

Literatura

1. Baldin K.V., Bryzgalov N.A., Rukosuev A.V. *Matematicheskoe programmirovaniye: uchebnyk*. M.: Dashkov i K, 2016. 218 s.
2. Boltneva A.A., Rykunov Ya.I., Stratutsa V.A. *Primeneniye metodov optimizatsii upravleniya k logisticheskim protsessam* // *Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnyye sistemy: modeli, analiz i upravleniye"*. 2018. Vyp. 1. S. 55–63.
3. Vecherskaya S.E. *Postanovka i algoritm resheniya zadachi optimizatsii upravleniya: uchebno-metodicheskoye posobie*. M.: RosNOU, 2018. S. 40.
4. Vecherskaya S.E., Gavryushova K.A. *Optimizatsiya i/ili reinzhiniring* // *Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnyye sistemy: modeli, analiz i upravleniye"*. 2016. Vyp. 4. S. 34–38.
5. Gubar' Yu.V. *Vvedeniye v matematicheskoye programmirovaniye*. 2-e izd. M.: Internet-universitet informatsionnykh tekhnologiy (INTUIT), 2016. 226 s.
6. Tarasov V.N., Bakhareeva N.F. *Matematicheskoye programmirovaniye: teoriya, algoritmy, programmy*. Samara: PGUTI, 2007.
7. Booking.com. URL: <https://www.booking.com/> (data obrashcheniya: 02.02.2019).
8. Hotel Kurfürst Gmb. URL: <http://www.kurfuerst.com/> (data obrashcheniya: 02.02.2019).
9. Why is GDS Important to the Travel Industry? // Travel Technology & Solutions. URL: <http://www.tts.com/blog/why-is-gds-important-to-the-travel-industry/> (date of the application: 02.02.2019).

DOI: 10.25586/RNUV9187.19.01.P.109

УДК 004.91; 614

О.В. Золотарев, А.Х. Хакимова, М.М. Шарнин

РАЗРАБОТКА МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО
АНАЛИЗА НАУЧНЫХ ПУБЛИКАЦИЙ
ДЛЯ МОНИТОРИНГА ПРИОРИТЕТНЫХ
НАПРАВЛЕНИЙ РАЗВИТИЯ ПРЕВЕНТИВНОЙ
И ПЕРСОНАЛИЗИРОВАННОЙ МЕДИЦИНЫ*

Рассматриваются вопросы интеллектуального анализа научных публикаций для мониторинга приоритетных направлений развития превентивной и персонализированной медицины. Показано динамичное развитие методов определения современных тенденций в развитии медицины. Описаны различные методы интеллектуального анализа научных публикаций. Проведен анализ публикаций медицинского электронного ресурса PubMed. Предложен метод выделения трендов в развитии направлений превентивной и персонализированной медицины на основе нейронных сетей.

Ключевые слова: интеллектуальный анализ данных, приоритетные направления, прогнозирование, нейронные сети, векторное представление слов.

* Работа выполнена при поддержке Российского фонда фундаментальных исследований, гранты № 16-07-00756, 16-29-09527, 18-07-00909 + 18-07-01111. Авторы благодарны РФФИ за поддержку и финансирование их проектов.

O.V. Zolotarev, A.K. Khakimova, M.M. Charnin

DEVELOPMENT OF METHODS OF INTELLECTUAL ANALYSIS
OF SCIENTIFIC PUBLICATIONS FOR MONITORING
THE PRIORITY DIRECTIONS OF DEVELOPMENT
OF PREVENTIVE AND PERSONALIZED MEDICINE

The article discusses the issues of intellectual analysis of scientific publications for monitoring the priority directions of development of preventive and personalized medicine. The dynamic development of methods for determining current trends in the development of medicine is shown. Various methods of intellectual analysis of scientific publications are described. The analysis of publications of the medical electronic resource PubMed has been carried out. A method for identifying trends in the development of areas of preventive and personalized medicine based on neural networks is proposed.

Keywords: Data mining, priority directions, prediction, neural networks, vector word representation.

