

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

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

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

МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЖУРНАЛОВ РЕГИСТРАЦИИ СОБЫТИЙ В ЗАДАЧАХ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ПРОЦЕССОВ

Предложены модели представления журналов регистрации событий, которые формируются корпоративной информационной системой при функционировании предприятия. При процессном управлении формируемые журналы регистрации событий отражают выполнение бизнес-процессов предприятия, а при функциональном управлении – реализацию отдельных функций подразделениями предприятия. Предложенные модели журналов создают возможность построения моделей бизнес-процессов методами process mining на основе обработки неструктурированных журналов регистрации событий, соответствующих функциональному подходу к управлению предприятием.

Запропоновано моделі подання журналів реєстрації подій, які формуються корпоративною інформаційною системою під час функціонування підприємства. У процесному управлінні формуються журнали реєстрації подій, що відображають виконання бізнес-процесів підприємства, а під час функціонального управління – реалізацію окремих функцій підрозділами підприємства. Запропоновані моделі журналів створюють можливість побудови моделей бізнес-процесів методами process mining на основі обробки неструктурованих журналів реєстрації подій, що відповідають функціональному підходу до управління підприємством.

The model of event logs, which are forming by corporate information system while enterprise functioning, viewed in article. In the process management generating event logs reflects the performance of business processes, whereas functional management reflects implementation of specific functions of the business units. The models of logs, which was propose in the article, creates the possibility of constructing business process models and process mining techniques on the basis of an unstructured event logs, which corresponding to the functional approach to business management.

Ключевые слова. Бизнес-процесс, модель журнала регистрации событий, отношения порядка.

Введение. В настоящее время активно развиваются методы интеллектуального анализа бизнес-процессов (БП), направленные на построение моделей таких процессов на основе анализа их журналов регистрации событий (ЖРС). Генерация журналов осуществляется соответствующей информационной системой. В журналах регистрации событий фиксируются события,

© В. М. Левыкин, С. Ф. Чалый, Е. О. Богатов, 2013

отражающие последовательность действий при выполнении бизнес-процессов или отдельных функциональных задач.

Методы интеллектуального анализа процессов (process mining) позволяют построить “черновую” модель бизнес-процессов (или уточнить существующую), выявить потенциал для оптимизации и “узкие места” бизнес-процессов. Результаты process mining дают возможность организовать эффективный контроль выполнения бизнес-процессов и при необходимости их реинжиниринг.

Однако следует отметить, что методы интеллектуального анализа процессов выдвигают ряд требований к структуре журналов регистрации событий, которые обычно выполняются только процессно-ориентированными информационными системами. В то же время существует значительное количество организаций, в которых бизнес-процессы слабо структурированы либо не определены. При этом обычно используются корпоративные информационные системы, ориентированные на функциональное управление, что вызывает трудности в описании и построении моделей существующих процессов. Текущие действия в данных информационных системах протоколируются с метками времени вне зависимости от того, к какому процессу и объекту принадлежит фиксируемое действие. Иными словами, регистрационные файлы в таких системах содержат информацию, необходимую для построения процессной модели, отражающей реально выполняющиеся бизнес-процессы. Но указанная в данных журналах информация не упорядочена в соответствии с существующими требованиями в process mining. Это, в свою очередь, значительно затрудняет процесс анализа.

Вышеизложенное определяет актуальность разработки моделей журналов регистрационного анализа данных.

Вопросы формирования структуры журнального файла как таковой рассмотрены в ряде трудов [1–4]. В статье [1] приведен анализ структуры журнальных файлов и обоснована достаточность вошедших в него элементов, а также записей для полноценного описания характеристик бизнес-процесса, таких как покрытие всех имеющихся трасс, отображающих все совокупности объектов БП, рассмотрено [2] место журнальных файлов в технологии интеллектуального анализа процессов, также приведены их примеры. В статье [3] основное внимание уделено структуре файлов MXML-формата, используемых для формализованного описания журнальных файлов и их последующего анализа. Также в статье [4] определен минимальный набор элементов для описания бизнес-процессов. Таким образом, в проведенных исследованиях основное внимание уделяется обоснованию и описанию структурных элементов журналов регистрации событий в процессных системах, однако не рассматриваются отношения упорядоченности между этими элементами. Также не рассматриваются с позиций интеллектуального анализа процессов структура и взаимосвязи между элементами журнальных файлов информационных систем, которые не ориентированы на процессное управление.

