

4. China's Progress Towards Green Growth: An International Perspective [Electronic resource]. – Mode of access: <http://www.oecd.org/env/country-reviews/PR-China-Green-Growth-Progress-Report-2018.pdf>

5. Возобновляемая энергетика: развитие за счет Китая [Электронный ресурс]. – Mode of access: <https://nangs.org/news/renewables/vozobnovlyаемaya-energetika-razvitie-za-schet-kitaya>

Владимирський О. А.,

*доктор технічних наук, професор
кафедри економіко-математичних дисциплін
та інформаційних технологій;*

Єршова О. Л.,

*кандидат економічних наук, доцент,
завідувач кафедри економіко-математичних дисциплін
та інформаційних технологій;*

Томашевська Т. В.,

*кандидат технічних наук, доцент,
доцент кафедри економіко-математичних дисциплін та
інформаційних технологій,
Національна академія статистики, обліку та аудиту, м. Київ*

МЕТОДИ PROCESS MINING: ПІДХОДИ ТА МОЖЛИВОСТІ ДЛЯ АНАЛІЗУ БІЗНЕС-ПРОЦЕСІВ В ІТ-КОМПАНІЯХ

Одним з ключових викликів, що поставлені сьогодні перед ІТ компаніями – здобуття корисної та цінної інформації з даних, які зберігаються в інформаційних системах, а саме – в журналах подій. Цим займається відносно новий напрямок під назвою Process Mining, або інтелектуальний аналіз процесів. Йому присвячена книга професора Віла Ван дер Аалста «Process Mining. Discovery. Conformance and Enhancement of Business Processes» [19]. Process Mining (надалі РМ) – новий потужний інструмент для підвищення ефективності бізнес-процесів у компаніях та організаціях, які працюють в різних галузях економіки. Зростання популярності цього методу обумовлено такими чинниками: усе більше подій фіксуються, забезпечуючи детальну інформацію, саме про історію процесів; виробники ПЗ для бізнес аналітики не виправдали очікувань споживачів. В основу РМ покладено технології моделювання бізнес процесів та підходи Data Mining. Автор [1] підкреслює, що РМ – більше ніж об'єднання існуючих підходів. РМ не обмежується виявленням процесів. Співставляючи дані про події та моделі процесів, РМ дає можливість перевіряти відповідність виявляти відхилення, здійснювати підтримку прийняття рішень та давати рекомендації.

Таким чином, в [1] виділені три ключові підходи, які застосовують в РМ: здобуття – побудова моделі на підставі записів про те, що відбувалося фактично; перевірка відповідності – дає відповіді на питання, де та чому реальний процес відхиляється від очікуваного; удосконалення – відповідає на питання, що слід змінити в моделі щоб покращити певні показники.

В [1] також вказано, що РМ має справу з двома типами процесів: так званими, Lasagna та Spaghetti процесами. Процеси Lasagna чітко структуровані та обробляються в визначеному порядку. Процес Spaghetti – заплутаний, не структурований процес.

Джерело [1] максимально детально описує підходи РМ, це єдина книга, яка повністю присвячена даному підходу. Одна із самих цінних ідей, описаних в книзі – можливість відтворити модель бізнес-процеса з лога.

В джерелі [2] зазначено, що серед великої кількості алгоритмів РМ, які були запропоновані останнім часом, не існує загальноприйнятих орієнтирів для оцінки та порівняння цих алгоритмів РМ. В результаті, може бути складно вибрати відповідний алгоритм РМ для конкретного підприємства або області застосування. Було розроблено систему для порівнянь характеристик алгоритмів для вирішення цієї проблеми. Проте, оцінюючи існуючі алгоритми РМ проти великого набору бізнес-моделей (наприклад, у великому підприємстві) їх порівняння заради вибору може бути утомливо, забирати багато часу та мати велику обчислювальну складність.

