

УДК 004.891.3

О.В. Чалая¹¹ХНУРЕ, г. Харьков, Украина, oksana.chala@nure.ua

ПРИНЦИП ТА МЕТОД ЕВОЛЮЦІЙНОЇ ПОБУДОВИ БАЗИ ЗНАНЬ НА ОСНОВІ АНАЛІЗУ ЛОГІВ ІС ПРОЦЕСНОГО УПРАВЛІННЯ

Запропоновано принцип еволюційного формування бази знань на основі аналізу логів інформаційної системи процесного управління. Принцип базується на поєднанні водоспадної та еволюційної стратегій та передбачає використання як формалізованих знань, так і екстерналізованих персональних знань виконавців для побудови бази знань. Формалізовані знання включають до складу бази знань при проектуванні або проведенні реінжиніринга бізнес-процесу. Екстерналізація знань виконавців виконується на основі аналізу логів, що містять у собі запис поведінки бізнес-процесу. Використання даного принципу дає можливість підвищити ефективність процесного управління за рахунок включення до бази знань персональних знань виконавців бізнес-процесу. На основі розробленого принципу запропоновано метод автоматизованої побудови бази знань інформаційної системи процесного управління шляхом виявлення зв'язків між атрибутами артефактів бізнес-процесу, а також значеннями цих атрибутів. Метод дає можливість на основі аналізу поведінки процесу отримати залежності між артефактами, що задають обмеження та правила вибору дій процесу, та включити ці залежності до складу бази знань.

ЗНАННЯ, БАЗА ЗНАНЬ, ЗНАННЯ-ЄМНИЙ БІЗНЕС-ПРОЦЕС, КОНТЕКСТ, АРТЕФАКТ

Вступ

Інформаційні системи (ІС) процесного управління потребують структуризації діяльності підприємства у вигляді послідовностей дій з виробництва продукції та послуг, а також побудови моделей множини бізнес-процесів (БП) організації. Кожен з бізнес-процесів описує послідовність дій, необхідну для вирішення відповідної функціональної задачі без врахування зв'язку з організаційною структурою підприємства [1].

Процесне управління підприємством передбачає управління бізнес-процесами з використанням їх моделей. Традиційно модель бізнес-процесу має вигляд workflow - графа [2]. При використанні процесного підходу реалізується управління по відхиленню. Необхідною умовою такого управління є адекватність моделей бізнес-процесів в умовах постійних змін вимог ринку, складу виконавців, замін ресурсів процесу.

Підтримка адекватності моделі для класу знання-ємних бізнес-процесів (ЗБП) пов'язана зі значними труднощами, тому що такі процеси можуть змінювати задану в моделі послідовність дій на основі контекстно-орієнтованих рішень виконавців [3, 4]. Це приводить до виникнення невідповідності між моделлю та бізнес-процесом, що виконується.

При ухваленні рішень про зміну ходу виконання бізнес-процесу виконавці використовують як загальнодоступні явні правила, так і персональні знання і досвід [4, 5]. Персональні знання зазвичай містять у собі неформалізовані правила та шаблони дій і тому не можуть бути апіорно включені в модель бізнес-процесу при її проектуванні.

Таким чином, при реалізації управління гнучкими бізнес-процесам, що адаптуються на основі знань, актуальною є проблема автоматизованої

побудови бази знань для забезпечення адекватності моделей гнучких процесів.

Для вирішення даної проблеми необхідно, по-перше, представити знання бізнес-процесу окремо від його традиційної workflow — моделі. По-друге, по мірі удосконалення ЗБП, необхідно проводити екстерналізацію знань виконавців — тобто виділити знання і представити їх у формальному вигляді, після чого доповнити базу знань інформаційної системи.

Задачі екстерналізації і формування бази знань вимагають розробки відповідних формальних принципів і підходів, що базуються на аналізі поведінки таких процесів. Можливість виділення розглянутих знань забезпечується тим, що до складу інформаційно-управляючої системи зазвичай входить модуль запису логів. Тому у системах процесного управління сліди застосування неявних знань виконавців відображаються у файлах логів бізнес-процесів. Це забезпечує умови для виявлення і подальшої формалізації неявних залежностей на основі аналізу логів БП методами інтелектуального аналізу процесів (process mining) [6].

Після перевірки виявлених залежностей ОПР вони можуть бути включені до складу бази знань інформаційно-управляючої системи. Проте питання побудови баз знань інформаційних систем процесного управління на основі аналізу логів процесів на сьогодні розроблені недостатньо. Зазначене свідчить про актуальність теми роботи.

