

А. С. Козицын, канд. физ.-мат. наук, вед. науч. сотр., alexanderkz@mail.ru,
МГУ им. М. В. Ломоносова

Алгоритмы тематического поиска данных в наукометрических системах

Рассмотрены разработанные автором методы тематического поиска журналов и конференций научно-технического характера с использованием программных средств, реализующих алгоритмы полнотекстового поиска и поиска экспертов. В начале статьи дан краткий обзор существующих систем тематического поиска данных наукометрического содержания с указанием их основных недостатков. Приведено описание разработанных автором алгоритмов, их программной реализации и апробации на больших данных наукометрической системы. Для выполнения запросов разработанные алгоритмы используют полнотекстовые описания объектов в наукометрических системах и графы взаимных связей между объектами, проводят фильтрацию данных по узконаправленным тематическим запросам и ранжируют результаты поиска по степени их авторитетности.

Ключевые слова: тематический анализ, наукометрия, информационные системы, библиография, журнал, конференция

Введение

При создании информационных наукометрических систем для анализа больших объемов данных необходимо использовать интеллектуальные методы поиска и анализа информации, которые заменяют или дополняют использование традиционного контекстного поиска и поиска по атрибутам. Одним из таких методов является тематический анализ информации. Тематический анализ применяется для уточнения наукометрических критериев, анализа основных тенденций научно-технического развития, для информационного поиска и решения других важных задач.

Наиболее распространенные критерии, которые традиционно используются для анализа результатов деятельности научных работников и определения эффективных методов их стимулирования [1], в значительной степени зависят от тематики научных исследований. Например, средние показатели индексов цитирования научных статей в различных тематических областях значительно отличаются [2, 3]. Использование методов тематического анализа позволяет вычислять более объективные наукометрические критерии, например, нормализованную цитируемость [4].

Применение методов тематического анализа также помогает более детально оценить основные тенденции развития мировой науки, наиболее значимые научные направления на уровне го-

сударства или отдельной организации. Для возможности проведения такой оценки большинство разрабатываемых наукометрических систем и систем цитирования имеют модули аналитической обработки и агрегирования данных [5, 6] и специализированные интерфейсы для визуализации тематического распределения данных [7, 8]. Например, наукометрическая информационно-аналитическая система "ИСТИНА" (ИАС "ИСТИНА" [1]) позволяет проводить анализ распределения показателей научных публикаций по тематическим рубрикам Scopus и ГРНТИ [9, 10]. Анализ основных тенденций развития различных тематических направлений науки и своевременное отслеживание их изменений необходимы, в том числе, для выбора эффективных способов стимулирования научной деятельности [11, 12].

В настоящей работе описаны разработанные автором методы тематического поиска журналов и конференций научно-технического содержания. Исходными данными для такого поиска являются описания объектов (статей, докладов, проектов и других результатов научно-технической деятельности), информация о которых загружена в наукометрическую систему. Используются два подхода для поиска: контекстный поиск по описаниям и тематический поиск экспертов, статьи и доклады которых представлены в журналах и на конференциях [13, 14].

Тематический поиск журналов

Актуальность задачи автоматизации тематической классификации и поиска журналов обусловлена в первую очередь большим числом публикуемых в мире журналов и сборников статей, в том числе, содержащих материалы проводимых конференций. Подобные материалы публикуют на русском, английском, немецком, французском и других языках, что также затрудняет задачу тематической классификации [15]. Например, в ИАС "ИСТИНА" [1, 16] зарегистрировано более 70 тысяч журналов и еще более 200 тысяч сборников научных публикаций. Увеличение числа журналов обусловлено в первую очередь увеличением общего числа исследователей, поскольку каждый новый журнал можно рассматривать как социологический результат деятельности вновь образовавшегося сообщества ученых-исследователей. Известный британский историк науки Дерек Джон де Солла Прайс в середине прошлого века предсказывал, что уже в начале XXI века число научных журналов превысит 1 млн [17]. Реальные темпы роста числа научных журналов значительно уступают этим смелым прогнозам. По современным оценкам их количество в мире продолжает увеличиваться ежегодно примерно на 4 % [18]. Однако даже такие объемы информации не позволяют использовать в поисковых системах исключительно ручную обработку данных экспертами и требует создания механизмов автоматизации процесса тематического анализа и поиска.

Существующие системы цитирования предоставляют интерфейс для проведения тематического анализа по заданным рубрикам (Scopus, ГРНТИ и др.). Однако уровень такой рубрикации слишком высок для проведения полноценного тематического анализа. Кроме того, представленные в работе [19] примеры показывают, что используемые в промышленных системах (РИНЦ, Антиплагиат) методы полнотекстового тематического анализа имеют определенные недостатки в области точности и полноты распознавания тематики текстов, а также имеют существенные сложности при сравнении тематики издающихся на разных языках журналов.