Постановка задачи. Цель заключается в формировании модели журнала регистрации событий для систем, которые не ориентированы на процессное управление и потому не структурируют ЖРС в формате последовательностей событий, относящихся к различным ситуациям. Последние отражают подпроцессы или бизнес-процессы полностью.

Исходными данными задачи являются:

- журнал регистрации событий процессно-ориентированной системы, содержащий информацию о частично-упорядоченной последовательности событий бизнес-процесса на основе процедур данного процесса и набора правил его исполнения;
- журнал регистрации событий, который содержит в общем случае неупорядоченное описание событий, происходящих на объекте управления с указанием меток времени.

Необходимо выявить ключевые взаимосвязи ЖРС процессно-ориентированной системы и на их основе выделить структурные составляющие и отношения порядка в модели ЖРС системы, ориентированной на традиционное функциональное управление.

Результаты исследования. В результате анализа бизнес-процессов с позиций process mining было выявлено несколько уровней их описания:

- уровень бизнес-правил;
- уровень процесса;
- уровень объектов, с которыми оперирует процесс.

Уровень бизнес-правил подразумевает представление бизнес-процессов в виде наборов зависимостей, отражающих взаимосвязи между его процедурами. В данном случае бизнес-правила задают ограничения на протекание БП и определяют порядок выполнения его действий. На уровне процесса рассматривается непосредственно поток работ. Способ отображения потока работ зависит от применяемого стандарта (нотации) БП. На уровне объектов изображают объекты БП в той последовательности, в которой они формируются (обрабатываются) процессом. Под объектами БП следует понимать объекты реального мира, которые используются процедурами процесса.

Журнальный файл, либо файл регистрации событий, чаще всего представляет собой один или несколько текстовых файлов, сформированных информационной системой (механизмами или службами системы) в результате мониторинга наступающих в процессе её функционирования событий. Журналирование позволяет представить все уровни рассматриваемого бизнес-процесса и связи между ними.

Одним из требований к любой информационной системе является поддержка процесса журналирования (регистрации) событий, в ряде систем сбор данных может производиться в неявном виде (в качестве фонового процесса) и быть доступным лишь администратору системы.

Жизненный цикл БП [5] включает в себя процессы формирования журнальных файлов (рис. 1).

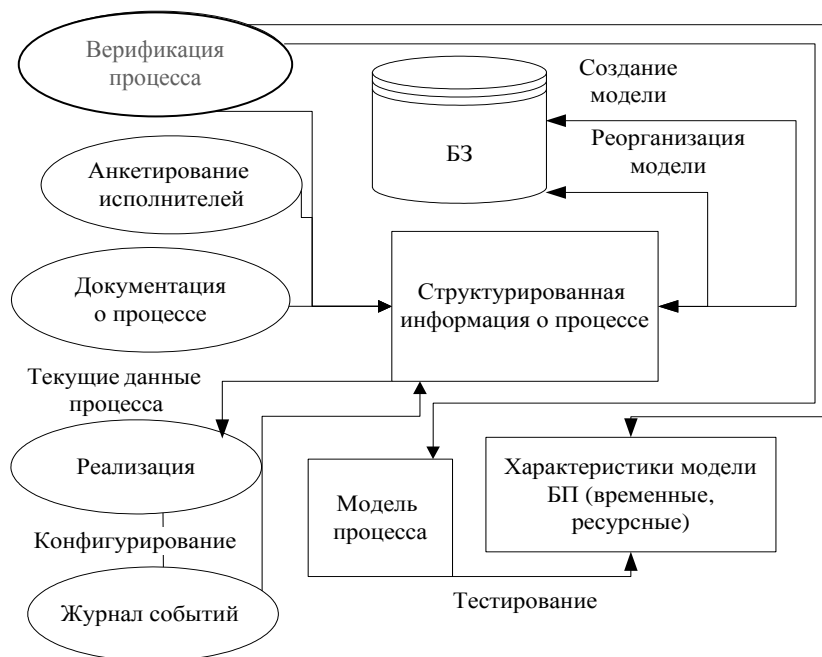


Рис. 1. Жизненный цикл БП

При построении журнальных файлов руководствуются следующими требованиями:

- отображаемые события являются атомарными;
- всегда отображаются временные показатели события;
- всегда отображаются ресурсы, связанные с событием.

