

Раздел IV. Анализ данных и управление знаниями

УДК 681.518

DOI 10.18522/2311-3103-2016-8-115122

Ю.Ю. Липко

ИССЛЕДОВАНИЕ МЕТОДОВ МОДЕЛИРОВАНИЯ РЕКОМЕНДАТЕЛЬНЫХ ВЕБ-СЕРВИСОВ ВЫБОРА ОБРАЗОВАТЕЛЬНЫХ РЕСУРСОВ ДЛЯ РЕАЛИЗАЦИИ ИНДИВИДУАЛЬНЫХ ОБРАЗОВАТЕЛЬНЫХ ТРАЕКТОРИЙ, ОБУЧАЮЩИХСЯ НА ОСНОВЕ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ*

Осуществлено исследование методов моделирования рекомендательных веб-сервисов выбора альтернатив образовательных ресурсов для реализации индивидуальных образовательных траекторий обучающихся на основе предпочтений освоивших предлагаемые образовательные ресурсы. Внедрение новых информационных технологий для реализации образовательного процесса, в частности асинхронные методы обучения, рассматривается как одно из наиболее перспективных направлений повышения качества высшего образования. Предложенный подход позволит обучающимся учитывать рекомендации уже освоивших предлагаемые веб-сервисом образовательные ресурсы, и выбирать наиболее предпочтительные с учетом выбранных образовательных траекторий. В процессе решения поставленной задачи был проведен сравнительный анализ методов поддержки выбора из альтернатив в веб-среде. Было обосновано использование метода коллаборативной фильтрации для реализации веб-сервиса выбора образовательных ресурсов для реализации индивидуальных образовательных траекторий студентов. В прикладной части работы осуществлено моделирование веб-сервиса на базе выбранного математического обеспечения и выявленных требований. Актуальность исследования обусловлена необходимостью развития информационных технологий экспертизы и сортировки контента, в которой участвуют как специалисты в исследуемой предметной области, так и пользователи, при этом альтернатива выбирается при помощи коллективной фильтрации на основе единых образовательных целей студентов при реализации индивидуальной работы. Исследуемые методы моделирования применимы для совершенствования технологий самостоятельной работы студентов, для реализации личного потенциала обучающегося с использованием веб-порталов образовательных ресурсов и организации дистанционного образования.

Информационные системы; технологии разработки информационных систем; формальное моделирование.

Yu. Yu. Lipko

STUDY OF METHODS FOR MODELING THE RECOMMENDATION WEB SERVICES OF EDUCATIONAL RESOURCES SELECTION FOR IMPLEMENTATION OF INDIVIDUAL EDUCATIONAL TRAJECTORIES BASED ON COLLABORATIVE FILTERING

Carried out is the study of methods for modeling the recommendation web services for selection of the educational resources for the purpose of implementing the individual educational trajectories of students basing on the preferences of those who mastered the proposed educational

* Работа выполнена при поддержке РФФИ при реализации проекта 15-07-04102 А «Разработка методов моделирования пользовательских данных в гибких конфигурируемых информационных системах».

resources. The introduction of new information technologies for educational process realization, such as asynchronous learning methods, is regarded as one of the most promising ways of increasing the quality of higher education. The proposed approach will allow students to take into account the recommendations of already mastered the proposed educational resources, and choose the most preferred in view of the selected educational trajectories. In the process of solving this problem a comparative analysis of the methods for support of selection from the alternatives in the web environment has been carried out. The use of the method for collaborative filtering for implementation of the web service of educational resource selection for the realization of individual educational trajectories of students has been justified. In the application part of the work the web services modeling based on the selected mathematical software and identified requirements is carried out. The relevance of the study is caused by the necessity to develop information technologies for expertise and content screening, which involve both the experts in the studied field and the users, with the alternative selected with the use of the collective filtering based on common educational goals of students when realizing the individual work. The studied modeling methods are applied to improve the technologies of students' independent work, realization of students' personal potential using web portal of educational resources and distance education.

Information systems; information systems development technology; formal modeling.

Введение. Число пользователей интернета в России с 2014 по 2015 год выросло на девять процентов и составило восемьдесят два миллиона человек (по данным Российской ассоциации электронных коммуникаций), это создает предпосылки для реализации новых возможностей обеспечения информацией для принятия решений в различных областях [1, 2]. Одним из направлений совершенствования веб-технологий направлено на развития возможностей использования контента, превосходящего возможности инстанций, которые осуществляют экспертизу в различных предметных областях [3].

