

Поддержка принятия решений при контроле документации при управлении проектами на базе поиска ассоциативных правил

В. А. Латыпова

Уфимский университет науки и технологий

vikvaphoto@yandex.ru

Аннотация. Интеллектуальная поддержка принятия решений, базирующаяся на интеллектуальном анализе данных, используется в управлении проектами для решения различных задач, однако она не охватывает процесс контроля проектной документации. Данный процесс является трудоемким, а также достаточно значимым, ведь от качества проектной документации зависит успешность реализуемого проекта. Предлагаемый подход к поддержке принятия решений при контроле проектной документации на базе поиска ассоциативных правил и интеллектуальная система направлены на решение данной проблемы. Апробация предлагаемого решения проводилась в вузе при проверке проектной документации.

Ключевые слова: поддержка принятия решений; контроль проектной документации; интеллектуальный анализ данных; поиск ассоциативных правил; интеллектуальный анализ комментариев

I. ВВЕДЕНИЕ

В настоящее время проекты, проектная деятельность широко используются в организациях. Интеллектуальный анализ данных применяют для обеспечения интеллектуальной поддержки принятия решений в различных видах управления проектами: например, при управлении рисками, ресурсами [1], расписанием [2]. Также первый используется на различных стадиях жизненного цикла проекта: например, при мониторинге для обнаружения отклонений, при планировании для оценки количества потенциальных дефектов [3], и при решении рутинных задач управления [4]. При этом используются различные методики интеллектуального анализа данных, включая: классификацию/кластеризацию, деревья решений, поиск ассоциативных правил и др., широко применяемые и в других областях, где возможен сбор большого массива данных, формируемого при взаимодействии с информационными системами (ИС) (например, в обучении [5] с использованием систем дистанционного обучения, в производстве [6] с использованием PDM-систем).

Тем не менее, в текущих исследованиях не уделяется достаточного внимания интеллектуальной поддержке процедуры контроля проектной документации, несмотря на то что данная процедура является значимой. Качество проектной документации напрямую влияет на исход проекта, будет ли он успешным, реализованным в рамках

требуемых затрат, сроков и качества, или нет. Также требуется много времени и усилий со стороны сотрудников проектного офиса для проведения контроля проектной документации.

Для решения данной проблемы в статье предлагается подход к поддержке принятия решений сотрудников проектного офиса при контроле проектной документации и соответствующая интеллектуальная система.

Структура статьи следующая. В главе II описаны существующие подходы, использующие поиск ассоциативных правил в проектном управлении, представлена информация об ассоциативных правилах и методах их поиска. Описание предлагаемого подхода и реализация интеллектуальной системы поддержки принятия решений (ИСППР) сотрудников проектного офиса представлены в главах III и IV, соответственно. В главе V описана апробация подхода и системы на данных, полученных при контроле проектной документации.

II. РАБОТЫ ПО ТЕМАТИКЕ ИССЛЕДОВАНИЯ

A. Существующие подходы, использующие поиск ассоциативных правил в проектном управлении

Методика поиска ассоциативных правил имеет широкое распространение в проектном управлении и используется для различных задач. В Таблице I представлено описание данных задач, типы проектов, в которых они решаются, приведены элементы-факторы, используемые в ассоциативных правилах при поиске решения.

Поиск ассоциативных правил может помочь при управлении сложными проектами с гибридными моделями, используя прошлый опыт, предлагая комбинации различных практик управления [7].

В [8] поиск ассоциативных правил используется для извлечения взаимосвязей между концептами онтологии в области проектного управления. Концепты могут получить статус: сильно связаны, связаны, не связаны. Наибольший эффект при выявлении взаимосвязей достигается при совместном использовании поиска ассоциативных правил и косинусного сходства.