В работе [19] предложено использовать графы соавторства [20] для проведения тематического анализа и оценки тематической близости научных журналов в наукометрических системах. Предложенный метод позволяет искать тематически близкие журналы на разных языках и проводить анализ без необходимости получения полных текстов статей. Такой подход значительно расширя-

ет область применения этого метода. В системе РИНЦ, например, полные тексты статей имеют менее 11 % журналов [21].

Исходными данными для поиска журналов являются граф соавторства, который можно получить из наукометрических систем или систем цитирования, а также исходное опорное множество авторов. В работе [19] в качестве опорного множества используется множество авторов выбранного журнала j_1 . На его основе строится множество $M(j_1, j_2)$ всех пар статей, которые были опубликованы в журналах j_1 и j_2 и имеют хотя бы одного совпадающего автора:

$$M(j_1, j_2) = \{d_1, d_2 \in D \mid d_1 \in D_{j_1}, d_2 \in D_{j_2}, A_{d_1} \cap A_{d_2} \neq \emptyset\},$$

где D — множество статей; D_j — множество статей в журнале j ; A_d — множество авторов статьи d .

Вес автора в каждой статье определяется как $w(a, d) = \frac{1}{2} + \frac{1}{2k}$ для первого автора и $w(a, d) = \frac{1}{2k}$ для остальных соавторов, где a — автор; k — число авторов.

Функция определения степени тематической связи между журналами при этом имеет вид

$$F(j_1, j_2) = \sum_a W(a, j_1, j_2), \quad (1)$$

где

$$W(a, j_1, j_2) = \min\left(\max_{d_1 \in D_{j_1}} (w(a, d_1)), \max_{d_2 \in D_{j_2}} (w(a, d_2))\right).$$

Представленный в работе [19] метод достаточно эффективен и обеспечивает точность 78 %. Существенным недостатком изложенного в работе [19] подхода является невозможность поиска журналов по тематике, заданной ключевыми словами. Для применения метода и построения первичного опорного множества необходимо знать хотя бы один журнал, тематика которого соответствует информационной потребности пользователя. Указанное ограничение не позволяет использовать этот метод для тематического поиска по ключевым словам.

В настоящей работе для построения опорного множества авторов и обеспечения возможности поиска журналов по ключевым словам предлагается использовать методы поиска экспертов [13, 14]. Согласно такому подходу, на первом шаге по поисковому запросу выполняется построение множества экспертов с учетом их "весового" соответствия поисковому запросу. Для этого с использованием алгоритма поиска экспертов для поискового

запроса q строится функция релевантности авторов поисковому запросу $v(a, q)$ и определяется множество авторов релевантных запросу $A_q = \{a \in A \mid v(a, q) > 0\}$. Функция тематического соответствия журнала запросу определяется как

$$F^*(q, j) \equiv \sum_{a \in A_q^* \cap A_j} W^*(a, q, j),$$

где $W^*(a, q, j) \equiv v(a, q) \max_{d \in D_j} (w(a, d))$.
 Схема работы алгоритма поиска представлена на рис. 1.

Программная реализация алгоритма выполнена на языке PL/SQL. Основная функция поиска имеет следующее описание на этом языке:

```
create or replace function find_journal4them
(keywords varchar2, --Список ключевых слов через запятую
type_id NUMBER default null, --Тип объектов для поиска опорного множества экспертов
is_graph NUMBER default 1, --Использование графа соавторства для поиска опорного множества экспертов
```

```
journalrum_list VARCHAR2 DEFAULT NULL, --Список рубрик для сужения области поиска
department_list VARCHAR2 DEFAULT NULL, --Список подразделения для сужения опорного множества авторов
if_oracle NUMBER DEFAULT 1, --флаг использования внутренних механизмов контекстного поиска
if_translate NUMBER default null --флаг использования переводов
)
RETURN mas_man_find PIPELINED
```

Для создания интерфейса пользователя использован механизм построения динамических отчетных форм на основе SQL-запросов ИАС "ИС-ТИНА" [22]. В интерфейсе поиска можно указать требуемые ключевые слова, область знаний по рубрике ГРНТИ, подразделение для поиска опорного множества авторов и дополнительные параметры поиска (рис. 2).