Введение

В настоящее время большое внимание уделяется вопросам прогнозирования и оценки развития научных дисциплин, выявления новых областей науки, выделения приоритетных научных направлений. Вычислительное моделирование эволюции науки, отслеживание «взлетов и падений» научных тем со временем приобретает значение для финансирования перспективных направлений исследований, для организации научных мероприятий, для корректировки стратегий развития технологических компаний.

В медицине для прогнозирования развития новых направлений используется широкий спектр методов и моделей, включая поточечные оценки, регрессионные и авторегрессионные модели, методы машинного обучения, основанные на байесовских сетях, искусственных нейронных сетях, методы фильтрации и др. Важную роль играет анализ публикаций, связанных с новыми направлениями в науке. В связи с громадным потоком неструктурированной информации и необходимостью обобщения результатов исследований в данной работе предлагается использовать методы анализа научных публикаций, основанные на рекуррентных нейронных сетях, с целью выделения наиболее приоритетных направлений развития превентивной и персонализированной медицины.

Методы прогнозирования в медицине

На основе анализа совместной встречаемости слов различных категорий и корреляции между ними был создан метод «co-word analysis» [8]. Он способствует выявлению ассоциаций среди терминов и на их базе построению сетей, отражающих картину эволюции любой дисциплины.

Метод семантического спектра позволяет исследовать динамику ключевых слов во времени и рассматривать аспекты структурных изменений в различных областях на основании анализа частот встречаемости ключевых слов [1]. По мере развития научного направления их частота растет, прекращение работ по какому-либо научному направлению приводит к уменьшению частоты встречаемости определенных терминов. Например, анализ по ключевым словам изменения тематики публикаций авторов из Казахстана

с 2004 по 2014 гг. показал, что темы «моделирование» и «астрофизика» в Казахстане находятся в развитии и в то же время отмечается снижение интереса к экологии водных ресурсов [2].

М. Steyvers и другие [12] предложили модель «автор – тема» (АТМ) [14]. Она представляет документ как продукт смеси тем авторов, где каждое слово генерируется при активации одной из тем автора этого документа, но временное упорядочение отсутствует. D.M. Blei и другие [7] на базе модели АТМ предложили сегментированную модель «автор – тема» (S-АТМ), которая интегрирует временные характеристики коллекции документов в генеративный процесс. S-АТМ показывает возможность выявления эволюции тем с течением времени.

DTM (динамическая тематическая модель) разработана для отслеживания эволюции темы путем последовательного группирования совокупности документов на базе предположения так, что темы в текущем временном интервале плавно эволюционировали из соответствующих тем в предыдущем временном интервале (Wang и McCallum, 2006 [14]; Blei и Lafferty, 2006 [7], Wang, Zhai и Roth, 2013 [15], Tang и другие, 2013 [13]).

Электронные ресурсы в области медицины

PubMed – один из ведущих мировых медицинских ресурсов, он ежедневно обрабатывает миллионы запросов и является важным инструментом для исследователей в области здравоохранения во всем мире [11].

Научные исследования в медицине очень активны, количество журналов и проиндексированных ссылок в MEDLINE, крупнейшей библиографической базе статей по медицинским наукам, увеличивается со значительной скоростью. PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>) является крупнейшим доступным ресурсом с 1996 г. Более 26 миллионов ссылок из MEDLINE, журналов по естественным наукам и онлайн-книг включены в PubMed в области биомедицины. На рисунке 1 представлены накопленные в MEDLINE и PubMed ссылки в период с 2013 по 2018 гг., по данным MEDLINE® STATISTICS (https://www.nlm.nih.gov/bsd/medline_pubmed_production_stats.html). С ростом объема публикаций в медицинских науках отслеживание эволюции научных областей и прогнозирование будущих тенденций исследований имеют большое значение для ученых.