1. Постановка задачі

Метою статті є розробка принципу та методу еволюційної побудови бази знань інформаційної системи процесного управління шляхом аналізу логів та подальшого виявлення закономірностей, що визначають обмеження та можливості вибору дій бізнес-процесу у конкретному контексті.

2. Операції з побудови представлення знань для ІС процесного управління

Гнучкі бізнес-процеси, що змінюються внаслідок використання знань виконавців, доцільно розділити на три складові [7,8]:

- контекст процесу;
- набір патернів, правил і обмежень, що визначають вибір дій в заданому контексті, з урахуванням поточного стану його елементів;
- набір послідовностей дій бізнес-процесу у вигляді його workflow — представлення (ланцюжків робіт).

Контекст містить у собі набір об'єктів предметної області, що використовуються бізнес-процесом, або впливають на його виконання. Такі об'єкти зазвичай іменуються артефактами БП. Стан контексту визначається станом артефактів, а також поточними взаємозв'язками між ними [9,10].

У відповідності до наведеної структури БП, представлення знань *Kr* ІС процесного управління повинно містити у собі як статичні залежності між атрибутами артефактів, так і змінні залежності між значеннями цих атрибутів. Перші доцільно розглядати як обмеження на виконання процесу, а другі — як правила вибору дій.

Наприклад, повторення атрибутів «Роль виконавця» і «Найменування операції» для всіх подій у вхідних даних свідчить об існування зв'язку між артефактами «Роль» і «Операція» у рамках опису бізнес-процесу. Тобто при побудові моделі бізнес-процесу для кожної дії необхідно визначити роль виконавця.

З іншого боку, зв'язки між найменуваннями операцій і конкретною роллю, а також ім'ям виконавця вказують на можливість вибору конкретного працівника на певну роль при виконанні заданої операції бізнес-процесу.

Слід зазначити, що окремі дії процесу також можуть бути представлені як артефакти, які використовує процес. Така типізація дозволяє єдиним чином представити як залежності контексту, так і правила вибору дій процесу.

У відповідності до наведених особливостей використання знань при процесному управлінні, представлення знань складається з наступних елементів:

$$Kr = \{Af, Rl, Cn, Rl^*, Cn^*\}, \quad (1)$$

де *Cn* — залежності, що визначають обмеження на виконання бізнес-процесу; *Rl* — правила, що задають вибір конкретних артефактів (наприклад, виконавців, дій) при виконанні процесу, *Rl** — інтегровані правила; *Cn** — інтегровані обмеження.

Обмеження та правила задаються через значення атрибутів артефактів:

$$\begin{aligned} Cn_d &\equiv a_{ij} \Rightarrow a_{lk}, \\ Rl_z &\equiv v_{ij}^s \Rightarrow v_{lk}^n | \\ \forall v_{ij}^s \exists t_{sm} \wedge \forall v_{lk}^n \exists t_{nm}, t_{sm}, t_{nm} &\in T \end{aligned} \quad (2)$$

де t_{sm} — мітка часу події логу, в описі якої вказано значення v_{ij}^s атрибуту a_{ij} ; t_{nm} — мітка часу події логу, в описі якої вказано значення v_{lk}^n атрибуту a_{lk} .

Інтегровані правила та обмеження об'єднують кілька залежностей між атрибутами артефакту.

Для формування наведеного представлення знань *Kr* використовуються наступні операції:

- виявлення атрибутів артефактів на основі аналізу подій логу інформаційної системи;
- екстерналізація контекстно-орієнтованих знань про управління бізнес-процесом у вигляді залежностей між артефактами і їх атрибутами;
- установка пріоритетів для залежностей в базі знань;
- ітеративне узагальнення та доповнення залежностей.

Представлений вибір операцій базується як на особливостях вхідних даних, представлених у формі логу процесу, так і на ітеративному характері процесного управління: після завершення бізнес-процесу виконується аналіз його результатів та, при необхідності, коригування моделі.

Знову процес запускається вже після уточнення моделі.

Операція виявлення атрибутів дозволяє встановити зв'язок між вхідними атрибутами подій логу і вихідними артефактами та їх атрибутами. У разі нестандартних логів може знадобитися введення додаткової інформації або уточнення наявних атрибутів безпосередньо користувачем.

