

ДОСЛІДЖЕННЯ МЕТОДІВ ВИЯВЛЕННЯ АНОМАЛЬНИХ СИТУАЦІЙ У МЕДИЧНИХ БІЗНЕС-ПРОЦЕСАХ

Мартиненко О.В., Чалий С.Ф.

e-mail: oleksii.martynenko@nure.ua, serhii.chalyi@nure.ua

Харківський національний університет радіоелектроніки, каф. ІУС
м. Харків, Україна

The study addresses the identification of anomalous situations in medical business processes, emphasizing the importance of combining data analysis and sequence detection. Existing anomaly detection methods typically focus on either data attributes or process sequence deviations. However, their integration into a unified approach is critical, particularly in healthcare, where patient waiting times, examination durations, and the order of medical procedures are paramount. The proposed methodology extends process mining techniques to incorporate fuzzy inference systems, enabling robust anomaly detection even with incomplete data.

Існуючі методи виявлення аномальних ситуацій аналізують відхилення в послідовності робіт процесу або в даних процесу. Проте відсутність поєднання аналізу обох аспектів бізнес-процесу у рамках єдиного підходу є проблемою. Вона особливо актуальна в медичній сфері, де ці обидва аспекти вважаються критично важливими. Дані, такі як час очікування пацієнтів, час проведення оглядів та досліджень, та послідовності виконання дій бізнес-процесів, зокрема порядок виконання процедур пацієнтами, мають рівнозначний вплив на якість надання медичних послуг. Відповідно, постає необхідність розробки такого методу, що дозволить виявляти аномальні ситуації у медичних бізнес-процесах з урахуванням послідовності дій процесів та даних. Для вирішення цієї задачі пропонується розширити існуючий підхід для виявлення аномальних ситуацій з використанням процесного аналізу [1], який дозволяє на основі логу бізнес-процесу виявити його типову модель.

Вирази (1), (2), (3) відображають відповідні залежності між активностями, трасами та логами:

$$A = \{a_1, a_2, \dots, a_n\}, \quad (1)$$

$$T = \{t_1, t_2, \dots, t_m\}, \quad t_i \in \overline{A^*}, \quad i = \overline{1, m}, \quad (2)$$

$$L \subseteq T'^2, \quad T' \subseteq \overline{T}, \quad (3)$$

де A – множина усіх активностей $a_i, i = \overline{1, n}$; A^* – множина усіх можливих комбінацій послідовностей активностей a_i ; T – множина усіх трас t_i ; L – лог, множина декартового добутку підмножини трас T' .

Для виявлення типової моделі бізнес-процесу, зокрема, з використанням генетичного алгоритму [2, 3], будується відповідна причинна матриця (causal matrix). Вона використовується для порівняння отриманих варіантів моделі з існуючими трасами логу. Відповідно, ключовим критерієм вибору типової моделі бізнес-процесу у цьому підході є відсоток трас, що відповідає їй – чим більший відсоток, тим краще підібрана модель. У результаті отримана типова модель бізнес-процесу використовується для виявлення відхилень у послідовності дій процесу – одного з аспектів розглянутої раніше проблеми виявлення аномальних ситуацій.

Розширення ж цього підходу має на меті додати новий етап, який дозволить також врахувати другий аспект проблеми, а саме виявлення відхилень у даних бізнес-процесу. Для цього пропонується використовувати системи нечіткого виведення [3] на базі правил, що враховуватимуть існуючі знання про предметну область медичних бізнес-процесів. Це дозволить виявляти аномалії у випадках, коли відсутня достатня кількість даних для застосування більш класичних методів, що базуються на використанні статистичних параметрів досліджуваних вибірок.

Відповідні системи нечіткого виведення складаються з низки етапів [4]: задання бази правил, фазифікації вхідних даних, агрегації умов правил, активізації висновків правил, акумуляції висновків, дефазифікації результату. Загалом, кожний з етапів є стандартизованим, проте для деяких з них було конкретизовано методи, які використовуються у процесі виконання системи. Зокрема, агрегація умов виконується з використанням нечіткої кон'юнкції та нечіткою диз'юнкції для операцій «І» та «АБО» відповідно:

$$p_i(w_j \wedge w_{j+1}) = \min\{w_j, w_{j+1}\}, \quad j = \overline{1, n-1}, \quad (4)$$

$$p_i(w_j \vee w_{j+1}) = \max\{w_j, w_{j+1}\}, \quad j = \overline{1, n-1}, \quad (5)$$

де p_i – результат агрегації умови i ; w_j – нечітке число, отримане у процесі фазифікації.

На заключному етапі дефазифікації використовується метод центру тяжіння (Centre of Gravity):

$$y = \frac{\int_{\min}^{\max} x\mu(x)dx}{\int_{\min}^{\max} \mu(x)dx}, \quad (6)$$

де y – результат етапу дефазифікації, дійсне число, яке відображає значення імовірності аномалії у відсотках; x – змінна, що відповідає лінгвістичній змінній; $\mu(x)$ – функція приналежності лінгвістичної

змінної; *min*, *max*–інтервал значень лінгвістичної змінної. У цьому випадку лінгвістична змінна це «імовірність аномалії».

На основі отриманої імовірності Y можна робити висновок про його аномальність на основі деякого порогу P , що може задаватись окремо для кожного процесу в моделі бізнес-процесів. Відповідно, траса з логів є аномальною, якщо виконується умова $Y > P$.

З використанням запропонованого методу на останньому етапі виконується аналіз даних процесів в залежності від застосованої бази правил.

У результаті отримуємо наступні етапи підходу до виявлення аномальних ситуацій: попередня обробка журналу логів бізнес-процесу – видаляються неповноцінні траси та несуттєві активності, що визначаються медичними фахівцями (спеціалістами); виявлення типової моделі бізнес-процесу з використанням генетичного алгоритму; виділення аномальних треків з логу шляхом порівняння послідовності активностей трас логу зі знайденою типовою моделлю бізнес процесу; виділення аномальних треків з множини нормальних, що залишились у результаті попереднього етапу – новий доданий етап, що використовує системи нечіткого виведення для аналізу даних бізнес-процесу.

Розроблений удосконалений метод підвищує точність виявлення аномальних ситуацій у медичних бізнес-процесах завдяки використанню процесного аналізу та систем нечіткого виведення, які об'єднують у єдиний підхід обидва аспекти зазначеної проблеми пошуку відхилень у послідовності дій процесу та в даних процесу.

Список використаних джерел:

1. Aalst W., Medeiros A, Weijters A. Genetic Process Mining. An experimental valuation, Data Mining and Knowledge Discovery. 2005. Vol. 14. P. 48–69.
2. Чала О. В., Богатов Є. О. Побудова моделі бізнес-процесу з використанням темпоральних знань при впровадженні процесного управління. Біоніка інтелекту. 2023. № 1 (99). С. 19–25.
3. Чалий С. Ф., Богатов Є. О. Технологія автоматизованої побудови моделі прототипу бізнес-процесу на основі попередньої обробки журналу подій. Вісник Національного технічного університету "ХПІ": зб. наук. пр. Серія : Системний аналіз, управління та інформаційні технології. 2020. № 2 (4). С. 57–63.
4. West F. Fuzzy Inference System. New-York: NY Research Press, 2015. 326 p.