ТАБЛИЦА I. ПОИСК АССОЦИАТИВНЫХ ПРАВИЛ ДЛЯ РЕШЕНИЯ ЗАДАЧ В ПРОЕКТНОМ УПРАВЛЕНИИ

№ задачи	Описание задач	Тип проекта	Элементы-факторы, влияющие на решение задачи	Публикация
1	оценка рисков спортивных мероприятий	организация мероприятия	ситуация по счету (равный счет, отстают, ведут), фаза игры (начало, конец, первая и вторая половина)	[9]
2	прогнозирование длительности проекта	Любой	параметры освоенного объема, показатели отклонений от плана (по времени и стоимости)	[10]
3	выбор наилучшего метода управления проектами	IoT, информационные технологии (ИТ)	тип, размер и ориентация проекта, уровень устранимого риска, проведение ежедневных встреч, размер команды, наличие аппаратного обеспечения, итерационность, роли	[11]
4	оценка рисков, связанных с запросами на внесение изменений	ИТ	тип запроса (новое требование, изменить требование, дефект), приоритет	[12]
5	выбор членов команды	//-//-//	фронтенд разработка (html+css+JS), бэкенд разработка (php+mysql), скорость работы, общительность, опыт работы, награды, членство в профессиональных организациях, возраст	[13]
6	прогнозирование усилий на этапах проекта	//-//-//	усилия на предыдущих этапах проекта (например, планирования, разработки спецификации, проектирования и т.д.)	[14]
7	оценка успеха проекта (с помощью выявления связи между мерами риска и результатом проекта)	//-//-//	уровни различного вида рисков: команды, заказчиков, требований, сложности проекта, управления проектом, окружения организации	[15]
8	оценка влияния факторов политики управления на атрибуты проекта (качество программного обеспечения, усилия на разработку, длительность проекта)	//-//-//	время задержки в адаптации и встраивании новых сотрудников при их найме	[16]
9	извлечение взаимосвязей между параметрами сообщений о проблеме в системе слежения за ошибками	//-//-//	как относится сообщение к проекту (приоритет, класс, серьезность), длительность устранения проблемы	[17]

В. Выявление ассоциативных правил

Ассоциативные правила – это продукционные правила, правила вида: ЕСЛИ факт1 ТО факт2 (другое представление: $fact1 \Rightarrow fact2$). В качестве факта1 выступает причина (антецедент правила), а факта2 – следствие (консеквент правила), т. о. правила отображают причинно-следственную связь. Данные правила ищут в наборах элементов-фактов в составе транзактов баз данных (БД) ИС.

Чтобы ассоциативное правило было выявлено, необходимо чтобы оно обладало на заданном уровне следующими свойствами:

1. поддержка $RS(fact1 \Rightarrow fact2)$: относительная частота (вероятность) совместной встречаемости факта1 и факта2 в наборах;
2. достоверность $RC(fact1 \Rightarrow fact2)$: частота срабатывания правила (условная вероятность факта2 при условии факта1);
3. зависимость фактов в правиле $RL(fact1 \Rightarrow fact2)$: насколько факт2 зависит от факта1 (отношение условной вероятности факта2 при условии факта1 к вероятности факта2).

Пороговые значения перечисленных выше свойств при поиске правил задаются экспертно. Однако, свойства RC и RL должны удовлетворять условиям: $RC > 0.5$, $RL > 1$. При $RC \leq 0.5$ правило срабатывает в 50 и менее процентов случаев. При $RL < 1$ отсутствие факта 1 приводит к факту2, т. е. срабатывает антиправило вида: $не(fact1) \Rightarrow fact2$. Все это является неприемлемым для выделения правила.

Значение свойства RS показывает, насколько редким является правило, не влияя на качество последнего, определяемого свойствами RC и RL . Чем выше RS , тем

более распространенным является правило, и регулируя его уровень, можно менять общее количество выявленных правил. Значимые правила имеют значения свойств RS , RC и RL на уровне, который превосходит пороговый.

С. Методы поиска ассоциативных правил

Для поиска ассоциативных правил используются методы (первые два наиболее популярны и активно модифицируются исследователями):

1) Apriori, базирующийся на итерационном процессе, включающем генерацию наборов-кандидатов (в том числе наборов, не содержащихся в транзакциях), их очистку от избыточных наборов и расчет параметра RS для каждого набора-кандидата. Вариации метода: “AprioriTID” [18]; “FApriori” [19], “R-Apriori”[20], “RDD-Apriori”[21], “TRMR-Apriori”[22], EAFIM [23]

2) FP-growth, базирующийся на создании FP-дерева (суффиксного дерева, построенного из частей элементов наборов, отсортированных в убывающем порядке) и его анализе. Вариации метода: “AFOP”, “NONORDFP”, “FPGROWTH*”, отличающиеся использованием различных структур данных для представления FP-дерева [24];

3) другие методы: Eclat [18], генетический алгоритм [15].

