнение процесса решения задачи. Каждое событие характеризуется набором атрибутов, включающих в себя в частности:

- наименование выполнившегося действия;
- временная метка, содержащая время возникновения события;
- исполнитель действия;
- страна, город, в котором выполняется процесс;
- подпроцесс (регулярный фрагмент процесса решения), которому принадлежит данное действие.

На практике данный уровень обычно реализован в виде лога (журнала событий) информационной системы.

Качественное описание процесса отражает алгоритм его действий, и формируется методами интеллектуального анализа процессов. Результирующая модель процесса решения задачи может быть представлена с использованием математических аппаратов сетей Петри, темпоральных логик, различных алгебр, графов событий а также EPC и UML – диаграмм [10-12]. Качественное описание процесса позволяет оценить достижимость его подмножества конечных состояний.

Таким образом, уровень качественного представления процесса позволяет выяснить, какой набор решений задачи  $\{PP_i\} \subseteq PP$  может быть получен в результате выполнения процесса при заданных условиях внешней среды  $\bigwedge_j Po_{ij}$ .

Данный уровень основан на дискретном представлении времени и потому не позволяет оценить продолжительность решения задачи.

Уровень количественного представления дает возможность получить количественные оценки времени для альтернативных вариантов процесса решения,

В качестве исходных данных для предлагаемого метода выступает лог процесса решения прецедентной задачи. Лог объединяет все реализованные на практике варианты решения задачи.

$$\Pi = \{\pi_k\}, k = \overline{1, K}, \quad (1)$$

где  $\pi_k$  – трасса процесса, отражающая его однократное выполнение.

Каждый процесс решения отображен в логе в виде трассы, состоящей из конечной последовательности событий, принадлежащих подмножеству,

Между событиями данной последовательности задано отношение переходов  $\rightarrow$  :

$$\pi_k = \langle e_{k,1} \rightarrow \dots \rightarrow e_{k,i} \rightarrow e_{k,i+1} \rightarrow \dots \rightarrow e_{k,l} \rangle, i = \overline{1, l}, e_{k,i} \in E_k, E_k \subseteq E, \quad (2)$$

где  $E$  – полное множество событий процесса.

Данное отношение сопоставляет каждому событию  $e_{k,i}$  последующее событие  $e_{k,i+1}$  на трассе  $\pi_k$  процесса решения задачи  $Z$ .

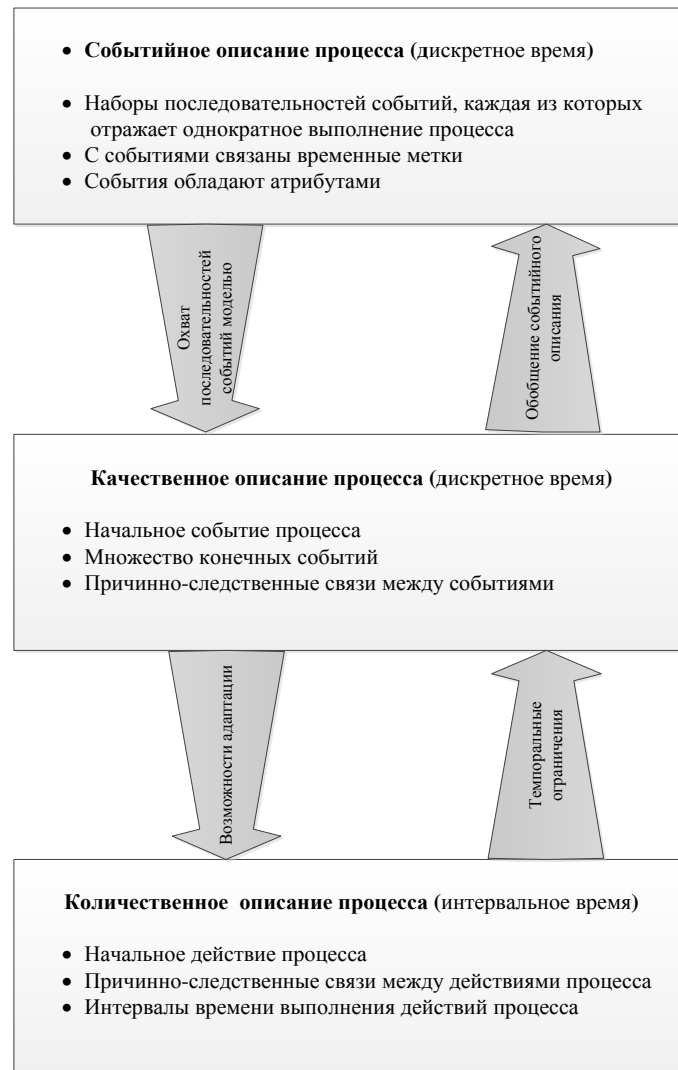


Рис. 1. Уровни описания процесса решения задачи как элемента прецедента

Поскольку трасса процесса  $\pi_k$  фиксирует последовательность выполнения действий процесса во времени, то для любой пары событий  $(e_{k,i}, e_{k,i+1})$ , связанных отношением переходов выполняется условие:

$$e_{k,i} \rightarrow e_{k,i+1} \vee e_{k,i+1} \rightarrow e_{k,i} \quad (3)$$

Следовательно, данное отношение является антирефлексивным. По этой же причине отношение перехода  $\rightarrow$  обладает свойством транзитивности:

$$e_{k,i} \rightarrow e_{k,i+1} \rightarrow e_{k,i+2} \equiv e_{k,i} F^+ e_{k,i+2}, \quad (4)$$

где модальный оператор  $F^+$  отображает транзитивное замыкание отношения переходов.