После выполнения поиска пользователю на экран выводится ранжированный список ре-

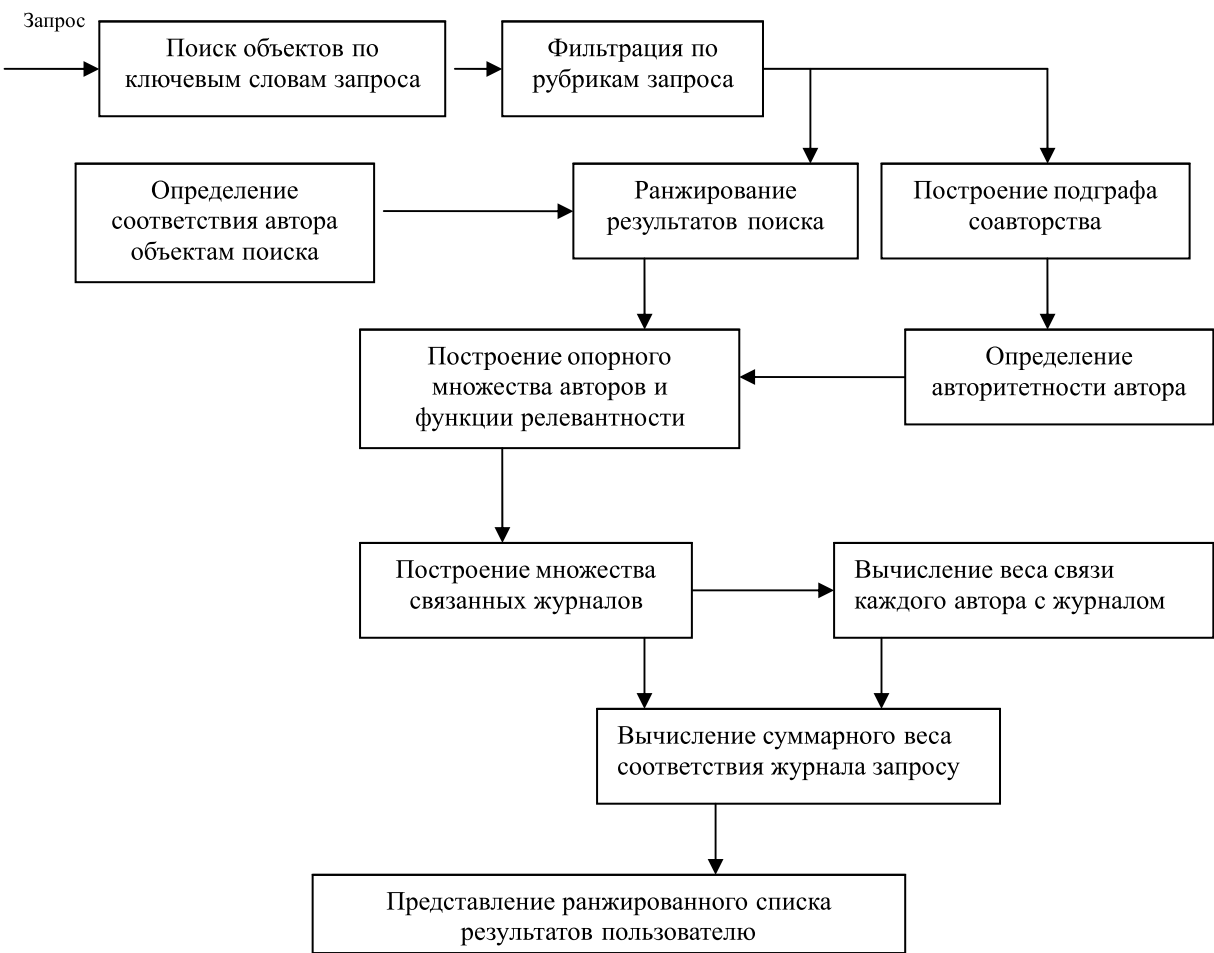


Рис. 1. Схема алгоритма тематического поиска журналов

Ключевые слова (через запятую)

титан

Область знаний

Все

Подразделение

Все

Встроенный поиск

Нет

Учет переводов

Нет

Отправить запрос

Рис. 2. Интерфейс тематического поиска журналов по ключевым словам

зультатов с указанием названия журнала "NAME" и степени его соответствия поисковому запросу "V". Для удобства навигации после нажатия на название журнала осуществляется переход в карточку журнала

с описанием его характеристик цитирования, списком последних статей и другой информацией (рис. 3).

Следует отметить, что данный алгоритм может применяться как самостоятельно, так и в сочета-

Show by 50 items

V	NAME
35.566	Доклады Академии наук
11.419	Russian Chemical Bulletin
10.832	Журнал физической химии
9.337	Вестник Московского университета. Серия 3: Физика, астрономия
9.165	Russian Journal of Physical Chemistry A
8.737	Journal of Physics: Conference Series
8.255	Вестник Московского университета. Серия 2: Химия
8.152	Известия Академии наук. Серия химическая
8.138	Mendeleeev Communications
7.623	Doklady Akademii nauk SSSR
7.612	Вестник Московского университета. Серия 4: Геология
6.912	Известия Российской академии наук. Серия физическая

Рис. 3. Интерфейс тематического поиска журналов по ключевым словам

нии с другими алгоритмами полнотекстового тематического анализа. Результатом работы алгоритма является ранжированный список журналов, соответствующий поисковому запросу пользователя.