III. ПОДХОД К ПОДДЕРЖКЕ ПРИНЯТИЯ РЕШЕНИЙ СОТРУДНИКОВ ПРОЕКТНОГО ОФИСА ПРИ КОНТРОЛЕ ПРОЕКТНОЙ ДОКУМЕНТАЦИИ НА БАЗЕ ПОИСКА АССОЦИАТИВНЫХ ПРАВИЛ

Предлагаемый подход основан на майнинге замечаний сотрудников проектного офиса по контролируемой проектной документации и включает положения:

1. при контроле проектной документации используется ИС, сохраняющая результаты первого;
2. замечания сотрудников проектного офиса хранятся в БД ИС в виде списка в типизированном виде;
3. замечания типизируются сотрудниками проектного офиса вручную, либо с использованием кластеризации пула первых;
4. находятся значимые ассоциативные правила между типизированными замечаниями сотрудников проектного офиса, которые далее используются для создания производной базы знаний (БЗ).

Свойства ассоциативных правил между типизированными замечаниями сотрудников проектного офиса $comm1$ и $comm2$ RS , RC , RL рассчитываются по формулам:

$$RS(comm1 \Rightarrow comm2) = \frac{count(comm1 \wedge comm2)}{LN},$$

где $count(comm1 \wedge comm2)$ – количество наборов с совместной встречаемостью $comm1$ и $comm2$; LN – количество выгруженных списков замечаний сотрудников проектного офиса;

$$RC(comm1 \Rightarrow comm2) = \frac{RS(comm1 \Rightarrow comm2)}{RS(comm1)},$$

где $RS(comm1)$ – относительная частота (вероятность) встречаемости $comm1$ в списках;

$$RL(comm1 \Rightarrow comm2) = \frac{RC(comm1 \Rightarrow comm2)}{RS(comm2)},$$

где $RS(comm2)$ – относительная частота (вероятность) встречаемости $comm2$ в списках;

5. используется ИСППР, содержащая БЗ взаимосвязи типизированных замечаний для поиска потенциальных ошибок по уже выявленным при проверке проектной документации.

Предлагаемый подход является развитием подходов, представленных авторами ранее в [25] (для сферы образования при контроле научных статей магистрантов) и [26] (для сферы производства при согласовании конструкторской документации).

Перед процедурой майнинга замечаний списки последних L_m выводятся из БД ИС в форме:

$$L_m = \{RI_{mn}|n=1,2, \dots, RN_m\}; m = 1,2, \dots, LN;$$

$$RI_m < RI_{m+1},$$

где RI_{mn} – код n -го замечания по проектной документации в БД ИС в m -м списке; RN_m – количество замечаний в m -м списке.

Алгоритм подготовки замечаний для майнинга и создания БЗ взаимосвязи типизированных замечаний в составе ИСППР сотрудников проектного офиса представлен на Рис. 1.

IV. РЕАЛИЗАЦИЯ ИСППР СОТРУДНИКОВ ПРОЕКТНОГО ОФИСА ПРИ КОНТРОЛЕ ПРОЕКТНОЙ ДОКУМЕНТАЦИИ

ИСППР встроена в систему контроля проектной документации и разработана на JavaScript, PHP с использованием MySQL. ИСППР включает:

- модуль подготовки данных для производционного вывода;
- библиотеку Таuprolog (для производционного вывода связанных замечаний сотрудников проектного офиса по проектной документации);
- модуль отображения связанных замечаний по проектной документации;
- модуль записи связанных замечаний сотрудников проектного офиса в БД;
- БЗ взаимосвязи замечаний.

Рис. 2 показывает архитектуру ИСППР сотрудников проектного офиса (UML диаграмму компонентов).

Для поиска значимых ассоциативных правил применяется программное средство SPMF [27], реализующее метод FP-growth. Скриншот системы при выводе связанных замечаний сотрудников проектного офиса по проектной документации представлен на рис. 3.