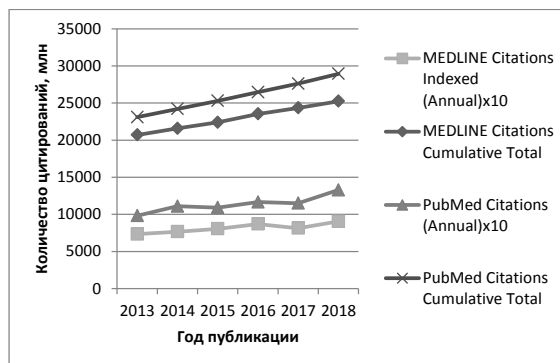


Рис. 1. Количество проиндексированных ссылок, накопленных, а также добавляемых в MEDLINE и PubMed в течение каждого года с 2013 по 2018 гг., по данным MEDLINE® STATISTICS (https://www.nlm.nih.gov/bsd/medline_pubmed_production_stats.html)

Метод анализа научных публикаций: тренд-генерация

При обработке текста публикации формируются словари слов (общий словарь слов, словарь слов предметной области, словарь текста), встречающихся в тексте, за исключением так называемых стоп-слов, т.е. общеупотребительных слов, таких как «также», «скорее всего», «может быть» и др. [4; 5]. При формировании словарей производится нормализация слов (перевод слов в единичное число, настоящее время, именительный падеж, неопределенную форму – для глаголов и т.д.) [3]. Далее подсчитывается частота встречаемости слова в тексте, формируются векторы слов. Структура словаря определяются следующим образом. Каждое слово в словаре представляется как вектор:

$$D = ((\text{Main}W, h), (W_1, n_1, v_1), (W_2, n_2, v_2), \dots, (W_k, n_k, v_k)), \quad (1)$$

где $\text{Main}W$ – основное слово словаря; h – частота встречаемости основного слова в текстах; W_k – k -е слово, стоящее в анализируемом тексте рядом с основным словом (парное слово); n_k – количество раз, которое парное слово встретилось рядом с основным словом; v_k – относительная частота встречаемости пары слов – основного и парного (v_k/h).

Каждый словарь представляется в виде матрицы векторов. Обычно при формировании матрицы векторов слов выбирается размерность вектора в зависимости от задачи от 300 до 500, размер словаря (множества векторов) неограничен.

В результате формируются три словаря слов (множества векторов): словарь текста, словарь предметной области, общий словарь.

D_{gen} – общий словарь;

$D_{sub} = \{D_1^s, D_2^s, \dots, D_k^s\}$ – множество словарей предметных областей;

$D_{doc} = \{D_1^d, D_2^d, \dots, D_m^d\}$ – множество словарей документов.

На основе этих словарей можно предсказывать наиболее вероятное следующее слово при анализе текстов. Наиболее популярные направления могут рассчитываться на основе вычисления максимального градиента относительной частоты встречаемости слов (терминов). Частота встречаемости пар терминов (частота события) называется статистической вероятностью. Для расчета градиента статистической вероятности необходимо ввести интервал дискредитации времени – частоту семплирования. Минимальный интервал семплирования устанавливается в размере одного месяца.

Для определения тенденций развития перспективных направлений может использоваться показатель частоты встречаемости терминов, что уже было описано. В данной работе предлагается использовать подход, основанный на увеличении частоты встречаемости множеств связанных пар терминов, которые образуют тему. Выделение тем может осуществляться разными способами. Один из наиболее употребительных на сегодня подходов – использование методов тематического моделирования. Эта методика основана на вероятностном определении множества тем документа.

В нашей работе для выделения наиболее перспективных направлений используется градиент статистической вероятности изменения частоты встречаемости пар терминов во времени (с учетом частоты семплирования). Чтобы добиться этого, необходимо формировать множество общих словарей попарной встречаемости терминов $\{D_{gen}\}$ в соответствии с частотой семплирования. В результате можно будет вычислять градиенты изменения статистической вероятности между множествами пар терминов по времени.