На сегодняшний день выделяют следующие представления журнала событий:

- текстовый файл;
- XES-файл, файл MXML формата;
- таблица БД.

Структуры журнальных файлов могут быть представлены в следующих формах: в соответствии со стандартом OpenXES; в виде текстового файла; в виде таблицы БД представлены на рис. 2–4 [2]. На рис. 2 выделены блоки: начало процедуры и событие, входящее в процедуру. Отметим, что на рис. 3 в виде текстового файла представлен журнал регистрации событий системы, ориентированной на функциональное управление.

```

<global scope="event">
<string key="concept: name" value="INVALID"/>
  <string key="lifecycle: transition" value="complete"/></global>
<classifier name="MXML Legacy Classifier" keys="concept: name lifecycle: transition"/>
<classifier name="Event Name" keys="concept: name"/>
<classifier name="Resource" keys="org: resource"/>
<string key="source" value="Rapid Synthesizer"/>
<string key="concept: name" value="bigger-example.mxml">
<string key="lifecycle:model" value="standard"

```

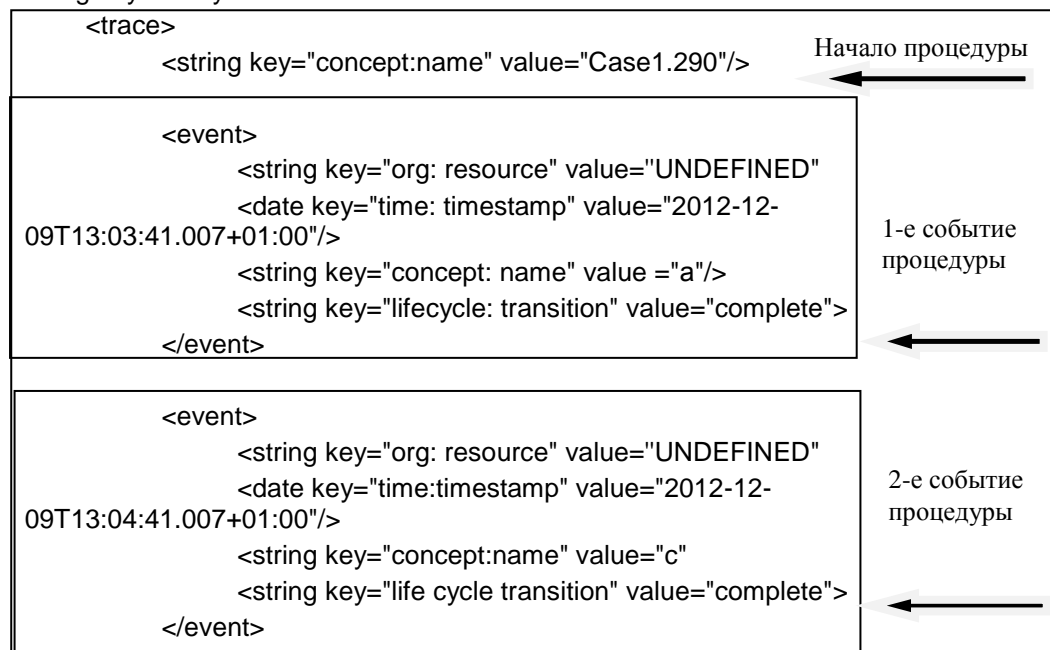


Рис. 2. Пример структуры журнального файла согласно стандарту OpenXES

12.01.2004;09:59:37;;конфигуратор;Подключение;Сеанс;0;НОВЫЙ СЕАНС : Компьютер 111-D29B45CA
 12.01.2004;10:02:03;;конфигуратор;Отключение;Сеанс;0;;;;
 12.01.2004;10:06:59;;1С:Предприятие;Подключение;Сеанс;0;НОВЫЙ СЕАНС : Компьютер 111-D29B45
 12.01.2004;10:07:00;;1С:Предприятие;Изменена точка актуальности;Изменение данных;2;Вперед :
 12.01.2004;10:13:07;;1С:Предприятие;Подключение;Сеанс;0;Компьютер 111-D29B45CA0D0(m);;;
 12.01.2004;10:13:08;;1С:Предприятие;Отключение;Сеанс;0;;;;
 12.01.2004;10:19:36;;конфигуратор;Подключение;Сеанс;0;Компьютер 111-D29B45CA0D0(m);;;
 12.01.2004;10:20:17;;конфигуратор;Отключение;Сеанс;0;;;;
 12.01.2004;10:22:57;;1С:Предприятие;Отключение;Сеанс;0;;;;
 13.01.2004;11:24:34;;1С:Предприятие;Подключение;Сеанс;0;НОВЫЙ СЕАНС : Компьютер 111-D29B45
 13.01.2004;11:25:13;;1С:Предприятие;Документ создан;Информация;3;;0/1254/0;Перемещение 000
 13.01.2004;11:25:17;;1С:Предприятие;Документ не записан;Информация;3;;0/1254/0;Перемещение
 13.01.2004;11:25:28;;1С:Предприятие;Отключение;Сеанс;0;;;;