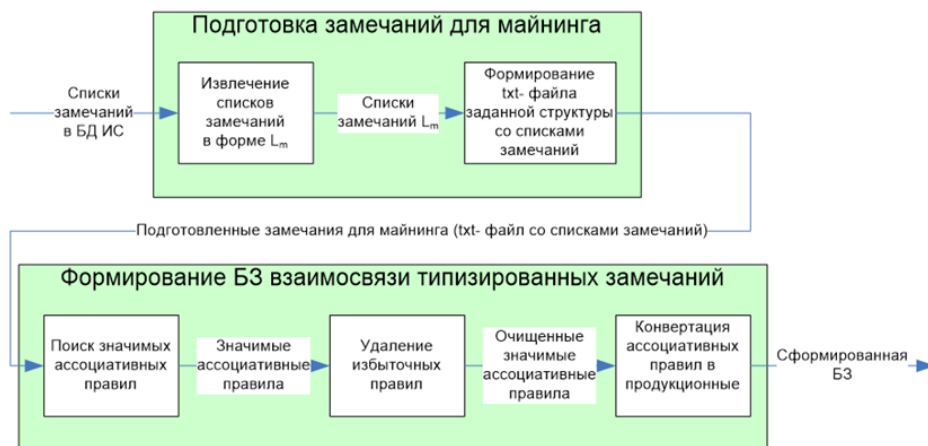


Рис. 1. Алгоритм подготовки замечаний для майнинга и создания БЗ взаимосвязи типизированных замечаний в составе ИСППР

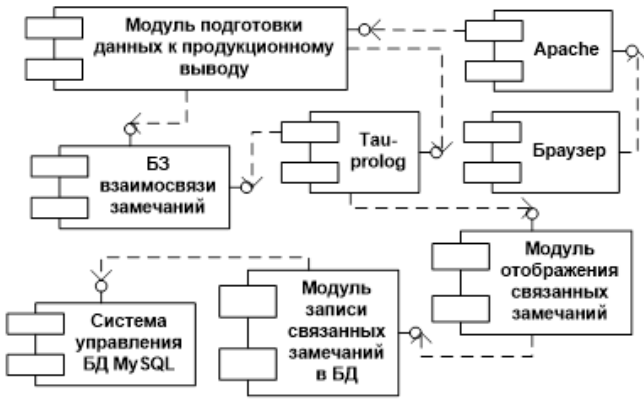


Рис. 2. Архитектура ИСППР сотрудников проектного офиса

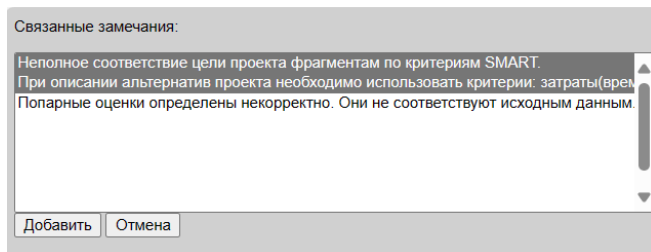


Рис. 3. Скриншот ИСППР при выводе связанных замечаний сотрудников проектного офиса по проектной документации

V. ОПИСАНИЕ ЭКСПЕРИМЕНТА И РЕЗУЛЬТАТЫ

Контроль проектной документации осуществлялся при обучении студентов направления прикладной информатики в курсе «Проектный практикум» при проверке расчетно-графических работ (РГР) в университете. Задание состояло в разработке проектной документации в заданной предметной области при управлении проектом одного из 3-х типов:

- внедрения ИС;
- разработки ИС;
- открытия организации.

Проектная документация в составе РГР должна включать элементы:

- паспорт проекта;
- иерархическую структуру работ проекта;
- календарно-ресурсный план, выполненный в MS Project.

Собраны данные за 3 семестра обучения с 2020 по 2022 гг., по 9 группам.

Фрагмент найденных значимых ассоциативных правил взаимосвязи типизированных замечаний при ограничениях: $RC=0.6$, $RS=0.1$, $RL=1.1$, показан в табл. II, а коды в БД и текст замечаний по проектной документации, относящийся к ним – в табл. III.

На основе данных правил была сформирована БЗ взаимосвязи замечаний ИСППР. Пример вывода данной системой связанных замечаний для поддержки принятия решений при контроле проектной документации в составе РГР представлен на рис. 3. Комментарии, предложенные системой, были оценены как полезные тьюторами-экспертами.