Для построения градиента сравниваются частоты встречаемости пар слов для последовательных во времени общих словарей – D_j и D_{j+1} . Для устойчивых фраз градиенты изменения для пар связанных терминов будут похожими (термины связаны, если градиент изменения больше нуля). Анализ выполняется последовательно, сначала для первого главного термина, затем для связанного с ним парного термина, который выбирается далее в качестве главного слова. Так как векторы слов строятся последовательно, то максимальная цепочка слов, соответствующая ненулевому градиенту, и будет перспективной или развивающейся темой, фразой, направлением.

Если обозначить частоту встречаемости парного термина с основным термином как n_{jk} для текущего варианта словаря D_j , то $n_{j+1,k}$ будет определять ту же частоту для варианта словаря для следующего периода семплирования. Тогда для любого главного слова необходимо найти градиент возрастания частоты употребления указанной пары терминов:

$$\text{grad} (D_{j+1} (W_k, n_k, v_k), D_j (W_k, n_k, v_k)). \quad (2)$$

Множества пар связанных терминов образуют рекуррентную нейронную сеть. На входе нейронной сети – связанные пары терминов, на выходе – темы, фразы, наиболее динамично развивающиеся направления в различных отраслях знаний. Входными вершинами нейронной сети служат термины, веса – градиенты возрастания частоты использования пары терминов.

На рисунке 2 представлена рекуррентная нейронная сеть с входным, выходным и скрытым слоем, которая используется для поиска новых направлений исследований. На рисунке представлен один уровень скрытого слоя. Входной слой представляет однословные термины, несколько скрытых слоев образуют сочетание нескольких терминов в единое высказывание. При этом, как уже упоминалось, объединяются термины, соответствующие увеличению градиента частоты совместной встречаемости. Количество скрытых слоев определяется длиной высказывания. Если высказывание состоит из двух слов, то скрытый слой совпадает с выходным слоем, который представляет собой окончательное высказывание (фразу, набирающую популярность идею и т.д.).

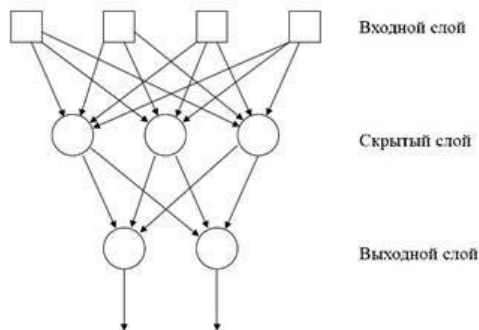


Рис. 2. Структура рекуррентной сети

Аналогичным образом можно строить словари предложений, но при этом не выполняется нормализация каждого слова и не удаляются стоп-слова. Каждое предложение входит в словарь в том виде, в котором оно встретилось в тексте. Подобный подход используется в современных системах автоматического перевода (Google, Яндекс и др.).

Настройка сети происходит на основе большого количества параллельных текстов на разных языках, при этом настраивается перевод между парами языков. В результате сопоставления появляется возможность опосредованного перевода с одного языка на другой (если нет прямой настроенной сети для перевода с одного языка на другой, но есть опосредованная связь через пары других языков).

Как уже было сказано, для анализа развития новых направлений можно использовать подход на основе методов тематического моделирования. В зависимости от длины текста может генерироваться заданное количество тем документа.

Аналогичным образом, как уже отмечалось, строятся три словаря тем, подсчитывается частота встречаемости конкретной темы в документе, в словаре предметной области, в общем словаре тем. Если частота встречаемости темы растет, то ее популярность увеличивается, в противном случае – уменьшается.

Описанные методы могут дополнять друг друга при сравнении результатов выделения определений тенденций при анализе публикаций.

Анализ тенденций публикаций PubMed

Мы провели анализ трендов в области 4P-медицины (4П – на русском языке) и медицинской генетики за период с 1940 по 2018 г. [10].

В медицинской среде есть такое понятие – «Медицина 4П», или «Медицина – 2030». «4П» – это аббревиатура от первых букв основных векторов новой медицины: предсказание, профилактика, персонализация и партисипативность (участие пациента).

