



УДК 004.021, 004.047, 004.622

Д. А. Вареников, М. Д. Шлей
Санкт-Петербургский национальный исследовательский
университет информационных технологий, механики и оптики

А. И. Водяхо, Н. А. Жукова,
Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Система информационных моделей поддержки научной составляющей образовательных процессов

Предлагается система интеллектуальной поддержки принятия решений по реализации научной составляющей в образовательных процессах в университетах в аспирантуре и магистратуре. Основное внимание уделяется проблеме формирования рекомендаций для участников образовательного процесса по выбору периодических изданий. В основу системы положена информационная модель научных интересов участников образовательного процесса, методы ее наполнения и использования. Система разрабатывалась как сервис, предназначенный для интеграции в информационные системы университетов. Показаны результаты работы системы при формировании рекомендаций по выбору изданий по заданным ключевым словам и тематикам для публикации результатов исследований обучающимися.

Научные интересы, научная деятельность, система поддержки принятия решений, информационные модели

В настоящее время проводится активная модернизация системы высшего образования с целью повышения ее качества. Значительное внимание уделяется вопросу интеграции научной, инновационной и образовательной деятельности, в связи с чем в различных прикладных областях возникает необходимость использовать наукоемкие технологии. В ближайшей перспективе новые технологии должны дополнить, а где-то и заменить традиционно используемые. Подобная ситуация уже наблюдалась в части внедрения новых технических средств. Сейчас она прослеживается в части внедрения IT-технологий. Даже в традиционно консервативных областях, например культуры и медицины, специалисты приходят к пониманию, что реализация лечебных и управленческих процессов невозможна без внедрения современных информационных систем.

Задачи использования наукоемких технологий в образовании и привлечения обучающихся к научной деятельности отнесены к ведущим

направлениям в соответствии с программой Горизонт 2020 [1]. Для выполнения поставленных в программе задач требуется привлечение обучающихся к научной деятельности на протяжении всего обучения, начиная с бакалавриата и заканчивая аспирантурой. Важно обеспечить:

- организацию коллабораций в рамках проектной и исследовательской деятельности;
- регулярное участие в значимых мероприятиях, в том числе на международном уровне;
- доступ к релевантным информационным источникам, в частности, статьям, журналам;
- привлечение к публикационной деятельности, в том числе публикацию научных результатов в высокорейтинговых изданиях;
- участие в грантах и конкурсах.

Поставленные задачи должны решаться информационными системами университета (ИСУ). Своевременное предоставление актуальной информации участниками образовательного процесса по направлению их научной деятельности яв-

ляется одним из основных требований, предъявляемых к информационным системам университета. В статье детально рассматривается одна из задач, являющаяся высоковольтребованной, особенно в технических университетах, – задача поиска и предоставления информации, необходимой для научной деятельности обучающихся [2].

Наука и образование в современном университете. В университетах научно-исследовательская деятельность ведется в рамках кафедр и/или лабораторий, созданных на их основе. Контроль за научно-исследовательской деятельностью магистрантов/аспирантов целиком и полностью ложится на научного руководителя. В связи с большой загруженностью научный руководитель не всегда имеет возможность своевременно проконтролировать и направить магистранта/аспиранта, а у магистранта/аспиранта зачастую не хватает опыта и необходимой информации.

В настоящее время существующими информационными системами накоплены большие объемы данных о деятельности университетов [3]. Для поиска информации в системах необходимо: сформировать четкие требования, которые варьируются в зависимости от целей и задач научных исследований, определить множество критериев оценки результатов и т. д. Такая задача предполагает наличие специализированных навыков, опыта работы с системой. Кроме того, поисковые системы значительной части ИСУ ВУЗ не всегда позволяют осуществить полномасштабный поиск. Для решения проблемы нехватки информации в условиях всеобщей информационной избыточности необходимо выполнять расширенный поиск по периодическим изданиям, проектам, а также научным конкурсам средствами ИСУ ВУЗ.

Текущие достижения технологий IT и опыт их широкого внедрения и использования в таких системах, как АСУ-Аналитика, STAT-RESEARCH, SAP, MetaStock Pro, CBR Express и CasePoint, показывают, что задача поиска может быть эффективно решена с использованием методов интеллектуальной поддержки принятия решений.