ТАБЛИЦА II. ФРАГМЕНТ НАЙДЕННЫХ ЗНАЧИМЫХ АССОЦИАТИВНЫХ ПРАВИЛ ВЗАИМОСВЯЗИ ТИПИЗИРОВАННЫХ ЗАМЕЧАНИЙ

№	Ассоциативное правило		RS	RC	RL
	Антецедент	Консеквент			
1	4973 4986	4970	0.101	1.0	2.379
2	5017	4986	0.101	0.875	1.509
3	4973	4970	0.13	0.818	1.947
4	4982 4985	4986	0.13	0.75	1.294
5	4993	4986	0.116	0.727	1.255
6	4975	4971	0.101	0.7	2.1
7	4974 4986	4970	0.101	0.7	1.666
8	4986	4971	0.101	0.7	2.1
9	4982 5007	4986	0.101	0.7	1.208
10	4982 4986	4970	0.159	0.688	1.636
11	5008	5009	0.116	0.667	3.286
12	4973	4971	0.101	0.636	1.91
13	4982	4986	0.232	0.64	1.104
14	4973	4970 4986	0.101	0.636	1.91
15	4971 4986	4970	0.13	0.6	1.428

ТАБЛИЦА III. Коды в БД и текст замечаний, относящихся к ним

Код в БД	Текст замечаний
4970	Неполное соответствие цели проекта фрагментам по критериям SMART
4971	При описании альтернатив проекта необходимо использовать критерии: затраты(временные), стоимость и их конкретные значения
4973	В поле контактной информации по заинтересованным лицам нужно указать конкретную э/почту, телефон
4974	Для количественных критериев должны быть представлены диапазоны их значений
4975	Не должно быть действий в продуктовой иерархической структуре работ
4982	Фрагменты/часть фрагментов цели проекта по критериям SMART не корректны
4985	Параметр CR должен быть менее 0,2. Большее значение показывает наличие несогласованности во мнениях эксперта при оценке альтернатив
4986	Попарные оценки определены некорректно. Они не соответствуют исходным данным
4993	Описание альтернатив проекта должно быть подробным, объемом 1/3–1/2 страницы
5007	Необходимо устранить дублирование при описании результатов
5008	Сроки для предлагаемого решения должны соответствовать пункту паспорта «Предварительная оценка продолжительности проекта»
5009	Затраты для предлагаемого решения должны соответствовать пункту паспорта «Предварительная оценка стоимости проекта»
5017	Итоговая продолжительность проекта отсутствует

VI. ЗАКЛЮЧЕНИЕ

Рассмотрены существующие подходы, использующие поиск ассоциативных правил в проектном управлении.

Предложен подход к поддержке принятия решений сотрудников проектного офиса при контроле проектной документации при управлении проектами на базе поиска ассоциативных правил и соответствующая интеллектуальная система. Подход апробирован при контроле проектной документации в рамках курса «Проектный практикум» в вузе.