Анализ тенденций публикаций PubMed, демонстрирующий возросший научный интерес к области 4P-медицины, охватывает период 1950–2018 г. Использовались ключевые слова поиска: predictive medicine, personalized medicine, personalized treatment. Из графика (рис. 3) очевиден резкий рост числа публикаций в перечисленных областях, по-видимому, благодаря деятельности созданной в 2008 г. Европейской ассоциации предиктивно-превентивной и персонализированной медицины (ЕРМА – European Association for Predictive, Preventive and Personalised Medicine), активно продвигающей идеологию 4P-медицины [9].

В работе [6] была показана возможность визуализации связанных объектов, выделенных из научных текстов с использованием A-Frame технологии.

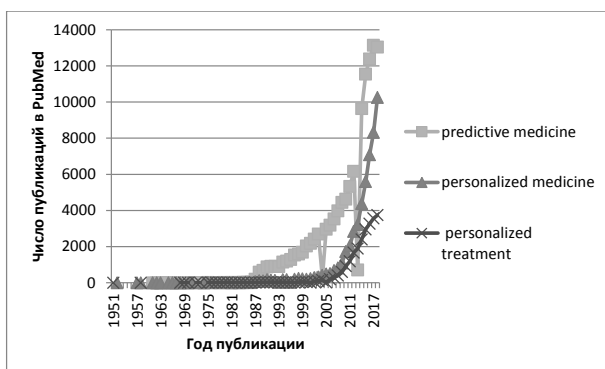


Рис. 3. Статистика публикаций PubMed в области предиктивной и персонализированной медицины

Заключение

В результате анализа научных публикаций в области медицины показано, что предсказательная и профилактическая медицина становятся реальным трендом современности. Дальнейшее развитие медицины неотрывно связано с развитием новых перспективных направлений.

В данной работе описан метод определения трендов при анализе научных публикаций в области медицины.

Описанный в данной статье метод тренд-генерации предполагается использовать в целях анализа научных публикаций для мониторинга приоритетных направлений развития превентивной и персонализированной медицины.

Литература

1. Агеев Б.А., Черноног С.Б., Розуменко С.Б. Анализ развития научных направлений методом «семантического спектра» // НТИ. Серия 1: Организация и методика информационной работы. 1990. № 5. С. 16–18.
2. Акоев М.А. Картирование науки и технологии, прогноз развития // Руководство по наукометрии: индикаторы развития науки и технологии. Екатеринбург: Изд-во Урал. ун-та, 2014. С. 164–200.
3. Золотарев О.В. Формализация знаний о предметной области на основе анализа естественно-языковых структур // Цивилизация знаний: проблема человека в науке XXI века: труды Двенадцатой Международной научной конференции (Москва, 22–23 апреля 2011 г.): в 2 ч. М.: РосНОУ, 2011. Ч. II. С. 78–80.
4. Золотарев О.В., Козеренко Е.Б., Шарнин М.М. Принципы построения моделей бизнес-процессов предметной области на основе обработки текстов естественного языка // Вестник Российского нового университета. 2014. Вып. 4. С. 82–88.
5. Золотарев О.В., Шарнин М.М., Клименко С.В., Мацкевич А.Г. Исследование методов автоматического формирования ассоциативно-иерархического портрета предметной области // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2018. Вып. 1. С. 91–96.
6. Золотарев О.В., Шарнин М.М., Клименко С.В., Мацкевич А.Г. Семантический подход к визуализации научных документов с использованием ВЕБ-графики 3D // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2018. Вып. 2. С. 46–53.
7. Blei D.M., Lafferty J.D. Dynamic Topic Models // ICML '06: Proceedings of the 23rd International Conference on Machine Learning. N. Y.: ACM Press, 2006. P. 113–120.
8. Courtial J.P., Callon M., Sigogneau M. Is Indexing Trustworthy – Classification of Articles Through Co-word Analysis // J. Inf. Sci. 1984. Vol. 9, № 2. P. 47–56.
9. Heyoung Y., Hyuck J.L. Research Trend Visualization by MeSH Terms from PubMed // International Journal of Environmental Research and Public Health. 2018. № 15 (6). P. 1113. DOI: 10.3390/ijerph15061113.
10. Hood L. Systems Biology and P4 Medicine: Past, Present, and Future // Rambam Maimonides. Medical Journal. 2013. Vol. 4 (2). DOI: 10.5041/RMMJ.10112.
11. Lu Z. PubMed and Beyond: A Survey of Web Tools for Searching Biomedical Literature. [S. l.]: Database, 2011. DOI: 10.1093/database/baq036.
12. Rosen-Zvi M., Griffiths T., Steyvers M., Smyth P. The Author-Topic Model for Authors and Documents // 20th Conference on Uncertainty in Artificial Intelligence: Proceedings. Banff: AUAI Press, 2004. P. 487–494.