У роботі [2] досліджуються масштабовані рішення, які можуть ефективно оцінювати, порівнювати і ранжувати алгоритми РМ. Також, було запропоновано новий фреймворк, який може ефективно вибирати алгоритми РМ, які є найбільш підходящі для даної моделі. Зокрема, за допомогою даного фреймворку, тільки частина моделей процесу потребує емпіричної оцінки. Інші можуть бути рекомендовані безпосередньо за допомогою регресійної моделі. Для подальшої оптимізації, у джерелі [2] також наведено техніку вибору еталонних моделей високої якості для отримання ефективних регресійних моделей. У якості прикладу наведено експеримент на штучних та реальних наборах даних для підтвердження того, що запропонований підхід є практичним і є кращим за традиційний підхід. Для підтвердження цього довелося ознайомитися з джерелами [3-4].

Багато алгоритмів РМ запропоновані на сьогоднішній день, наприклад у [3, 4]. Однак, як зазначено у [5], загальноприйнятий спосіб оцінки та порівняння різних алгоритмів РМ досі не визначений. Наприклад, існує чимало методик відновлення процесу за журналом подій і різні науковці пропонують свою. Тим не менш, не існує єдиної думки щодо якості цих методів.

Власне, сам термін РМ пов'язаний із перетворенням інформації про реальні дії у структурований опис процесу [6], його завданням є отримання якісних моделей для відтворення робочих процесів з мінімальною кількістю інформації, як це можливо. Важливим і очевидним критерієм якості моделі робочого процесу є узгодженість між здобутою моделлю і журналом операцій (логом).

Критерії якості (в тому числі: доречність, простота, точність і узагальнення), які засновані на журналах подій, широко обговорювалися. Але головним завданням є досягнення балансу між цими критеріями якості [5], тому що всі критерії важливі і конкурують один з одним. Крім того, журнали подій часто дуже великі і важкі до сприйняття людиною.

Окрім критеріїв якості, неодноразово розглядалася оцінка якості моделей бізнес-процесів і розуміння того, що є «хорошою» моделлю бізнес-процесу (наприклад, [7-9]).

Наприклад, деякі моделі бізнес-процесів (у тому числі моделі з WFNNets [10] та дані [11, 12]) занадто прості, щоб диференціювати продуктивність цих широко використовуваних алгоритмів РМ.

Перша спроба розробки загальної основи оцінки алгоритмів РМ була зроблена Розинатом та ін. [13, 14].

Їх фреймворк оцінювання складається з чотирьох компонентів, а саме:

- журнал подій та репозиторій;
- модуль генерації логу;
- інструменти модифікації і валідації (обробки логу);
- модуль оцінювання і порівняння.

Основна увага приділяється оцінці відповідності між журналом подій і отриманою моделлю процесу.

Найголовніше, як було зазначено авторами в [13], повний набір критеріїв (в ідеалі, що містять як штучні і реальні набори даних) до теперішнього часу відсутній. В цьому джерелі [13] адаптовано і розширено концепцію емпіричної оцінки алгоритмів РМ на основі поведінкової подібності та структурної схожості між еталонними моделями процесу і отриманою моделлю.

Запропоновано багато підходів до подібності для моделей бізнес-процесів, наприклад, [14-17]. Слід розглянути підхід з [18] зі структурною мірою подібності, оскільки він добре підходить для визнання відносної відстані між початковими моделями та отриманими моделями.

Поведінкова подібність вимірюється за принципом перехідної послідовності (PTS) [19], так як він може мати справу з багатьма складними структурами, такими, як необмежені мережі Петрі, і т.п., і може бути визначена ефективно, як зазначено в [20].

Проте, фреймворк оцінювання легко може бути використаний в інших схемах вимірювання, щоб задовольнити будь-які специфічні вимоги. Цими алгоритмами РМ є: α алгоритм [21], генетичний алгоритм [22], heuristic miner і region miner [23]. Вважається, що ці чотири алгоритми представляють чотири найбільш популярні класи алгоритмів РМ. І, нарешті, описано ідею використання регресії для оцінки алгоритмів РМ на моделях бізнес-процесів.