СПИСОК ЛИТЕРАТУРЫ

- [1] Obaid T., Abu-Naser S.S. Big data analytics in project management: a key to success // IJAER. 2023. Vol. 7. Iss. 7. pp. 1-8.
- [2] Wei W., Rana M.E. Software project schedule management using machine learning & data mining // Int. J. of Scientific & Technology Research. 2019. Vol.8. Iss. 9. pp. 1385-1389.
- [3] Pospieszny P. Application of data mining techniques in project management – an overview // Collegium of Economic Analysis Annals. 2017. Iss. 43. pp.199-220.
- [4] Vasiliev A.A., Goryachev A.V. Data mining in project management // Proceedings of the 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Russian Federation. 2022. pp. 482-486.
- [5] Latypova V. Automated system for checking works with free response using intelligent tutor's comment analysis in engineering education // Proceedings of the 2022 VI International Conference on Information Technologies in Engineering Education (Inforino), Moscow, Russian Federation. 2022. pp. 1-5. DOI: 10.1109/Inforino53888.2022.9782966.
- [6] Kabirova A., Kulagin I., Emaletdinova L. Development and software implementation of the design documentation verification model based on association rules // LNEE, vol. 986. Springer, Cham, 2023. pp. 12–22.
- [7] Bianchi M.J., Amaral D.C. A systematic review of association rules in project management: opportunities for hybrid models // Product. 2020. Vol.18. No.2. pp.136-144. DOI:10.4322/pmd.2020.033.
- [8] Gulla J.A., Brasethvik T., Kvarv G.S. Using association rules to learn concept relationships in ontologies // Proceedings of the 10th International Conference on Enterprise Information Systems (ISAS). 2008. pp. 58-65. DOI: 10.5220/0001686200580065.
- [9] Pan Y., Chen Z. Research on risk management of large-scale sports events based on association rule mining algorithm// Proceedings of the 2022 2nd International Conference on Networking, Communications and Information Technology. 2022. pp. 170-174.
- [10] Iranmanesh S., Mokhtari Z. Application of data mining tools to predicate completion time of a project // World Academy of Science: Engineering and Technology. 2008. Vol. 42.
- [11] Merzouk S., Gandoul R., Marzak A., Sael N. Toward new data for IT and IoT project management method prediction // Mathematical modeling and computing. 2023. Vol. 10. No. 2. pp. 557–565.
- [12] Mello O.C., Fontoura L.M. Improving the evaluation of change requests using past cases // Int. J. of Information Systems and Project Management. 2023. Vol. 11. No. 1. pp. 74-89.
- [13] Sulova S. Association rule mining for improvement of it project management // TEM J. 2018. Vol.7. Iss.4. pp. 717-722.
- [14] Azzeh M., Cowling P., Neagu D. Software stage-effort estimation based on association rule mining and fuzzy set theory // Proceedings of the 10th IEEE International Conference on Computer and Information Technology. 2010. pp. 249–256.
- [15] Shan X., Jiang G., Huang T. A framework of estimating software project success potential based on association rule mining // Proceedings of the 2009 International Conference on Management and Service Science, Beijing, China. 2009. pp. 1-4.
- [16] García M.N.M., Roman I.R., Peñalvo F.J.G., Bonilla M.T. An association rule mining method for estimating the impact of project management policies on software quality, development time and effort // Expert Systems with Applications. 2008. Vol.34. pp.522-529.
- [17] Nayak R., Qiu T. A data mining application: analysis of problems occurring during a software project development process // Int. J. of Software Engineering and Knowledge Engineering. 2005. Vol.15.
- [18] Saabith A.L.S., Sundararajan E., Bakar A.A. Comparative analysis of different versions of association rule mining algorithm on AWS-EC2/ Badioze Zaman, H., et al. // Lecture Notes in Computer Science, vol. 9429. Springer, Cham, 2015.
- [19] Patel M.R., Rana D.P., Mehta R.G. FApriori: A modified Apriori algorithm based on checkpoint // 2013 International Conference on Information Systems and Computer Networks. 2013. pp. 50-53.
- [20] Rathee S., Kashyap A., Kaul M. R-Apriori: An efficient apriori based algorithm on Spark // PIKM'15. 2015.
- [21] Singh P., Singh S., Mishra P.K., Garg R. A data structure perspective to the RDD-based Apriori algorithm on Spark// Int. j. inf. tecnol. 2022. Vol.14. pp.1585–1594. DOI: 10.1007/s41870-019-00337-3
- [22] Saabith S., Sundararajan E., Abu Bakar A. A parallel Apriori-transaction reduction algorithm using Hadoop-Mapreduce in Cloud // Asian J. of Computer Science and Information Technology. 2018. Vol. 1. DOI: 10.9734/AJRCOS/2018/41045.
- [23] Raj S., Ramesh D., Sreenu M., Sethi K.K. EAFIM: efficient apriori-based frequent itemset mining algorithm on Spark for big transactional data // Knowl Inf Syst. 2020. Vol.62. pp. 3565–3583.
- [24] Said A.M., Dominic D., Abdullah A.B. A comparative study of FP-growth variations // Int. J. of Computer Science and Network Security. 2009. Vol.9. No.5. pp.266-272.
- [25] Latypova V. Decision support based on analysis of relationship between errors using association rule mining on the example of graduate students' scientific papers // Proceedings of the 2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT), Yekaterinburg, Russian Federation. 2023. pp. 217-220. DOI: 10.1109/USBREIT58508.2023.10158892.
- [26] Latypova V. Intelligent system of approving person information support based on association rule mining between errors in design documentation // Advances in Automation V, RusAutoCon 2023, Lecture Notes in Electrical Engineering, vol 1130. Springer, Cham, 2024. pp. 311–320. DOI: 10.1007/978-3-031-51127-1_30.
- [27] Fournier-Viger P., Lin C.W., Gomariz A., Gueniche T., Soltani A., Deng Z., Lam H.T. The SPMF open-source data mining library version 2 // LNCS, vol 9853. Springer, Cham, 2016. pp. 36-40.