13. *Tang S., Zhang Y., Wang H., Chen M., Wu F., Zhuang Y.* The Discovery of Burst Topic and Its Intermittent Evolution in Our Real World. [S. l.]: China Commun, 2013. DOI: 10.1109/CC.2013.6488826.
14. *Wang X., McCallum A.* Topics Over Time: A Non-Markov Continuous-Time Model of Topical Trends // 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: Proceedings. [S. l.]: ACM, 2006. P. 424–433.
15. *Wang X., Zhai C., Roth D.* Understanding Evolution of Research Themes: A Probabilistic Generative Model for Citations // 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: Proceedings. [S. l.]: ACM, 2013. P. 1115–1123.

Literatura

1. *Ageev B.A., Chernonog S.B., Rozumenko S.B.* Analiz razvitiya nauchnykh napravlenij metodom “semanticheskogo spektra” // NTI. Seriya 1: Organizatsiya i metodika informatsionnoj raboty. 1990. № 5. S. 16–18.
2. *Akoev M.A.* Kartirovanie nauki i tekhnologii, prognoz razvitiya // Rukovodstvo po naukometrii: indikatory razvitiya nauki i tekhnologii. Ekaterinburg: Izd-vo Ural. un-ta, 2014. S. 164–200.
3. *Zolotarev O.V.* Formalizatsiya znaniy o predmetnoj oblasti na osnove analiza estestvenno-yazykovykh struktur // Tsvivilizatsiya znaniy: problema cheloveka v nauke XXI veka: trudy Dvenadtsatoj Mezhdunarodnoj nauchnoj konferentsii (Moskva, 22–23 aprelya 2011 g.): v 2 ch. M.: RosNOU, 2011. Ch. II. S. 78–80.
4. *Zolotarev O.V., Kozerenko E.B., Sharnin M.M.* Printsipy postroeniya modelej biznes-protsessov predmetnoj oblasti na osnove obrabotki tekstov estestvennogo yazyka // Vestnik Rossijskogo novogo universiteta. 2014. Vyp. 4. S. 82–88.
5. *Zolotarev O.V., Sharnin M.M., Klimenko S.V., Matskevich A.G.* Issledovanie metodov avtomaticheskogo formirovaniya assotsiativno-ierarkhicheskogo portreta predmetnoj oblasti // Vestnik Rossijskogo novogo universiteta. Seriya “Slozhnye sistemy: modeli, analiz i upravlenie”. 2018. Vyp. 1. S. 91–96.
6. *Zolotarev O.V., Sharnin M.M., Klimenko S.V., Matskevich A.G.* Semanticheskij podkhod k vizualizatsii nauchnykh dokumentov s ispol'zovaniem VEB-grafiki 3D // Vestnik Rossijskogo novogo universiteta. Seriya “Slozhnye sistemy: modeli, analiz i upravlenie”. 2018. Vyp. 2. S. 46–53.
7. *Blei D.M., Lafferty J.D.* Dynamic Topic Models // ICML '06: Proceedings of the 23rd International Conference on Machine Learning. N. Y.: ACM Press, 2006. P. 113–120.
8. *Courtial J.P., Callon M., Sigogneau M.* Is Indexing Trustworthy – Classification of Articles Through Co-word Analysis // J. Inf. Sci. 1984. Vol. 9, № 2. P. 47–56.
9. *Heyoung Y., Hyuck J.L.* Research Trend Visualization by MeSH Terms from PubMed // International Journal of Environmental Research and Public Health. 2018. № 15 (6). P. 1113. DOI: 10.3390/ijerph15061113.
10. *Hood L.* Systems Biology and P4 Medicine: Past, Present, and Future // Rambam Maimonides. Medical Journal. 2013. Vol. 4 (2). DOI: 10.5041/RMMJ.10112.
11. *Lu Z.* PubMed and Beyond: A Survey of Web Tools for Searching Biomedical Literature. [S. l.]: Database, 2011. DOI: 10.1093/database/baq036.
12. *Rosen-Zvi M., Griffiths T., Steyvers M., Smyth P.* The Author-Topic Model for Authors and Documents // 20th Conference on Uncertainty in Artificial Intelligence: Proceedings. Banff: AUAI Press, 2004. P. 487–494.
13. *Tang S., Zhang Y., Wang H., Chen M., Wu F., Zhuang Y.* The Discovery of Burst Topic and Its Intermittent Evolution in Our Real World. [S. l.]: China Commun, 2013. DOI: 10.1109/CC.2013.6488826.