В джерелі [24] йде мова про те, що сучасні інформаційні системи (наприклад, WFM, ERP, CRM, SCM і B2B системи) записують події в так звані журнали подій (логи). РМ використовує ці журнали, щоб відтворити процес, управління, дані, організаційні і соціальні структури. Хоча багато дослідників розробляють нові і більш потужні методи РМ, розробники програмного забезпечення все одно вбудовують логування в свої програмні продукти. Деякі з найбільш передових методів видобутку процесів були протестовані на реальних життєвих процесах. У [24] описується застосування РМ в одному з провінційних відділень голландського відділу національних

громадських робіт (Dutch national public work department), відповідальному за будівництво та утримання дорожньої та водної інфраструктури.

Використовуючи різні методи РМ, було проаналізовано обробку рахунків-фактур, що посилаються різними субпідрядниками та постачальниками з трьох різних точок зору:

- процес (the process perspective).
- організація (the organizational perspective).
- випадок (the case perspective).

Для цього було використано деякі з інструментів, розроблених із застосуванням фреймворку ProM.

В роботі [25] детально розглянуті можливість та доцільність використання методів РМ в соціальних мережах через взаємодію шаблонів бізнес-процесів.

Є багато алгоритмів РМ з різними теоретичними основами і цілями, порушуючи питання про те, як вибрати найкращий для конкретної ситуації. У джерелі [26] запропоновано фреймворк для об'єктивного порівняння алгоритмів РМ, що використовуються для відновлення бізнес-процесів. Результати експериментальної оцінки п'яти алгоритмів на нескладних процесах підтверджують правильність такого підходу.

В роботі [27] зроблено огляд великої кількості літератури щодо РМ, створеної, починаючи з середини 1990-х років.

В роботі [28] розглянуто необхідність всеосяжного фреймворку для РМ, що включає в себе тестові дані, інструменти і методи порівняння. Описано та протестовано різні методи оцінки. Аргументовано необхідність великої точності інформації, отриманої у результаті РМ.

Аналіз відповідності та метрики для оцінки моделей процесів також описано в [29], включно з плагіном перевірки відповідності для ProM. Тестування відповідності до певної структури процесів описано в роботі [30], в контексті тестування генетичного алгоритму інтелектуального аналізу на певних структурах. У [31] наведено порівняння поведінки різних алгоритмів на різних моделях, з використанням принципу мінімальної довжини опису (Minimum description length) в цілях стиснення логу.

В [32] CPN Tools і ProM plug-in використовується для створення та моделювання мереж Петрі, контролю ймовірності, із застосуванням ProMimport для завантаження журналів подій в ProM. Зворотний процес описано в роботі [33], поряд з проблемами визначення ймовірностей настання подій у процесах.

В [33] автор запропонував власний підхід до відновлення бізнес-процесів по журналам подій IC, за допомогою яких ведеться конкретний бізнес – розглядає задачу отримання актуальної моделі процесу діяльності у вигляді мережі Петрі, її аналіз та побудову ефективного розподілу ресурсів для реалізації цього процесу.

В [34] для дослідження обраний найбільш відомий фреймворк, спеціально розроблений для ПМ – ProM. Він має відкритий код і в якості вхідного формату має MXML. Результатом застосування даного інструменту

отримаємо актуальну модель у вигляді мережі Петрі, а також модель організаційного виконання процесів. Для даної задачі мережі Петрі дають можливість з заданою точністю представляти розгалужені, паралельні та циклічні процеси, мають засоби для аналізу та моделювання в реальному масштабі часу.