**Тематический поиск
научно-технических конференций**

Использование данных наукометрических систем позволяет создать поисковые механизмы тематического поиска регулярных научно-технических конференций с возможностью оценки их авторитетности. В качестве основной гипотезы используется объективная гипотеза, что авторитетные ученые и эксперты посещают авторитетные конференции по своей тематике. Автором публикации разработано три алгоритма тематического поиска конференций: поиск конференций, похожих по тематике на заданный журнал; поиск конференций, похожих на заданную конференцию, поиск конференций с использованием поиска экспертов.

Поиск конференций, похожих на заданный журнал, использует в качестве опорного множества авторов публикаций журнала. После этого выполняется расчет функции определения степени тематической связи между конференциями и журналами аналогично расчету функции определения степени тематической связи между журналами (1). Алгоритм, показанный на рис. 4, применим, если пользователь знает один или несколько журналов по своей тематике.

Основой программной реализации алгоритма является вычисляемая хранимыми в базе данных запросами таблица результатов расчета близости журналов и конференций:

```
JOURNAL_CONF_THEM_LINK
(CNT      NUMBER --Степень соответствия
JOURNAL_ID NUMBER --Уникальный ключ журнала
CONF_ID   NUMBER --Уникальный ключ конференции
)
```

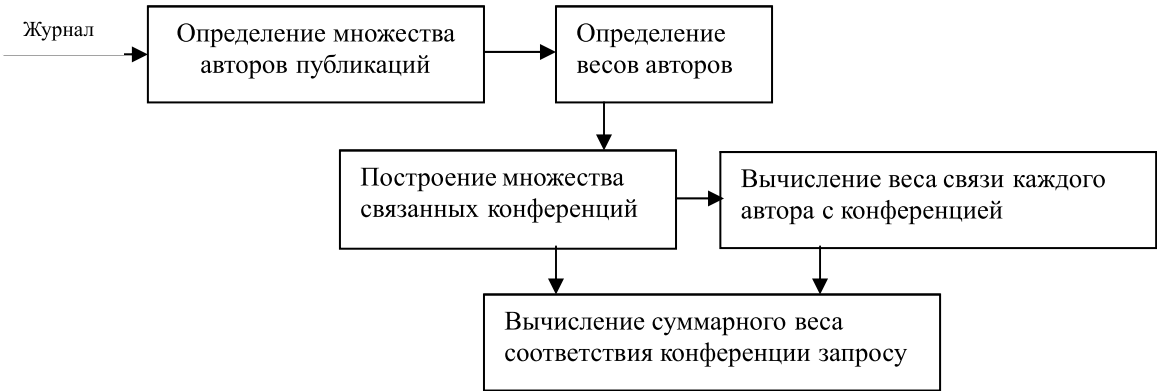


Рис. 4. Алгоритм поиска конференций по журналу, в котором публикуется автор

Использование такого подхода позволяет заранее рассчитать граф соответствия между журналами и конференциями для ускорения выполнения пользовательских запросов.

Поиск конференций, похожих на заданную конференцию, использует в качестве опорного множества докладчиков заданной конференции. После этого также проводится расчет функции *F*. Алгоритм, представленный на рис. 5, применим, если пользователь уже знает одну или несколько конференций по своей тематике и хочет расширить круг своего научного общения.

Основой программной реализации алгоритма является вычисляемая хранимыми в базе данных запросами таблица результатов расчета близости конференций:

```
CONF_THEM_LINK
(CNT      NUMBER --Степень соответствия
CONF_ID1 NUMBER --Уникальный ключ конференции
CONF_ID2 NUMBER --Уникальный ключ связанной
)
```

При поиске конференций с использованием методов и средств поиска экспертов на первом этапе обработки запроса, как и при тематическом поиске журналов, строится функция релевантности авторов поисковому запросу $v(a, q)$ и определяется множество авторов, релевантных запросу. После этого рассчитывается функция тематического соответствия F^* . Таким образом, алгоритм, приведенный на рис. 6, находит конференции, которые чаще всего посещают наиболее авторитетные специалисты по заданной в запросе тематике.

Для удобной визуализации результатов поиска разработан специальный интерфейс пользователя. На рис. 7 представлен пример отображения результата поиска конференций, тематически близких к журналу "Нейрокомпьютеры: разработка, применение".