Таким образом, задача внедрения научной деятельности в образовательные процессы сводится к расширению ИСУ за счет разработки и интеграции нового сервиса, который представляет собой систему интеллектуальной поддержки принятия решений.

Модели и технологии для естественно-научного образования. Процесс поиска можно разделить на 2 этапа. На первом этапе должен

осуществляться отбор альтернатив на уровне ключевых слов, тематик или научных интересов. На втором этапе выполняется анализ и отбор на основе критериев, сформированных участниками образовательного процесса. Операции поиска, анализа и отбора основаны на использовании информационных моделей пользователей ИСУ, публикаций, мероприятий.

Модель пользователя информационной системы университета. Для реализации процедуры поиска была разработана модель пользователя информационной системы университета, содержащая общую информацию о пользователе, описание его научно-практической активности и научных интересов [4]:

$$P = \langle D, P_a, P_p, M, K \rangle,$$

где D – множество описательных характеристик пользователя, включая ученую степень, ученое звание, специальность, владение языком, сведения об обучении, научное руководство; P_a – множество публикаций пользователя; P_p – множество научных проектов, в которых пользователь принимал участие или являлся руководителем; M – множество научных мероприятий, в которых пользователь принимал непосредственное участие; K – множество научных интересов пользователя [5]. Здесь

$$K = K^P \cup K^a,$$

где K^P – множество ключевых слов, указанных пользователем; K^a – множество ключевых слов, автоматически выбранных с учетом частоты их появления в информационной системе:

$$K^a = \{q \in Q : f(q) > c\}, \quad Q = \bigcup_{i=1}^h Q_i;$$

Q – множество автоматически полученных ключевых слов; $f(q)$ – частота появления ключевого слова q ; c – пороговое значение для частоты появления ключевого слова; h – количество источников, на основании которых формируется множество Q . Множество Q формируется за счет:

1) анализа посещения пользователем информационной системы; Q_1 – множество ключевых слов, полученных по результатам посещения страниц: $Q_1 = \bigcup_{j=1}^m L_j$, где L_j – множество ключе-

вых слов j -й страницы, содержащей ключевые слова и тематики; m – количество страниц;

2) анализа схожих интересов между пользователями, посетившими одинаковые страницы; Q_2 – множество ключевых слов пользователей со схожими интересами:

$$Q_2 = \left\{ K_v : v \in \bigcup_{j=1}^m V_j, \text{Jaccard}(K_v, K^P) > s \right\},$$

где K_v – множество ключевых слов пользователя; v , V_j – множество пользователей, посетивших j -ю страницу, за исключением рассматриваемого

пользователя; $\text{Jaccard}(K_v, K^P) = \frac{|K_v \cap K^P|}{|K_v \cup K^P|}$ – мера

Жаккара ($0 \leq \text{Jaccard}(K_v, K^P) \leq 1$) [5]; s – пороговое значение схожести;

3) анализа тематик публикаций, автором которых является пользователь; Q_3 – множество ключевых слов, полученных на основе анализа публикационной активности пользователя:

$$Q_3 = \bigcup_{j=1}^g Kp_j, \text{ где } Kp_j \text{ – множество ключевых}$$

слов j -й публикации; g – количество публикаций пользователя;

4) анализа схожести интересов между соавторами публикаций, автором которых является пользователь; Q_4 – множество ключевых слов соавторов публикаций пользователя: $Q_4 =$

$$\left\{ K_{ca} : c_a \in \bigcup_{j=1}^g Ca_j, \text{Jaccard}(K_{ca}, K^P) > s \right\}, \text{ где}$$

K_{ca} – множество ключевых слов пользователя p_{ca} ; Ca_j – множество соавторов публикации j , за исключением рассматриваемого пользователя; $\text{Jaccard}(K_{ca}, K^P)$ – мера Жаккара; s – пороговое значение схожести [4];

5) получения сведений о подписке пользователя на рассылки в информационной системе; Q_5 – множество ключевых слов, указанных пользователям для получения рассылок в информационной системе;

6) анализа профиля пользователя в наукометрических базах данных; Q_6 – множество ключевых слов пользователя, полученных из наукометриче-

ских баз данных: $Q_6 = \left\{ K_{apr} : a_{pr} \in \bigcup_{j=1}^d Apr_j \right\}$, где

K_{apr} – множество ключевых слов пользовательского профиля a_{pr} ; Apr_j – множество авторских профилей наукометрической базы данных j для рассматриваемого пользователя; d – количество наукометрических баз данных.