Основні проблеми, що існують досі в області Process Mining-у викладені [1]: зареєстровані дані можуть бути неправильні або неповні через проблеми зі створенням (шум); завдання, які існують, але не можуть бути знайдені в даних (приховані завдання); два вузла процесу можуть стосуватися однієї і тієї ж моделі процесу (повторювані завдання); керований вибір, може залежати від вибору, зробленого в іншій частині моделі процесу (невільний вибір); процес може бути виконаний в кілька разів, петлі можуть бути простими за участю одного або декількох подій або більш складним (видобуток петель); події процесу можуть бути додані з додатковою інформацією для процесу виявлення (різні перспективи); порівняння моделі процесу і еталонної моделі для перевірки схожості чи невідповідності (дельта аналіз); результати виявлення процесу можуть бути представлені в графічній формі в термінах панелі менеджменту (візуалізація результатів); доступ до інформаційних систем, заснованим на різних платформах (гетерогенні результати); видобуток процесів, що відбуваються в той же самий час (одночасні процеси); локальні стратегії обмежують простір пошуку і менш складні, глобальні стратегії є складними, але мають більше шансів знайти оптимальне рішення (локальний та глобальний пошук); вибір алгоритму виявлення, який може повторно виявити клас моделей процесу з повного журналу процесу (процес повторного виявлення).

Не дивлячись на те, що багато які з завдань процесу виявлення можуть бути вирішені за допомогою комбінації модифікованих підходів інтелектуального аналізу даних і призначених для користувача вбудованих алгоритмів, не існує єдиного підходу, який може вирішити всі проблеми, з якими стикаються в процесі виявлення. Багато існуючих алгоритмів інтелектуального аналізу користувальницьких процесів, як правило, вирішують тільки одну або дві проблеми. Найбільш широко до задач Process Mining застосовуються генетичні алгоритми. Такі підходи показали хороші результати в області зниження шуму і виявлення прихованих операцій. Спроб використовувати технології нейронних мереж для подібних завдань виявлення було менше, можливо, через їх складність.

Список використаних джерел:

1. van der Aalst W.M.P. Process Mining / M.P.W. van der Aalst. – Berlin: Springer-Verlag Berlin Heidelberg, 2016. – 467 p.
2. Efficient Selection of Process Mining Algorithms / J. Wang et. al. [Electronic resource]. – Mode of access: <http://www.cse.unsw.edu.au/~wong/papers/tsc12.pdf> (Last accessed: 1.03.2020)
3. van der Aalst W.M.P. Workflow mining: discovering process models from event logs / W.M.P. van der Aalst, A.J.M.M. Weijters, L. Maruster // IEEE

Transactions on Knowledge and Data Engineering. – 2004. – Vol. 16 (9). – P. 1128–1142.

4. Wen L. Mining process models with non-free-choice constructs / L. Wen, W.M.P. van der Aalst, J. Wang, J. Sun // Data Mining and Knowledge Discovery. – 2007. – Vol. 15 (2). – P. 145–180.

5. IEEE Task Force on Process Mining. Process Mining Manifesto // Lecture Notes in Business Information Processing. – Berlin: SpringerVerlag, Berlin, 2011. – Vol. 99. – P. 169–194.

6. Workflow mining: A survey of issues and approaches / W.M.P. van der Aalst, B.F. van Dongen, J. Herbst, L. Maruster, G. Schimm, A.J.M.M. Weijters // Data and Knowledge Engineering. – 2003. – P. 237–267.

7. Mendling J. Testing density as a complexity metric for EPCs [Electronic resource] / J. Mendling // German EPC Workshop on Density of Process Models. – 2006. – Mode of access: <https://www.semanticscholar.org/paper/Testing-Density-as-a-Complexity-Metric-for-EPCs-Mendling/ce8e8be0bed01fd034088e8e4e5bab2895e11687>

8. Mendling J., Reijers H.A., Cardoso J. What makes process models understandable? [Electronic resource] / J. Mendling, H. A. Reijers, J. Cardoso. – Mode of access: https://link.springer.com/chapter/10.1007/978-3-540-75183-0_4

9. Quality metrics for business process models / I. Vanderfeesten, J. Cardoso, J. Mendling, H.A. Reijers, W.M.P. van der Aalst // BPM and workflow handbook. – 2007. – P. 179–190.

10. van der Aalst W.M.P. The application of Petri nets to workflow management / M.P.W. van der Aalst // The Journal of Circuits, Systems and Computers. – 1998. – No.8 (1). – P. 21–66.