 13.01.2004;11:25:33;;конфигуратор;Подключение;Сеанс;0;НОВЫЙ СЕАНС : Компьютер 111-D29B45CA
 13.01.2004;11:28:47;;конфигуратор;Отключение;Сеанс;0;;;;

Рис. 3. Представление журнального файла в текстовом формате
(процессно-ориентированная система)

	A	B	C	D	G	H	I
1	Service ID	Operation	Start Date	End Date	Product	Service Type	Agent
2	Case 1	Inbound Call	9.3.10 8:05	9.3.10 8:10	MacBook Pro	Referred to Servicer	Helen
3	Case 1	Handle Case	11.3.10 10:30	11.3.10 10:32	MacBook Pro	Referred to Servicer	Helen
4	Case 1	Call Outbound	11.3.10 11:45	11.3.10 11:52	MacBook Pro	Referred to Servicer	Henk
5	Case 2	Inbound Call	4.3.10 11:43	4.3.10 11:46	MacBook Pro	Referred to Servicer	Susi
6	Case 3	Inbound Call	25.3.10 9:32	25.3.10 9:33	MacBook Pro	Referred to Servicer	Mary
7	Case 4	Inbound Call	6.3.10 11:41	6.3.10 11:51	iPhone	Referred to Servicer	Fred
8	Case 5	Inbound Call	18.3.10 10:54	18.3.10 11:01	MacBook Pro	Product Assistance	Kenny
9	Case 6	Inbound Call	25.3.10 17:09	25.3.10 17:13	MacBook Pro	Referred to Servicer	Harold

Рис. 4. Представление журнального файла в форме таблицы БД

Таким образом, можно выделить следующие схожие элементы большинства журналов регистрации событий процессно-ориентированных систем [3]:

- наименование процесса;
- возникшая ситуация, которая обычно определяется на основе обрабатываемого в рамках процесса объекта (часто таким объектом является документ);
- действие в рамках возникшей ситуации (например, оплата сырья, поставка комплектующих);
- событие, определяющее начало, окончание действия либо его промежуточные фазы;
- временные параметры события.

На основе проведенного анализа [6] можно утверждать, что журналы регистрации событий структурированного бизнес-процесса включают следующие обязательные элементы:

- идентификатор последовательности выполнения процедур в определенной ситуации;
- временной параметр;
- событие, отражающее действие процесса;
- объект, с которым выполняется данное действие процесса;
- исполнитель действия.

Перечисленные выше атрибуты являются обязательными, но также журнальные файлы могут содержать ряд дополнительных атрибутов: стоимость, затраченные ресурсы и др.