В системе учитываются наиболее распространенные наукометрические базы данных и их идентификаторы авторских профилей [6]:

– РИНЦ (русский индекс научного цитирования) – используется уникальный идентификатор SPIN-код;

– Web of Science – база данных по научному цитированию института научной информации (Institute of Scientific Information – ISI) – используемый уникальный идентификатор ResearcherID;

– Scopus – единая мультидисциплинарная реферативная база данных, представляющая уникальную систему оценки частоты цитирования. Используемый уникальный идентификатор ORCID.

Авторский профиль из наукометрических баз данных представляется следующим образом [6]:

$$Apr = \langle K_{apr}, P_{pr}, Ind_{pr} \rangle,$$

где P_{pr} – множество публикаций авторского профиля; Ind_{pr} – идентификатор авторского профиля.

Модель публикации имеет вид

$$At = \langle Iz, Aat, Kat \rangle,$$

где Iz – профиль издания; Aat – множество авторов публикаций; Kat – множество тематических характеристик публикации;

$$Iz = \langle Diz, Niz, Siz, Pbiz, Kiz \rangle;$$

Diz – множество описательных характеристик издания, таких, как место издания, издательство, вид издания, сайт издания, язык издания, вид материалов, ISSN, ISBN, e-ISSN; Niz – множество наукометрических показателей (IF, IPP, SJR, SNIP); Siz – множество показателей статусов публикаций согласно индексированию в наукометрических базах данных; $Pbiz$ – множество категорий авторов, публикующихся в данном издании, таких, как студенты, аспиранты, сотрудники, кандидаты наук, доктора наук, иностранные сотрудники; Kiz – множество тематических характеристик издания.

Модель научного мероприятия представляется следующим образом:

$$Kf = \langle Dkf, Pkf, Pkfv, Kkf \rangle,$$

где Dkf – множество описательных характеристик научных мероприятий, таких, как сумма финансирования, страна, язык мероприятия, сроки проведения, сайт мероприятия; Pkf – множество категорий пользователей, участвовавших в данном мероприятии; $Pkfv$ – множество категорий пользователей, посетивших данное мероприятие; Kkf – множество тематических характеристик научных мероприятий.

Алгоритм интеллектуализации поиска периодических изданий по заданным ключевым словам и тематикам.

Входные данные: множество тематик

$T = \{t_y\}_{y=1}^k$, k – количество тематик, описывающих область исследования; множество критериев $C = \{c_j\}_{j=1}^m$, m – количество критериев и веса v_j для каждого c_j .

Выходные данные: множество вариантов периодических изданий $R = \langle A, U, D_a \rangle$, где A – множество периодических изданий, подходящих по тематикам T ; U – рассчитанная функция полезности; D_a – дополнительная информация, обосновывающая найденные периодические издания A .

1. Из всех I_z выбирается множество таких периодических изданий A , у которых тематические характеристики Kiz совпадают с тематиками поиска T ; $A = \{a_i\}_{i=1}^n$, где n – количество периодических изданий.

2. Оценка периодических изданий A по критериям C , где $X = \{x_{ij}\}, i = [1..n], j = [1..m]$ – множество оценок.