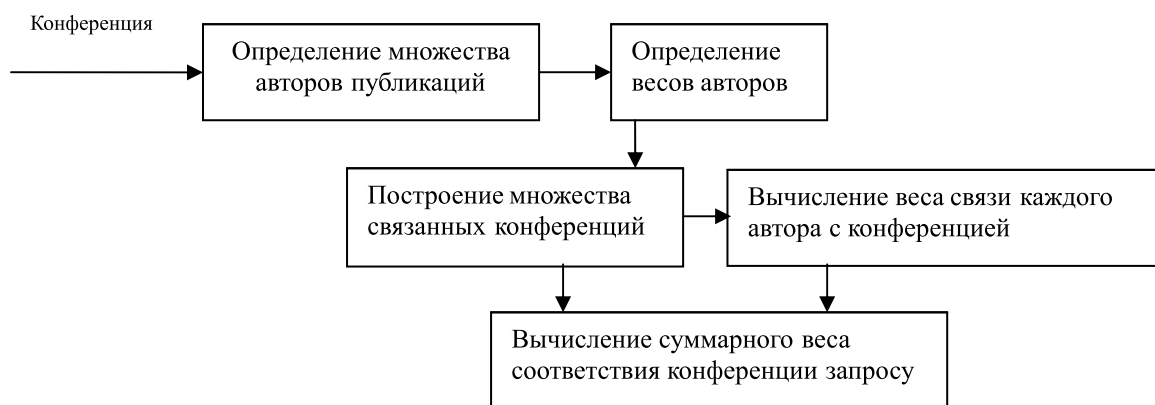


Рис. 5. Алгоритм поиска конференций по заранее выбранной пользователем конференции



Рис. 6. Алгоритм поиска конференций с использованием тематического поиска экспертов

В интерфейсе отображается название найденной конференции, степень соответствия, число загруженных в наукометрическую систему докладов с этой конференции и ссылки для дальнейшего поиска близких конференций и журналов, для осуществления которого можно нажать на

слова "конференции" и "журналы" в соответствующей строке. На рис. 8 представлены результаты перехода по одной из этих ссылок для поиска конференций, близких по тематике к конференции "XVIII Всероссийская научная конференция "Нейрокомпьютеры и их применение".

N	Конференция	Вес	Количество докладов	Похожие конференции	Похожие журналы
1	Ломоносовские чтения 2020. Секция философских наук(2020)	2,5	82	конференции	журналы
2	CAICS 2020: NATIONAL CONGRESS ON COGNITIVE RESEARCH, ARTIFICIAL INTELLIGENCE AND NEUROINFORMATICS(2020)	2,42	19	конференции	журналы
3	I Национальный Конгресс по когнитивным исследованиям, искусственному интеллекту и нейроинформатике(2020)	2,38	78	конференции	журналы
4	XVIII Всероссийская научная конференция "Нейрокомпьютеры и их применение". Москва. 17 марта 2020(2020)	2,08	4	конференции	журналы
5	61-е заседание международного междисциплинарного семинара "Нейрофилософия" Научного совета по методологии исследований искусственного интеллекта при Президиуме РАН(2021)	2	2	конференции	журналы
6	Нейронаука для медицины и психологии: XVII Международный междисциплинарный конгресс.(2021)	1,93	18	конференции	журналы
7	XII Международная научная конференция «Интеллектуальные системы и компьютерные науки»(2021)	1,92	77	конференции	журналы
8	Ломоносовские чтения - 2021, секция "Философские науки"(2021)	1,5	77	конференции	журналы
9	Международная междисциплинарная конференция «Искусственный интеллект в новой коммуникативной реальности» (2020)	1,25	17	конференции	журналы
10	Нейроинформатика 2021(2021)	1,17	3	конференции	журналы

Рис. 7. Результаты поиска для журнала "Нейрокомпьютеры: разработка, применение"

Список похожих конференций

Show by 10 items Search:

N	Название конференции	Вес	Количество докладов	Похожие конференции	Похожие журналы
1	I Национальный Конгресс по когнитивным исследованиям, искусственному интеллекту и нейроинформатике(2020)	0,75	78	конференции	журналы
2	CAICS 2020: NATIONAL CONGRESS ON COGNITIVE RESEARCH, ARTIFICIAL INTELLIGENCE AND NEUROINFORMATICS(2020)	0,67	19	конференции	журналы
3	VI Международная конференция и молодёжная школа «Информационные технологии и нанотехнологии» (ИТНТ-2020)(2020)	0,5	4	конференции	журналы
4	Нейронаука для медицины и психологии: XVII Международный междисциплинарный конгресс.(2021)	0,33	18	конференции	журналы

Items 1 to 4 shown. Total: 4

First Previous 1 Next Last

Рис. 8. Результаты поиска для конференции "XVIII Всероссийская научная конференция Нейрокомпьютеры и их применение"

Апробация алгоритмов проводилась на данных наукометрической системы ИАС "ИСТИНА". Для тестирования было сделано по пять поисковых запросов для каждой из трех программных реализаций представленных алгоритмов. Лучшие десять результатов для каждого запроса оценивались по следующей шкале: 2 — полное соответствие запросу; 1 — частичное соответствие запросу; 0 — несоответствие запросу. Интегральная оценка для каждой программной реализации рассчитывалась как сумма набранных баллов. Результаты тестирования представлены ниже.