Таким образом, журнальные файлы процессно-ориентированных систем можно описать в виде наборов событий G_{B_p} , каждый из которых отражает упорядоченное во времени подмножество действий, происходивших в i -й ситуации c_i :

$$G_{B_p} = c_i \cdot \quad (1)$$

Для каждой ситуации бизнес-процесса задается отношение порядка относительно времени \prod_t для каждой пары соседних элементов (e_{t_k}, e_{t_l}) :

$$c_i = E_j, E_j \subseteq E, \forall e_{t_k}, e_{t_l} \in E_j \exists \prod_t (e_{t_k}, e_{t_l}). \quad (2)$$

где E – множество событий в ЖРС;

e_{t_k}, e_{t_l} – события, которые произошли (начали выполняться) в моменты времени t_k и t_l .

Данное отношение порядка определяет, что между событиями e_{t_k}, e_{t_l} на шкале времени в ЖРС не существует промежуточных событий:

$$\begin{aligned} (\forall e_{t_k}, e_{t_l} \in E_j) (\neg \exists e_{t_m} \in E_j) ((t_k < t_l) \wedge (t_l > t_m) \vee (t_m > t_k)) \\ \Rightarrow \prod_t (e_{t_k}, e_{t_l}). \end{aligned} \quad (3)$$

В свою очередь, каждое событие процесса e_n характеризуется следующим набором параметров:

$$e_n = t_n^*, t_n^{\ddot{}}, obj_s, isp_p, \quad (4)$$

где $t_n^*, t_n^{\ddot{}}$ – моменты времени начала и завершения события e_n ;

obj_s – объект, с которым работает процедура процесса, отраженная в ЖРС в виде события e_n ;

isp_p – исполнитель процедуры процесса, отраженной в ЖРС в виде события e_n .

Отметим, что выражения (1–4) определяют модель журнала регистрации событий процессно-ориентированной системы. В такой модели ключевым является временной порядок событий в рамках ситуаций. Связи с организационной структурой и обрабатываемыми объектами рассматриваются как неявно влияющие на временной порядок и отражаемые им.

Между исполнителями действий БП существует отношение порядка $\prod_{Os} (isp_p, isp_r)$, отражающее иерархическую организационную структуру предприятия Os и задающее отношение “начальник – подчиненный” для каждой пары исполнителей (isp_p, isp_r) .

Следовательно, для событий e_n можно задать отношение порядка на основе организационной структуры предприятия.

Выделение данного отношения порядка позволяет отобразить бизнес-процессы (или их фрагменты, описываемые в ЖРС ситуациями) на организационную структуру предприятия. Более того, данное отношение оказывает непосредственное влияние на порядок действий бизнес-процесса и отражающих их событий в ЖРС.

Анализ существующих бизнес-процессов, а также журналов регистрации событий показывает, что объекты, с которыми оперирует процесс, семантически связаны между собой. Семантика таких связей влияет на порядок их обработки. Иными словами, для событий e_n можно определить отношение порядка на основе обрабатываемых объектов $\prod_{Obj}(e_k, e_l)$. Данное отношение порядка задает ограничения на последовательность обработ-

ки объектов и, следовательно, на последовательность событий бизнес-процесса.

Отношения порядка $\prod_{Obj}(e_k, e_l)$ и $\prod_{Os}(e_k, e_l)$ позволяют определить G_{F_p} – модель ЖРС, формируемого информационной системой, ориентированной на традиционный подход к управлению:

$$G_{F_p} = \{E_j\}, E_j \subseteq E, E_j = \{e_n\},$$

$$e_n = (t_n^*, t_n^{(**)}, obj_s, isp_p)(\forall e_k, e_l \in E_j)(\exists \prod_{obj, Os, t}(e_k, e_l)), \quad (5)$$

где $\prod_{obj, Os, t}(e_k, e_l)$ – отношение порядка на множестве объектов, на множестве исполнителей и по времени.