3. Нормирование множества оценок X :

$$\text{ЕСЛИ } x_{ij} = 0$$

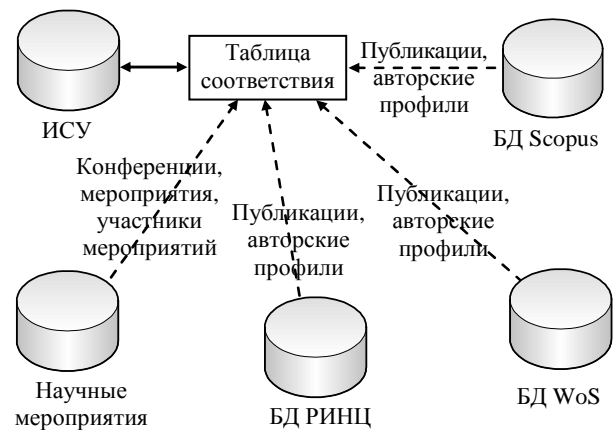
$$\text{ТО нормирование значения } x'_{ij} = 0$$

$$\text{ИНАЧЕ } x'_{ij} = \frac{x_{ij}}{w_j}, \text{ где } w_j = \sum_{i=1}^n x_{ij}.$$

4. Вычисление функции полезности для каждого периодического издания a_i : $U_i = \sum_{j=1}^m x'_{ij}$.

5 Формирование множества R на основе результатов ранжирования функции полезности по убыванию.

Описание и внедрение СППР. Система ППР представляет собой сервис, который интегрируется в ИСУ. Типовая ИСУ – это комплексная система, которая состоит из четырех базовых компонентов: административного, учебного, проектного и финансово-хозяйственного комплексов. В части проектного комплекса, как правило, поддерживается взаимодействие с наукометрическими базами данных и базами научных мероприятий, что расширяет информационное пространство для реализации ППР. На рисунке рассмотрена ИСУ, построенная на базе платформы Oracle Application Express и СУБД Oracle.



Система ППР интегрируется в ИСУ в части проектного комплекса. Разработанная система представляет собой набор программных модулей, которые подключаются к подсистемам ИСУ. Вызов и работа с модулями реализуются подключением соответствующих библиотек. Интерфейс управления дает возможность пользователям задавать входные параметры, которые определяют критерии отбора, например ключевые слова, определяющие тематическую область поиска. Посредством интерфейса управления пользователям представляется сформированная рекомендация по информационным источникам с обоснованием каждого предложенного варианта. Модуль обработки входных данных реализует алгоритм интеллектуального поиска, приведенный ранее.

Результаты экспериментальных исследований. Работа инструмента формирования рекомендаций демонстрируется на примере поиска иностранных журналов для публикации материалов по

Критерий		Результаты вычислений критериев					
Наименование	Вес	Journal of Intelligent Systems	International Journal of General Systems	Journal of Automation and Information Sciences	Chemometrics and Intelligent Laboratory Systems	Problems of Information Transmission	IEEE Transactions on Information Theory
Количество заданных тематик, совпадающих с тематиками журнала	10	0.23	0.15	0.15	0.15	0.15	0.15
Наличие журнала в Scopus	8	0.17	0.17	0.17	0.17	0.17	0.17
Наличие журнала в Web of Science	8	0	0.2	0.2	0.2	0.2	0.2
Наличие электронной версии журнала	2	0.2	0.2	0.2	0	0.2	0.2
IF (импакт-фактор)	9	0	0.13	0.06	0.37	0.06	0.37
SJR	7	0.03	0.13	0.05	0.13	0.06	0.59
Кол-во публикаций сотрудников и аспирантов	4	0.04	0.08	0.17	0.04	0.17	0.5
Кол-во авторов (сотрудников и аспирантов)	3	0.2	0.05	0.15	0.15	0.20	0.25
Значение функции полезности	–	4.98	7.42	6.91	9.40	7.12	15.16

итогах научных исследований. Для формирования рекомендаций использовался следующий набор ключевых слов, описывающих тематическую область, – «information systems, software, computer science applications, artificial intelligence».

Первый этап анализа позволил выявить 38 журналов, у которых ключевые слова и тематики схожи с описанием, заданным в запросе. Для наглядной демонстрации работы алгоритма были выбраны 6 журналов, наиболее схожих по тематике с заданной.

Для проведения второго этапа анализа, позволяющего получить более качественные оценки альтернатив, были сформулированы критерии отбора и определены веса критериев (таблица, столбец «Критерий»).