Алгоритм.....	Точность, %
Поиск по журналу	81
Поиск по конференции	76
Использование поиска экспертов.....	89

Представленные результаты тестирования показывают, что наиболее точным способом поиска является использование алгоритма с тематическим поиском экспертов. Поиск по журналу имеет преимущество перед поиском похожих конференций вследствие более представительного опорного множества авторов.

Реализация подобных механизмов значительно упрощает для пользователей процесс информационного поиска конференций, соответствующих их информационным потребностям.

Заключение

Представленные в работе алгоритмы позволяют проводить тематический поиск журналов и конференций с использованием данных наукометрических систем. Представленные алгоритмы могут успешно применяться для поиска объектов на других языках. Применение этих алгоритмов возможно, как автономно, так и в сочетании с классическими алгоритмами полнотекстового тематического анализа. В отличие от полнотекстового поиска алгоритмы учитывают авторитетность конференций.

Использование представленных алгоритмов тематического анализа информации позволяет: проводить нормализацию наукометрических показателей в зависимости от тематики научных исследований; проводить выделение и анализ основных тематических направлений развития науки; осуществлять тематический поиск объектов (журналов и конференций) в целях расширения научных контактов и создания новых научных коллективов.

Список литературы

1. **Интеллектуальная** система тематического исследования научно-технической информации (ИСТИНА) / В. А. Садовничий, В. А. Васенин и др. М.: Изд-во Московского ун-та, 2014. 262 с.
2. **Орлов А. И.** Наукометрия и управление научной деятельностью // Управление большими системами. Специальный выпуск 44: Наукометрия и экспертиза в управлении наукой. Институт проблем управления им. В. А. Трапезникова РАН. 2013. С. 538—568.
3. **Полянин А. Д.** Недостатки индексов цитируемости и Хирша. Индексы максимальной цитируемости. URL: http://eqworld.ipmnet.ru/ru/info/sci-edu/Polyanin_IndexH_2014.html
4. **Бричковский В. В.** Наукометрический анализ в информационном обеспечении инновационной деятельности // В мире науки. 2017. № 8 (174). С. 64—67.
5. **Афонин С. А., Козицын А. С., Шачнев Д. А.** Программные механизмы агрегации данных, основанные на онтологическом представлении структуры реляционной базы наукометрических данных // Программная инженерия. 2016. Том 7, № 9. С. 408—413.
6. **Witten I. H., Frank E., Hall M. A., Pal C. J.** Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016. 558 p.
7. **Васенин В. А., Занчурин М. А., Козицын А. С., Кривчиков М. А., Шачнев Д. А.** Архитектурно-технологические аспекты разработки и сопровождения больших информационно-аналитических систем в сфере науки и образования // Программная инженерия. 2017. Том 8, № 10. С. 448—455.
8. **Vasenin V. A., Zanchurin M. A., Kozitsyn A. S., Krivchikov M. A., Shachnev D. A.** Architectural and technological aspects of the cloud data analysis system development, case of istina system // CEUR Workshop Proceedings. 5. Ser. "APSSE 2017 — Proceedings of the 5th International Conference on Actual Problems of System and Software Engineering" 2017. P. 90—96.
9. **Анализ** статистических данных в системе ИСТИНА. URL: <https://istina.msu.ru/statistics/organization/214524/dynamic/>
10. **Васенин В. А., Афонин С. А., Козицын А. С.** Автоматизированная система тематического анализа информации // Информационные технологии. 2009. № 7. Приложение. С. 2—32.
11. **Садовничий В. А., Васенин В. А., Афонин С. А., Козицын А. С., Голомазов Д. Д.** Информационная система "ИСТИНА" как big data - инструментарий в области управления на основе анализа наукометрических данных // Знания — Онтологии — Теории (ЗОНТ-2015) Материалы Всероссийской конференции с международным участием. Новосибирск, 2015. С. 115—123.
12. **Васенин В. А., Зензинов А. А., Лунев К. В.** Использование наукометрических информационно-аналитических систем для автоматизации проведения конкурсных процедур на примере информационно-аналитической системы ИСТИНА // Программная инженерия. 2016. Том 7, № 10. С. 472—480.
13. **Shachnev D. A.** Searching for activity results and experts in a given subject area, taking results significance into account // Программная инженерия. 2021. Том 12, № 5. С. 260—266.
14. **Kozitsin A., Afonin S., Shachnev D.** Algorithm for searching experts in scientometric systems // CEUR Workshop Proceedings. Proceedings of the 23rd Conference on Scientific Services & Internet (SSI 2021). Vol. 3066. CEUR, 2022. P. 59—68.
15. **Плагиат** в научных статьях: трудности обнаружения перевода. URL: http://ai-news.ru/2018/01/plagiat_v_nauchnyh_statyah_trudnosti_obnaruzheniya_perevoda.html
16. **Афонин С. А., Голомазов Д. Д., Козицын А. С.** Использование систем семантического анализа для организации поиска научно-технической информации // Программная инженерия. 2012. № 2. С. 29—34.