Наприклад, в описі подій початкового логу може бути вказана роль виконавця «*org:role*» або оброблюваний об'єкт «*product*» в закодованому виді, як показано на рис. 1.

```
<string key=>org:role value=>A2_1/>,  
<string key=>product value=>PROD542/>
```

Рис. 1. Приклад неповної початкової інформації

У цьому прикладі семантика «*A2_1*», «*PROD542*» абсолютно неочевидна, що вимагає введення додаткової інформації для розшифровки приведених кодів.

Операція узагальнення знань про предметну область полягає в інтеграції залежностей, які є істинними у рамках конкретного домену. Приклад інтегрованої залежності наведений на рис.2.

ВИКОНАВЕЦЬ.Роль.І'мя
зазвичай виконує операцію
ОПЕРАЦІЯ.Найменування
бізнес-процесу
ПРОЦЕС.Назва.

Рис. 2. Приклад інтегральної залежності

У цьому ілюстративному прикладі представлені об'єкти *ВИКОНАВЕЦЬ*, *ОПЕРАЦІЯ*, *ПРОЦЕС*, між атрибутами яких існує загальна залежність.

Ця залежність дозволяє доповнювати процесну модель у разі неповноти початкових даних. Так, якщо в описі подій конкретного логу, або подій

окремої траси логу пропущені атрибути з ім'ям виконавця, то інформація про виконавця відсутня в моделі процесу. Тоді приведене загальне правило дозволяє приймати рішення по виконанню *ОПЕРАЦІЯ.Найменування*, призначенню виконавцем *ВИКОНАВЕЦЬ.Ім'я*, що виконує роль *ВИКОНАВЕЦЬ.Роль*.

Операція установки пріоритетів для розміщених у базі знань залежностей реалізується через визначення вагових відношень між артефактами і їх атрибутами. У загальному випадку для встановлення таких коефіцієнтів можна використати два підходи:

- значення коефіцієнтів для залежностей визначає користувач відповідно до свого розуміння важливості цього правила для відповідної задачі процесного управління;

- значення вагових коефіцієнтів задаються за допомогою одного з відомих методів навчання.

При рішенні завдань процесного управління доцільним представляється застосувати комбінацію обох підходів. Спочатку в якості навчальної вибірки використовується один або декілька логів. Навчання на основі трас одного логу дозволяє адаптувати отримані залежності до специфіки конкретного підприємства або підрозділу підприємства. Використання в якості навчальної вибірки декількох логів дозволяє узагальнити значення вагових коефіцієнтів для групи аналогічних бізнес-процесів в заданій предметній області.

Надалі результати навчання можуть бути відкориговані ЛПП з урахуванням персональних контекстно-залежних знань.

Операція ітеративного уточнення і доповнення залежностей використовується при рішенні наступних задач:

- усунення помилок і неточностей у виділених залежностях;

- підвищення адекватності формальних знань про бізнес-процеси у разі використання персональних знань і досвіду виконавців при управлінні бізнес-процесом.

Як було показано раніше, зміна ходу процесу в результаті застосування персональних знань фіксується у лозі. Кожна нова реалізація записується у вигляді окремої траси логу. Порівняння трас дозволяє виділити залежності між артефактами і їх атрибутами, які сформувалися внаслідок управлінських рішень виконавців. Виконання перших трьох операцій по формуванню представлення знань дозволяє доповнити БЗ і тим самим підвищити адекватність процесної моделі.

3. Метод побудови бази знань для процесного управління

У відповідності до традиційного життєвого циклу бізнес-процесу вимоги до БП відомі апріорно. Відповідно, структура процесу повністю, з

урахуванням усіх деталей, задається при його побудові/реінжинірингу і не змінюється під час його виконання.

При адаптації та удосконаленні таких процесів реалізується інкрементний підхід. Останній використовує модифікацію відомої водоспадної стратегії, що застосовується в області розробки програмного забезпечення. Удосконалення моделі процесу виконується після завершення однієї або декількох його реалізацій, за результатами оцінки ефективності БП.

Для знання-ємних процесів характерна зміна послідовності дій внаслідок використання неформалізованих знань виконавців. Використання виконавцями своїх знань та досвіду при управлінні бізнес-процесами зазвичай виникає після змін у контексті виконання дій БП. Останні виникають внаслідок неконтрольованого впливу зовнішнього середовища, зміни вимог до процесу, тощо.

В результаті застосування таких персональних знань реальний процес буде відрізнятися від його моделі, що потребує постійного поповнення бази знань.