Оператор  $F^+$  указывает на то, что в один из моментов в будущем обязательно возникнет событие  $e_{k,i+2}$  при условии истинности события  $e_{k,i}$  в текущий момент.

Правомочность использования оператора модальной логики для отображения транзитивного замыкания определяется тем, что трасса событий логически упорядочена по времени согласно (2).

Предлагаемый метод основан на объединении событийного, качественного и количественного представлений в единую модель процесса решения задачи и включает в себя следующие этапы:

Этап 1. Построение качественного представления процесса решения задачи на основе анализа логических методами process mining.

Этап 2. Формализация набора базовых операций над интервалами событий с целью получения интервальных оценок времени выполнения отдельных действий процесса.

Этап 3. Формализация набора операций на последовательностях интервалов событий, обеспечивающих слияние нескольких альтернативных вариантов решения задачи.

Этап 4. Формирование интервального представления процесса решения задачи в составе прецедента:

- формирование подмножеств событий, соответствующих одному действию процесса;
- определение временных интервалов для каждого действия (операции) процесса на основе сравнения временных меток соответствующих событий;
- определение отношений последовательного, параллельного выполнения между интервалами на основе отношений между соответствующими событиями, отображенных в полученной на этапе 1 модели.
- формирование алгоритма действий процесса с учетом полученных отношений между интервалами.

Этап 5. Уточнение интервальной модели с учетом атрибутов событий процесса:

- выделение подмножеств атрибутов, а также их значений, задающих необходимые и достаточные условия для перехода между действиями и интервалами ожидания процесса;
- уточнение модели путем выделения интервалов ожидания;
- уточнение модели путем объединения интервалов для отдельных действий процесса.

#### Выводы

Предложен метод формирования модели процесса решения задачи на основе анализа логов. Метод включает в себя этапы построения моделей процесса с событийным и интервальным представлением времени, а также уточнение интервальной модели с учетом атрибутов событий процесса.

Полученная в результате применения метода модель процесса решения задачи обладает возможностью адаптации с учетом ограничений предметной области и ожидаемых результатов. Оценка ограничений выполняется на базе интервального представления времени. Данная оценка позволяет выявить возможность повышения эффективности процесса путем распараллеливания его действий.

Адаптация выполняется путем удаления из модели трасс процесса (или фрагментов трасс), которые не удовлетворяют временным ограничениям, а также смещения фрагментов трасс на временной линии выполнения процесса. Выбор удаляемых трасс (или их фрагментов) осуществляется на основе качественного описания процесса.

#### Список использованной литературы

1. Richter M. M. Case-Based Reasoning. A Textbook [Text] / Michael M. Richter, Rosina O. Weber // Springer. — 2013. — 546 p.
2. Kolodner J. Case-Based Reasoning [Text] / J Kolodner // Magazin Kaufmann. San Mateo. 1993. — 386 p.
3. Carbonell, J.G. Learning by analogy: Formulating and generalizing plans from past experience. [Text] / R.S. Michalski, J.G. Carbonell, & T.M. Mitchell (Eds.) // Machine learning, an artificial intelligence approach (Vol. 1), Palo Alto, CA: Tioga Press. — 1983. — P. 137 — 162.
4. Van der Aalst W.M.P. Process Mining: Discovery, Conformance and Enhancement of Business Processes [Text] / W.M.P. Van der Aalst // Springer-Verlag, Berlin — 2011. — 352 p.
5. Чалый С.Ф. Разработка обобщенной процессной модели прецедента, метода его формирования и использования [Текст] / С.Ф. Чалый, И. В. Левыкин // Журнал управляющие системы и машины. — 2016. — №3 — С. 23 — 29.
6. Николайчук О.А. Прототип интеллектуальной системы для исследования технического состояния механических систем [Текст] / О.А.Николайчук, А.Ю. Юрин // Искусственный интеллект. — 2006. — №4 — С. 459 — 468.

7. Николайчук О.А. Применение прецедентного подхода для автоматизированной идентификации технического состояния деталей механических систем [Текст] / О.А. Николайчук., А.Ю. Юрин // Автоматизация и современные технологии. — 2009. — №5 — С.3 — 12.
8. Берман А.Ф. Интеллектуальная система для анализа отказов сложных технических систем [Текст] / А.Ф.Берман, О.А.Николайчук, А.И.Павлов, М.А. Грищенко, А.Ю.Юрин // Труды Тринадцатой национальной конференции по искусственному интеллекту с международным участием КИИ-2012 (16-20 сентября 2012 г., г. Белгород, Россия): Труды конференции. — М.: Физматлит. — 2012. — Т.3. — С. 146— 154.
9. Aamodt A., Plaza E. Case-Based Reasoning: Foundational issues, methodological variations, and system approaches[Text] / A. Aamodt, E.Plaza // AI Communications. 1994. IOS Press, Vol.7:1, P. 39— 59.
10. Van der Aalst W.M.P. Process Mining in the Large: A Tutorial. [Text] / W.M.P. Van der Aalst // In E. Zimnyi, editor, Business Intelligence (eBISS 2013), volume 172 of Lecture Notes in Business Information Processing, Springer-Verlag, Berlin, — 2014 — P. 33— 76.
11. Christian W. Gunther. Process Mining in Flexible Environments. [Text] / W. Gunther Christian // PhD thesis, Eindhoven University of Technology, Berlin — 2008. — 228 p.
12. Gunther Christian W. Fuzzy mining- adaptive process simplification based on multiperspective metrics [Text] / Christian W, Gunther W.M.P. Van der Aalst // In Gustavo Alonso, Peter Dadam, and Michael Rosemann, editors, Business Process Management, volume 4714 of LNCS, — 2007. — P 328 — 343.