17. Домнина Т. Н., Хачко О. А. Научные журналы: количество, темпы роста // XIX Научно-практический семинар с международным участием "Информационное обеспечение науки: новые технологии". 2015. С. 83–96.

18. De Solla Price D. D. Little Science, Big Science. New York, 1963. 109 p.

19. Козицын А. С., Афонин С. А., Шачнев Д. А. Метод оценки тематической близости научных журналов // Программная инженерия. 2020. Том 11, № 6. С. 335–341.

20. Краснов Ф. В. Анализ методов построения графа соавторства: подход на основе двудольного графа // International Journal of Open Information Technologies. 2018. Vol. 6, № 2. P. 31–37.

21. Статистика по наличию журналов РИНЦ. URL: https://www.elibrary.ru/project_free_access.asp

22. Afonin S., Kozitsyn A., Astapov I. Sqlreports: Yet another relational database reporting system // Proceedings of the 9th International Conference on Software Engineering and Applications. 2014. P. 529–534.

Thematic Search Algorithms for Experts, Journals and Conferences

A. S. Kozitsyn, alexanderkz@mail.ru, Lomonosov Moscow State University, Moscow, 119192, Russian Federation

Corresponding author:

Kozitsyn Alexander S., Leading Researcher, Lomonosov Moscow State University, 119192, Moscow, Russian Federation

E-mail: alexanderkz@mail.ru

Received on April 02, 2022

Accepted on April 14, 2022

The paper discusses methods for thematic search of experts, journals and conferences, provides a brief overview of existing systems with an indication of their main shortcomings, and describes the algorithms for the thematic search of experts developed by the authors, as well as algorithms for thematic search of journals and conferences. A brief description of the software implementation of the developed algorithms is given. User interfaces for searching for conferences are presented. The results of testing software implementations of algorithms based on the data of the scientometric system "ISTINA" are presented. The algorithms developed by the authors use data from scientometric systems, including graphs of mutual relations between objects, are able to filter data by narrowly focused thematic queries and rank search results by authority. The approach used makes it possible to perform a thematic search for conferences in different languages with the determination of their authority and the determination of their thematic proximity. Such thematic analysis makes it possible to normalize scientometric indicators depending on the subject of scientific research, to identify and analyze the main thematic directions in the development of science, to carry out a thematic search for objects (journals and conferences) in order to expand scientific contacts and create new research teams.

Keywords: thematic analysis, scientometrics, information systems, bibliography, author, graph

For citation:

Kozitsyn A. S. Thematic Search Algorithms for Experts, Journals and Conferences, *Programmnaya Ingeneria*, 2022, vol. 13, no. 6, pp. 291–300.

DOI: 10.17587/prin.13.291-300

References

1. *Intelligent system of case study of scientific and technical information (ISTINA)* / V. A. Sadovnichy, V. A. Vasenin et al. Moscow, Moscow University Press, 2014. 262 p. (in Russian).

2. Orlov A. I. Scientometrics and management of scientific activities, *Management of large systems. Special Issue 44: Scientometrics and Expertise in Science Management*, Institute of Management Problems V. A. Trapeznikova RAS, 2013, pp. 538–568 (in Russian).

3. Polljanin A. Disadvantages of citation and hirsch indices. Maximum Citation Indices, available at: http://eqworld.ipmnet.ru/ru/info/sci-edu/polyanin_indexh_2014.html (in Russian).

4. Brichkovsky V. V. Scientometric analysis in the information support of innovative activity, *V mire nauki*, 2017, no. 8 (174), pp. 64–67 (in Russian).

5. Afonin S. A., Kozitsyn A. S., Shachnev D. A. Software data aggregation mechanisms based on the ontological representation of the structure of the relational database of scientometric data, *Programmnaya Ingeneria*, 2016, vol. 7, no. 9. pp. 408–413 (in Russian).