Таке поповнення може бути реалізовано на основі запропонованого принципу еволюційної побудови бази знань для процесного управління.

Принцип еволюційної побудови бази знань заснований на поєднанні водоспадної і еволюційної стратегій. Еволюційна стратегія спрямована на ітеративне розширення можливостей процесного управління шляхом включення до бази знань нових залежностей, виявлених під час виконання процесу. Останні відображають невідомі раніше взаємозв'язку між елементами контексту і діями бізнес-процесу.

Запропонований принцип ґрунтується на тому, що лог інформаційної системи процесного управління поповнюється по мірі виконання бізнес-процесу. Кожна дія процесу фіксується у вигляді події логу. Сукупність подій для однієї реалізації процесу складає трасу. Порівняння поточної траси процесу з поточною моделлю процесу, а також з іншими трасами логу дає можливість виявити відхилення в ході процесу, яке виникло внаслідок застосування виконавцями своїх знань для зміни ходу процесу.

Після виявлення відхилення необхідно провести його аналіз, тобто виявити (доповнити) залежності між артефактами контексту, що можуть бути причиною для зміни ходу виконання БП.

Семантика еволюційної стратегії в рамках запропонованого принципу може бути відображена у вигляді наступного правила:

ЯКЩО

- 1) послідовність подій поточної траси є унікальною для всіх інших трас логу даного бізнес-процесу;

- 2) дана траса не відображена в моделі БП;

ТО

1) виділити унікальну підмножину подій траси логу;

2) на основі порівняння значень атрибутів відфільтрованої множини подій та існуючих трас логу виявити або уточнити обмеження та правила вибору дій бізнес-процесу;

3) на основі рішення ОПР доповнити базу знань.

Еволюційний характер принципу підтверджується тим, що наведене правило може бути виконано циклічно по мірі реалізації бізнес-процесу. При зміні останнього внаслідок врахування виконавцями нових вимог або зовнішніх впливів база знань буде постійно поповнюватись.

На основі запропонованого принципу розроблено метод побудови бази знань системи процесного управління.

Розроблений метод використовує операції виявлення атрибутів артефактів і зв'язків між ними та реалізує принцип інкрементного поповнення бази знань. Загальна ідея методу полягає у виявленні залежностей на різних рівнях бізнес-процесу:

- на рівні послідовності операцій бізнес-процесу виявляються залежності між подіями логу; події та операції

- на рівні контексту бізнес-процесу визначаються обмеження на виконання його дій;

- на рівні правил визначаються патерни вибору дій в залежності від поточного стану контексту.

Метод містить у собі наступні ключові етапи:

Етап 1. Виявлення артефактів та їх атрибутів на основі аналізу логу бізнес-процесу [11].

На даному етапі виявляються підмножини подій трас логу, між атрибутами яких можуть існувати залежності, а також значення атрибутів цих подій. Події логу фіксують виконання відповідних дій бізнес-процесу. В файлі логу при описі подій вказуються атрибути артефактів бізнес-процесу, які були використані при виконанні відповідних дій. Тобто опис подій логу містить інформацію не лише про виконані дії, а й про контекст цих дій. Це дозволяє визначити відповідні артефакти та значення їх атрибутів з мітками часу.

Етап 2. Побудова правил та обмежень по управлінню бізнес-процесами з використанням відомих формалізованих знань. Як було показано раніше, при управлінні бізнес-процесами використовуються як документовані та формалізовані знання, так і персональні знання виконавців. На даному етапі побудови БЗ використовуються формалізовані знання і тому реалізуються традиційні підходи інженерії знань.

Етап 3. Побудова прототипів правил та обмежень за результатами аналізу логів. Задача даного етапу полягає в попередньому виявленні залежностей, що відображають персональні знання виконавців, які були використані при адаптації/удосконаленні бізнес-процесу у відповідності до

поточного стану контексту. Побудова прототипів виконується на основі фактор-графу. Фактор-граф дозволяє визначити взаємозв'язки між декількома атрибутами артефактів (значеннями цих атрибутів). Приклади зв'язків: послідовний зв'язок, ідентичні атрибути, тощо. Змінні, що відповідають вершинам фактор-графу можуть бути двох типів: з відомим значенням; зі значенням, яке потрібно знайти. Для останніх визначається умовна ймовірність кожного кандидата.

Результуючий граф із визначеними умовними ймовірностями є прототипом правил та обмежень по управлінню бізнес-процесами.