Постановка задачи. В качестве исследуемой предметной области была выбрана индивидуальная траектория образования, которая направлена на реализацию личностного потенциала студентов. Предложенный подход позволит выбрать наиболее эффективных образовательных ресурсов с учетом предпочтений и образовательных целей студентов при реализации индивидуальной работы.

Сравнительный анализ. Для анализа в качестве альтернатив были выбраны следующие методы, на основе которых возможна реализация рекомендательных веб-сервисов: признаковые описания (content-based), коллаборативная фильтрация (collaborative filtering), гибридный подход [4–7]. Суть метода признаковых описаний заключается в использовании инструментов сбора данных как о пользователях сервисов, так и о объектах рекомендации, например, с помощью анкет. Метод коллаборативной фильтрацией заключается в выработке рекомендаций только на основании взаимодействия пользователей с объектами. Гибридный подход основан на одновременном использовании обоих вышеперечисленных подходов при реализации веб-сервисов.

Методы на основе признаковых описаний и гибридный подход требуют четкой классификации предоставляемого контента для выявления его параметров, которые в дальнейшем используются системой для выбора образовательных ресурсов. В рассматриваемой предметной области (образовательные ресурсы) сложно выделить критерии сравнения, так как рекомендуемые ресурсы являются разнородными, поэтому использование метода признаковых описаний алгоритмов невозможно. Таким образом, для решения поставленной задачи исследования будем использовать метод коллаборативной фильтрации [8–10].

Методы. Метод коллаборативной фильтрации основывается на оценках пользователей тому или иному образовательному ресурсу, эти оценки представляются в виде матрицы пользователь-ресурс, по которой в дальнейшем можно делать прогнозы относительно того, какие образовательные ресурсы лучше предложить для реализации образовательных траекторий студентов [3].

Осуществим формальное описание задачи:

$s \in S$ – множество студентов, использующих веб-сервис.

$e \in E$ – множество образовательных ресурсов.

$(q_{ue}, s, e) \in Q$ – оценки, проставляемые пользователями ресурсов.

Требуется предсказать, какую оценку выставит i -й пользователь веб-сервиса k -му образовательному ресурсу.

Основой рассматриваемого метода коллаборативной фильтрации является предположение, что похожим пользователям веб-сервиса будут полезны похожие образовательные ресурсы. Необходимо выбрать метрику, которая бы позволила определять схожесть двух пользователей. В работе рассматривается косинусная мера. Входной информацией для данного метода является матрица студент – образовательный ресурс, в которой отражены пользовательские оценки. Все оценки одного пользователя представляют собой вектор. Косинусная мера позволяет оценивать схожесть векторов по значению косинуса угла между ними.

$$\cos(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}| * |\vec{y}|}$$

Алгоритм коллаборативной фильтрации состоит из следующих этапов:

1. Выбрать пользователей, чьи предпочтения более всего совпадают с предпочтениями анализируемого пользователя. Критерием схожести выступает выбранная мера – косинусная.
2. Произвести калибровку множества оценок похожих пользователей: каждую оценку умножить на коэффициент схожести с рассматриваемым пользователем.
3. Чтобы получить итоговую предполагаемую оценку конкретного образовательного ресурса, необходимо просуммировать калиброванные оценки выбранных похожих пользователей и нормировать их, разделив на сумму мер выбранных пользователей.

Математическая модель. Формальное представление алгоритма реализации предложенного метода:

$$q_{s,e} = k * \sum_{s' \in S} \text{sim}(s, s') * r_{s',e},$$

где $k = \sum_{s' \in S} |\text{sim}(s, s')|$ – нормировочный коэффициент;

функция sim – косинусная мера схожести пользователей;

$q_{s',e}$ – оценка образовательному ресурсу от пользователя u' ;

$q_{s,e}$ – оценка образовательному ресурсу от пользователя u ;

S – множество пользователей, s – рассматриваемый пользователь;

s' – пользователь из множества похожих.