11. Faulty EPCs in the SAP reference model / J. Mendling, M. Moser, G. Neumann, H.M.W. Verbeek, B.F. van Dongen, W.M.P. van der Aalst // BPM'06. – 2006. – P. 451–457.

12. Detection and prediction of errors in EPCs of the SAP reference model / J. Mendling, H.M.W. Verbeek, B.F. van Dongen, W.M.P. van der Aalst, G. Neumann // Data and Knowledge Engineering. – 2008. – No. 64 (1). – P. 312–329.

13. Towards an evaluation framework for process mining algorithm [Electronic resource] / A. Rozinat, A.K. Alves de Medeiros, C.W. Gunther, A.J.M.M. Weijters, W.M.P. van der Aalst. – Mode of access: <https://pdfs.semanticscholar.org/d690/739283b54929cb9b536387e5775a8cd41d54.pdf>

14. Rozinat A. The need for a process mining evaluation framework in research and practice [Electronic resource] / A. Rozinat, A.K. Alves de Medeiros, A.J.M.M. Weijters, W.M.P. van der Aalst. – Mode of access: https://www.researchgate.net/publication/220268118_The_Need_for_a_Process_Mining_Evaluation_Framework_in_Research_and_Practice

15. Dijkman R.M. Graph matching algorithms for business process model similarity search / R.M. Dijkman, M. Dumas, L. Garc'ia-Ba nuelos. // Lecture Notes in Computer Science. BPM 2009 / ed. U. Dayal et al. – Berlin: SpringerVerlag, 2009. – Vol. 5701. – P. 48–63.

16. Dumas M. Similarity search of business process models / M. Dumas, L. Garc'ia-Ba nuelos, R.M. Dijkman // Bulletin of the IEEE Technical Committee on Data Engineering. – 2009. – No. 32 (3). – P. 25–30.

17. A work- flow net similarity measure based on transition adjacency relations / H. Zha, J. Wang, L. Wen, C. Wang, J. Sun // Computers in Industry. – 2010. – No. 61 (5). – P. 463–471.

18. Development of distance measures for process mining, discovery and integration / J. Bae, L. Liu, J. Caverlee, L-J. Zhang, H. Bae // Int. J. Web Service Research. – 2007. – No. 4 (4). – P. 1–17.

19. A behavioral similarity measure between labeled Petri nets based on principal transition sequences - (short paper) / J. Wang, T. He, L. Wen, N. Wu, A.H.M. ter Hofstede, J. Su // OTM Conferences. – 2010. – No. 1. – P. 394–401.

20. Becker M., Laue R. Analysing differences between business process similarity measures. International Workshop on Process Model Collections. Clermont-Ferrand, France. 2011.

21. de Medeiros A.K.A. Genetic process mining: an experimental evaluation / A.K.A. de Medeiros, A.J.M.M. Weijters, W.M.P. van der Aalst // Data & Knowledge Engineering. – 2007. – No. 14. – P. 245–304.

22. Weijters A.J.M.M. Process mining with Heuristics Miner- algorithm [Electronic resource] / A.J.M.M. Weijters, W.M.P. van der Aalst/ – Mode of access: https://www.researchgate.net/publication/306014995_Process_mining_with_the_heuristics_miner-algorithm

23. Carmona J. A regionbased algorithm for discovering Petri nets from event logs / J. Carmona, J. Cortadella, M. Kishinevsky // BPM'08. – 2008. – P. 358–373.

24. On recommendation of process mining algorithms / J. Wang, R.K. Wong, J. Ding, Q. Guo, L. Wen // IEEE ICWS. – 2012. – P. 311-318.

25. van der Aalst W.M.P. Mining Social Networks: Uncovering Interaction Patterns in Business Processes / W.M.P. van der Aalst, M. Song // International Conference on Business Process Management (BPM 2004): Lecture Notes in Computer Science / ed. J. Desel, B. Pernici, M. Weske. – Berlin: Springer-Verlag. – 2004. – Vol. 3080. – P. 244–260.