Сапожников Г.П. Редукция множества показателей мониторингового...

14. Wang X., McCallum A. Topics Over Time: A Non-Markov Continuous-Time Model of Topical Trends // 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: Proceedings. [S. l.]: ACM, 2006. P. 424–433.

15. Wang X., Zhai C., Roth D. Understanding Evolution of Research Themes: A Probabilistic Generative Model for Citations // 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: Proceedings. [S. l.]: ACM, 2013. P. 1115–1123.

DOI: 10.25586/RNUV9187.19.01.P.117

УДК 681.3

Г.П. Сапожников

РЕДУКЦИЯ МНОЖЕСТВА ПОКАЗАТЕЛЕЙ МОНИТОРИНГОВОГО
ОЦЕНИВАНИЯ ДЕЯТЕЛЬНОСТИ НЕКОММЕРЧЕСКОЙ
ОБРАЗОВАТЕЛЬНОЙ ОРГАНИЗАЦИИ
ДЛЯ ОПТИМИЗАЦИИ РЕСУРСОЭФФЕКТИВНОСТИ
ЕЕ ФУНКЦИОНИРОВАНИЯ

Описываются этапы обработки многолетних данных информационного мониторинга деятельности некоммерческой образовательной организации, направленные на возможность последующего использования при оптимизации ресурсоэффективности ее функционирования за счет редукции исходного множества показателей. С этой целью показатели, влияющие на экономические составляющие ресурсоэффективности, и сами эти составляющие рассматриваются как зависимые случайные процессы. Экспертно-визуальный анализ изменений тесноты корреляционных связей от временного сдвига позволяет построить ранговую последовательность на множестве показателей и выделить на первом этапе наиболее влиятельные. Второй этап редукции множества влиятельных показателей осуществляется на основе оптимизационной модели линейного булева программирования, которая решается методом ветвей и границ с вычислением верхней оценки путем решения эквивалентной непрерывной задачи линейного программирования.

Ключевые слова: ресурсоэффективность, оптимизация, информационный мониторинг, редукция, ранжирование, булево программирование.

G.P. Sapozhnikov

REDUCTION OF THE VARIETY OF INDICATORS
OF MONITORING EVALUATION OF THE ACTIVITY
OF NON-COMMERCIAL EDUCATIONAL ORGANIZATION
TO OPTIMIZE THE RESOURCE EFFICIENCY
OF ITS OPERATION

The stages of processing multi-year data of information monitoring of the activities of a non-profit educational organization are described. They are aimed at the possibility of subsequent use in optimizing the resource efficiency of its operation by reducing the initial set of indicators. To this end, indicators affecting the economic components of resource efficiency, and these components themselves are considered as dependent random processes. The expert-visual analysis of changes in the closeness of correlations