Существует несколько вариантов детализации модели ЖРС (5), которые зависят от способа задания порядка на отношениях $\prod_{Obj}(e_{t_k}, e_{t_1})$, $\prod_{Os}(e_{t_k}, e_{t_1})$ и $\prod_t(e_{t_k}, e_{t_1})$, а также выбора подмножеств объектов и исполнителей. Так, упорядоченность по объектам, а затем по времени позволяет сформировать ситуации (и выделить подмножества E_j), связанные с последовательными цепочками действий по обработке отдельных объектов без учета организационной структуры предприятия.

Упорядочивание на основе $\prod_{Os}(e_{t_k}, e_{t_1})$, а затем по времени позволяет осуществить привязку полученных в результате интеллектуального анализа процессов моделей БП к существующей организационной структуре. При этом появляется возможность выстроить иерархию подмножеств E_j на основе семантики “начальник – подчиненный”. Более того, набор подмножеств E_j , соответствующий начальнику, будет отражать процессы, которые выполняются в подчиненных ему подразделениях.

Отбор событий E_j по критерию пересечения связанных с ними множеств объектов и исполнителей создает условия для решения задачи выделения бизнес-процессов по границам подразделений. Решение этой задачи позволяет на уровне подразделений устранить противоречие между “горизонтальной”, “сквозной” структурой бизнес-процессов и

вертикальной иерархией подразделений. Очевидно, что отобранные таким образом события необходимо упорядочить одним из рассмотренных способов.

Выводы. Предложены модели журналов регистрации событий для процессно-ориентированных и функционально-ориентированных информационных систем. Модель ЖРС для процессно-ориентированной информационной системы отличается от существующих содержанием отношений порядка относительно времени для событий, принадлежащих каждой ситуации, входящей в состав журнала, что позволяет формировать модель бизнес-процесса как совокупность параллельно выполняющихся ситуаций на заданном интервале времени и тем самым упростить построение результирующей модели БП методами интеллектуального анализа процессов.

Модель ЖРС для информационной системы, ориентированной на традиционное функциональное управление, включает в себя набор подмножеств событий, отражающих выполнение функциональных задач, а также отношения порядка на основе связей между обрабатываемыми процессом объектами, иерархических связей исполнителей и временной последовательности событий. Данная модель создает возможности для структуризации журнала по ситуациям бизнес-процесса с использованием заданных отношений, что позволяет построить модели БП для организаций с функциональным управлением методами интеллектуального анализа процессов.

В практическом аспекте применение полученных моделей в задачах интеллектуального анализа процессов создает условия для дальнейшего реинжиниринга и совершенствования БП с учетом фактически выполняющихся в рамках процесса последовательностей действий.

Литература

1. Hedong Yang Estimating Completeness of Event Logs / Hedong Yang, B. F. van Dongen, Arthur H.M. ter Hofstede // Tsinghua University. – Beijing. – 2010. – 27 с.
2. W. M. P. van der Aalst Workflow Mining: Discovering Process Models from Event Logs / A. J. M. M. Weijters, L. Maruster // IEEE Transactions on Knowledge and Data Engineering 16 (9). – Eindhoven : Department of Technology Management, Eindhoven University of Technology, 2004. – P. 1128–1142.
3. Saravanan M. S. Application of Mining Algorithms using ProM and Weka Tools / M. S. Saravanan, Dr. R. J. Rama Sree // International Journal of Computer Science and Telecommunications. – October 2011. – Vol. 9. – No. 10. – P. 331–337.
4. Чалый С. Ф. Використання ієрархії шаблонів у process mining / С. Ф. Чалый, Є. О. Богатов, Альшейх Алі Джаміль // Системи обробки інформації. – 2010. – Вип. № 6 (87). – С. 257–259.
5. Чалый С. Ф. Верифікація бізнес-правил на основі журналу подій / С. Ф. Чалый, Є. О. Богатов, І. Б. Буцукина // Системи обробки інформації. – 2010. – Вип. № 7 (88). – С. 174–176.
6. Метод предварительной обработки журналов регистрации событий в задачах интеллектуального анализа процессов / В. М. Левыкин, С. Ф. Чалый, Е. О. Богатов, М. И. Дикусар // Системи обробки інформації. – 2011. – № 4 (20). – С. 203–207.