26. Weber P. A framework for comparing process mining algorithms. Birmingham [Electronic resource] / P. Weber, B. Bordbar, P. Tino. – Mode of access: https://www.researchgate.net/publication/260652522_A_Framework_for_the_Analysis_of_Process_Mining_Algorithms

27. Tiwari A. A review of business process mining: state-of-the-art and future trends. Bus / A. Tiwari, C.J. Turner, B. Majeed // Process Manage Journal. – 2008. – Vol. 14. – No. 1. – P. 5– 22.

28. Towards an evaluation framework for process mining algorithms [Electronic resource] / A. Rozinat, A.K. Alves De Medeiros, C.W. Günther, A.J.M.M. Weijters, W.M.P. Aalst, van der. 2007. – Mode of access: <https://research.tue.nl/en/publications/towards-an-evaluation-framework-for-process-mining-algorithms>

29. Rozinat A. Conformance checking of processes based on monitoring real behavior / A. Rozinat, W.M.P. van der Aalst // Information System. – 2008. – Vol. 33. – No. 1. – P. 64–95.

30. De Medeiros A.K.A. Genetic process mining: An experimental evaluation / A.K.A. de Medeiros, A.J.M.M. Weijters, W.M.P. van der Aalst // Data Mining and Knowledge Discovery. – 2007. – Vol. 14. – No. 2. – P. 245–304.

31. Calders T. Using minimum description length for process mining / T. Calders, C. W. Gunther, M. Pechenizkiy, A. Rozinat // SAC '09: Proceedings of the 2009 ACM symposium on Applied Computing, New York, NY, USA, 2009. NY, 2009. P. 1451–1455.

32. Rozinat A. Discovering colored Petri nets from event logs / A. Rozinat, R. S. Mans, M. Song, W.M.P. van der Aalst // J. Softw. Tools Technol. Transf. – 2008. – Vol. 10. – No. 1. – P. 57–74.

33. Rozinat A. Discovering simulation models / A. Rozinat, R. S. Mans, M. Song, W.M.P. van der Aalst // Information System. – 2009. – Vol. 34. – No. 3. – P. 305–327.

34. van der Aalst W.M.P. Prom: The Process Mining Toolkit [Electronic resource] / W.M.P. van der Aalst, B.F. van Dongen1, C. Günther. – Mode of access: <http://www.wis.win.tue.nl/~wvdaalst/publications/p555.pdf>

Жукович І. А.,

*кандидат економічних наук, старший науковий співробітник,
провідний науковий співробітник офісу оцінювання діяльності
наукових установ НАН України,*

*ДУ «Центр оцінювання діяльності наукових установ та
наукового забезпечення розвитку регіонів України НАН України», м. Київ*

ОСОБЛИВОСТІ ФІНАНСУВАННЯ ТА ОЦІНЮВАННЯ НАУКОВО-ДОСЛІДНИХ УСТАНОВ ЧЕХІЇ

Впровадження системи фінансування на основі ефективності є одним із центральних механізмів, завдяки якому багато держав-членів ЄС намагаються підвищити ефективність та результативність своїх систем наукових досліджень у публічному секторі.

Для реалізації Національної політики розвитку досліджень та інновацій Чеської Республіки на 2016–2020 рр. [1], уряд Чехії поставив за мету до 2020 року збільшити видатки на наукові дослідження та розробки до 2% ВВП, що буде сприяти віднесенню Чехії до десятки кращих країн Європи. Збільшення видатків на науку буде відбуватись не лише за рахунок державних витрат, а також через стимулювання інвестиції у дослідження серед підприємств з одночасним залученням європейських ресурсів.

Ключовими галузями, що будуть формувати Національну політику розвитку досліджень та інновацій визначені біотехнологія та нанотехнології, цифрова економіка, автомобільна та авіаційна промисловість та залізничні перевезення, а також традиційні сектори, такі як машини, електроніка, лиття сталі та енергетика. Також значна увага буде приділятися культурно-творчим галузям.