В столбце «Результаты вычислений критериев» показаны приведенные матрицы решений и значение функции полезности для соответствующих журналов.

Экспертный анализ результатов работы системы показал, что дополнительно полученная

информация о журналах, соответствующих тематике исследований, полезна обучающимся при выборе журнала для публикаций научных статей.

Результаты экспериментальных исследований системы показывают, что предложенный подход, основанный на использовании информационной модели участников, позволяет решить различные актуальные задачи научно-практической деятельности, такие, как выбор периодических изданий для публикаций научных результатов, выбор научных мероприятий для участия и презентации научно-исследовательской работы, выбор научных проектов для участия, а также формирование проектных команд в современном вузе.

Перспективным направлением развития предложенного решения является использование системы СППР не ограничиваясь рамками одного университета. Для этого необходимо обеспечить интеграцию информационных систем различных вузов. В настоящее время многие системы уже объединены, и можно ожидать, что в ближайшее время ИСУ объединятся в единую сеть.

СПИСОК ЛИТЕРАТУРЫ

1. Horizon 2020. URL: <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-education>. свободный.

2. Университет ИТМО сегодня. URL: http://www.ifmo.ru/ru/page/207/ob_universitete.htm. свободный.

3. Система поддержки проектной деятельности в Университете ИТМО / Ф. А. Казин, А. С. Биккулов, А. Н. Зленко, Н. Р. Тойвонен, И. А. Попова, М. Д. Шлей, Д. А. Вареников // Инновации. 2014. № 8 (190). С. 77–83.

4. Вареников Д. А., Шлей М. Д., Муромцев Д. И. Построение научных профилей участников научно-образовательного процесса в информационной системе университета // Современные проблемы науки и образования. 2015. № 2. С. 178.

5. Ефимов М. Н., Шлей М. Д., Вареников Д. А. Разработка системы формирования рекомендаций для пользователей корпоративного портала университета // Фундаментальные исследования. 2015. № 12. С. 688–692.

6. Вареников Д. А., Муромцев Д. И., Шлей М. Д. Подходы автоматизации обработки данных научных баз данных // Компьютерные инструменты в образовании. 2015. № 2. С. 3–13.

D. A. Varenikov, M. D. Shley

Saint Petersburg National Research University of Information Technologies, Mechanics and Optics

A. I. Vodyaho, N. A. Zhukova

Saint Petersburg Electrotechnical University «LETI»

SYSTEM OF INFORMATION MODELS FOR SUPPORTING SCIENTIFIC COMPONENT OF EDUCATIONAL PROCESSES

Decision support system for supporting scientific component of educational processes in universities, which is to be used by master and post graduate students is described. Main attention is paid to solving problem of forming recommendations for master and post graduate for choosing periodic editions, where they can publish results of researches. Suggested approach is based on information model of scientific interests of educational process participants. Methods of the model population and usage are discussed. The system is implemented as service, so the system can be easy integrated into interuniversity enterprise information system. The results of operation of the system are presented for forming recommendations for users for choosing editions for publication of researches using key words and thematic of research.

Scientific interests, scientific activity, decision support systems, information models

УДК 681.3

В. Е. Синев

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина)

Повышение уровня безопасности протокола групповой цифровой подписи, основанного на механизме маскирования открытых ключей

Впервые рассматривается задача повышения уровня безопасности протоколов групповой подписи, основанных на механизме маскирования открытых ключей подписантов и предлагается вариант ее решения, представляющий практический интерес. Разработан протокол, основанный на вычислительной трудности одновременного решения задачи факторизации и задачи дискретного логарифмирования по простому модулю.

Криптографические алгоритмы, криптографические протоколы, групповая цифровая подпись, маскирование открытых ключей

Криптографические алгоритмы и протоколы широко применяются в современных информационно-телекоммуникационных системах для защиты информации и формирования электронных цифровых подписей (ЭЦП) к электронным доку-

ментам [1]. Одним из основных требований, предъявляемых к криптосхемам, является стойкость, которая выражается вычислительной трудоемкостью W лучшего известного алгоритма взлома криптосхемы. Также важна оценка веро-