6. Witten I. H., Frank E., Hall M. A., Pal C. J. *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann, 2016, 558 p.
7. Vasenin V. A., Zanchurin M. A., Kozitsyn A. S., Krivchikov M. A., Shachnev D. A. Architectural and Technological Aspects of the Development and Maintenance of Large Information Analysis Systems in the Area of Science and Education, *Programmnaya Ingeneria*, 2017, vol. 8, no. 10, pp. 448–455 (in Russian).
8. Vasenin V. A., Zanchurin M. A., Kozitsyn A. S., Krivchikov M. A., Shachnev D. A. Architectural and technological aspects of the cloud data analysis system development, case of istina system, *CEUR Workshop Proceedings. 5. Cep. "APSSE 2017 — Proceedings of the 5th International Conference on Actual Problems of System and Software Engineering"* 2017, pp. 90–96.
9. Analysis of statistical data in the system ISTINA, available at: <http://gis-lab.info/qa/srtm.html> (in Russian).
10. Vasenin V. A., Afonin S. A., Kozitsyn A. S. Avtomatizirovannaya sistema tematicheskogo analiza informatsii, *Informatsionnye tekhnologii*, 2009, no. 7, supl., pp. 2–32 (in Russian).
11. Sadovnichy V. A., Vasenin V. A., Afonin S. A., Kozitsyn A. S., Golomazov D. D. Information system "ISTINA" as big data — a tool in the field of control based on the analysis of scientometric data, *Knowledge — Ontologies — Theories (ZONT-2015) Materials of the All-Russian Conference with international participation*, Novosibirsk, 2015, pp. 115–123 (in Russian).
12. Vasenin V. A., Zenzinov A. A., Lunev K. V. Using scientometric information-analytical systems to automate competitive procedures using the example of the information-analytical system ISTINA, *Programmnaya Ingeneria*, 2016, vol. 7, no. 10, pp. 472–480 (in Russian).
13. Shachnev D. A. Searching for activity results and experts in a given subject area, taking results significance into account, *Programmnaya Ingeneria*, 2021, vol. 12, no. 5, pp. 260–266.
14. Kozitsin A., Afonin S., Shachnev D. Algorithm for searching experts in scientometric systems, *CEUR Workshop Proceedings. Proceedings of the 23rd Conference on Scientific Services & Internet (SSI 2021)*, vol. 3066, CEUR, 2022, pp. 59–68.
15. Plagiat v nauchnykh statiyakh: trudnosti obnaruzheniya perevoda, available at: http://ai-news.ru/2018/01/plagiat_v_nauchnykh_statyakh_trudnosti_obnaruzheniya_perevoda.html (in Russian).
16. Afonin S. A., Golomazov D. D., Kozitsyn A. S. The use of semantic analysis systems to organize the search for scientific and technical information, *Programmnaya Ingeneria*, 2012, no. 2, pp. 29–34 (in Russian).
17. Domnina T. N., Khachko O. A. Nauchnye zhurnaly: kolichestvo, tempy rosta, *XIX Nauchno-prakticheskii seminar s mezhdunarodnym uchastiem "Informatsionnoe obespechenie nauki: novye tekhnologii"*, 2015, pp. 83–96 (in Russian).
18. De Solla Price D. D. Little Science, Big Science. New York, 1963, 109 p.
19. Kozitsyn A. S., Afonin S. A., Shachnev D. A. Metod otsenki tematicheskoi blizosti nauchnykh zhurnalov, *Programmnaya Ingeneria*, 2020, vol. 11, no. 6, pp. 335–341 (in Russian).
20. Krasnov F. V. Analysis of co-authorship graph construction methods: bipartite graph approach, *International Journal of Open Information Technologies*, 2018, vol. 6, no. 2, pp. 31–37 (in Russian).
21. Statistika po nalichiiu zhurnalov RINTS, available at: https://www.elibrary.ru/project_free_access.asp (in Russian).
22. Afonin S., Kozitsyn A., Astapov I. Sqlreports: Yet another relational database reporting system, *Proceedings of the 9th International Conference on Software Engineering and Applications*, 2014, pp. 529–534.

ИНФОРМАЦИЯ

В рамках Международного конгресса "Суперкомпьютерные дни в России" Нижегородский государственный университет им. Н. И. Лобачевского проводит международную конференцию и молодежную школу "Математическое моделирование и суперкомпьютерные технологии". Проведение конференции (14–15 ноября 2022 г.) и молодежной школы (16–18 ноября 2022 г.) будет содействовать дальнейшему повышению синергетического эффекта от взаимовыгодного взаимодействия двух приоритетных направлений развития науки и техники – математического моделирования и суперкомпьютерных технологий.

В программе конференции – лекции приглашенных экспертов, доклады известных ученых и специалистов, сообщения молодых исследователей. Школа будет включать мастер-классы и практикумы по применению актуальных инструментов и технологий.

Конференция и молодежная школа проводятся под эгидой **Суперкомпьютерного консорциума университетов России** при партнерстве с Национальным центром НТИ «Технологии машинного обучения и когнитивные технологии» (СПбГУ ИТМО) и Национальным центром НТИ «Технологии хранения и анализа больших данных» (МГУ).

Подробности: <http://agora.guru.ru/hpc2022/>