Етап 4. Визначення вагових коефіцієнтів для отриманих на етапі 3 правил та обмежень. При обчисленні вагових коефіцієнтів враховується кількість появ правила в трасах логу.

Етап 5. Відбір отриманих правил ОПР. Результатом даного етапу є впорядкований за ваговими коефіцієнтами набір правил вибору дій та/або обмежень на допустимі дії, які в подальшому можуть бути вносені в базу знань.

Наведений метод дозволяє знаходити правила на різних рівнях знання-ємого бізнес-процесу в залежності від підмножини вхідних даних.

В тому випадку, якщо використовуються події логу зі вказаними в атрибутах назвами відповідних дій процесу, то результатами методу є залежності між діями на рівні workflow бізнес-процесу.

При використанні назв атрибутів артефактів результатом методу є залежності між артефактами контексту, які доцільно розглядати як обмеження на виконання бізнес-процесу.

Якщо у якості вхідних даних використовуються значення атрибутів артефактів, то результатом методу є контекстно-орієнтовані правила вибору дій бізнес-процесу.

Висновок

Запропоновано принцип еволюційного формування бази знань на основі аналізу логів інформаційної системи процесного управління. Принцип базується на поєднанні водоспадної та еволюційної стратегій та передбачає використання як формалізованих знань, так і екстерналізованих персональних знань виконавців для побудови БЗ.

Формалізовані знання включають до складу бази знань при проектуванні або реінжинірингу бізнес-процесу. Екстерналізація знань виконавців виконується на основі аналізу логів, що містять у собі запис поведінки БП.

Використання даного принципу дає можливість підвищити ефективність процесного управління за рахунок включення до бази знань персональних знань виконавців бізнес-процесу.

На основі розробленого принципу запропоновано метод автоматизованої побудови бази знань ІС процесного управління шляхом виявлення

зв'язків між атрибутами артефактів бізнес-процесу, а також значеннями цих атрибутів. Метод передбачає визначення умовної ймовірності значень атрибутів, що входять до складу патерна або правила, а також визначення вагових коефіцієнтів отриманих патернів. Метод дає можливість на основі аналізу поведінки процесу отримати залежності між артефактами, що задають обмеження та правила вибору дій процесу, та включити ці залежності до складу бази знань.

Список літератури:

1. *Weske M.* Business Process Management: Concepts, Languages, Architectures. Second Edition/ M. Weske. — Springer, 2012. — 403 p.
2. *Vom Brocke, J.* Handbook on Business Process Management 1. Introduction, Methods, and Information Systems / J. vom Brocke, M. Rosemann. — Springer-Verlag Berlin Heidelberg, 2015. — 709 p.
3. *Gronau N.* KMDL-Capturing, Analysing and Improving Knowledge-Intensive Business Processes / N. Gronau, C. Müller, R. Korf // Journal of Universal Computer Science. - 2005. — №11(4). — pp. 452- 472.
4. *Gronau, N.* Modeling and Analyzing knowledge intensive business processes with KMDL: Comprehensive insights into theory and practice (English) / N. Gronau. — Gito, 2012. — 522 p.
5. *Polanyi M.* Personal Knowledge: Towards a Post-Critical Philosophy/ M. Polanyi. — University of Chicago Press, 1958. — 493 p.
6. *Van der Aalst, W. M. P.* Process Mining: Discovery, Conformance and Enhancement of Business Processes / W. M. P. Van der Aalst. — Springer Berlin Heidelberg, 2011. — 352 p.
7. *Чалая О.В.* Структуризация знание-емких бизнес-процессов // Системы управління навігації та зв'язку. — П.: ПолтНТУ, 2015. — Вип.2 (35) — С. 151-155.
8. *Левыкин В. М.* Модель знание-емкого бизнес-процесса для задач интеллектуального анализа процессов// В. М. Левыкин, О.В. Чалая // УСиМ, 2016, №6. — С. 59-66.
9. *Cohn D.* Business artifacts: A data-centric approach to modeling business operations and processes/ Cohn D., Hull R. // IEEE Data Eng. Bull. - 2009. — №32. — pp. 3-9.
10. *Bhattacharya K.* Artifact-centered operational modeling: Lessons from customer engagements/ K. Bhattacharya, N. S. Caswell, S. Kumaran, A. Nigam, F. Y. Wu // IBM Systems Journal. - 2007. — №46 (4). — pp. 703-721.
11. *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.

Надійшла до редколегії 05.04.2017