Основным недостатком предложенного метода является сложность реализации при отсутствии оценок пользователей тому или иному образовательному ресурсу или для нового пользователя, который не вносил еще никаких оценок, сложно сделать прогноз, это называется «холодным стартом» и присуще всем системам предоставления персонализированных рекомендаций. В работе предлагается следующий алгоритм: при недостатке данных о студенте или о существующих образовательных ресурсах, заполнять список альтернатив образовательных ресурсов, которые попадают в одну категорию с запрашиваемым ресурсом.

Веб-сервис. Рассмотрим особенности реализации веб-сервиса выбора образовательных ресурсов для реализации индивидуальных образовательных траекторий студентов [11–16].

Рекомендательный веб-сервис должна предоставлять студенту следующие функциональные возможности: регистрация в системе, поиск образовательных ресурсов, просмотр результатов поиска по запросу с включенными альтернативными вариантами, выставление оценки ресурсу [17–19].

Модульная структура веб-сервиса, она представлена на рис. 1 .

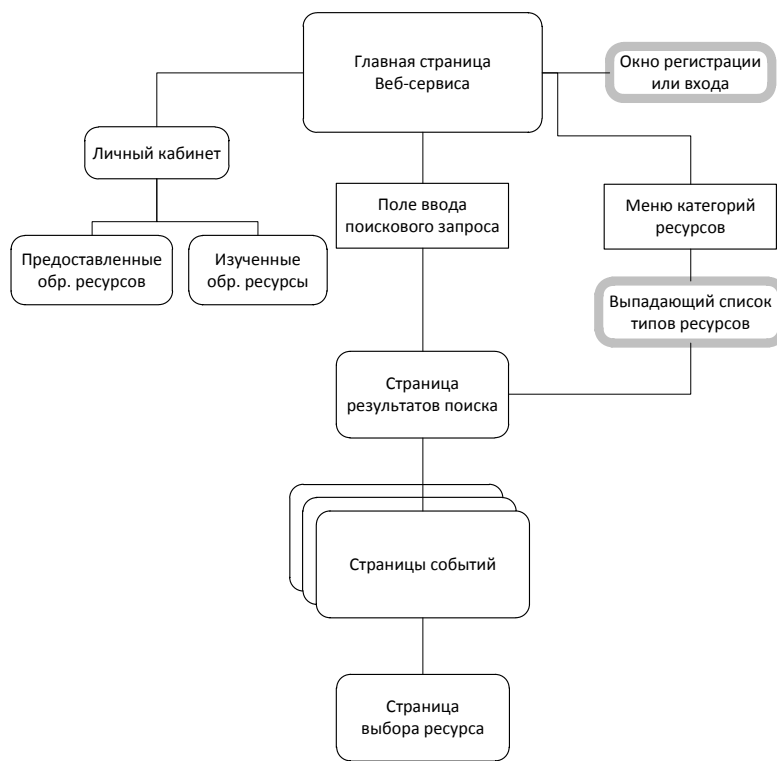


Рис. 1. Структура модуля потребителя образовательных ресурсов

Модуль потребителя образовательных ресурсов состоит из главной страницы, которая содержит блоки «Окно регистрации или входа», который предназначен для входа пользователя в систему для получения образовательных ресурсов; поле ввода поискового запроса, в котором пользователь может ввести желаемый ресурс для поиска; меню категорий образовательных ресурсов с подкатегориями в выпадающем списке для поиска без конкретного запроса. С главной страницы пользователь может перейти на страницу результатов поиска либо по вводу поискового запроса, либо по выбору типа образовательного ресурса. В личном кабинете обучающегося находятся все образовательные ресурсы: предоставленные и изученные. Пользователь может просматривать список ресурсов, а уже изученным выставлять оценки. Страница результатов поиска содержит список ресурсов, из которого обучающийся может выбрать интересующее его и перейти на страницу этого ресурса.

Диаграмма состояний описывает состояния, в которых может находиться веб-сервис, а также события, которые вызывают изменения этих состояний. Диаграмма приведена на рис. 2.

Для работы с системой обучающийся должен ввести запрос, при этом возникает вероятность опечатки. Если не устранять опечатки, то система не сможет найти ни одного подходящего результата. Для коррекции ввода будет применяться расстояние Левенштейна или редакционное расстояние. Это метрика, обозначающая разницу между двумя строками – минимальное количество операций вставки одного символа, удаления одного символа и замены одного символа на другой, необходимых для получения из входной строки выходной. Чтобы применять рас-

стояние Левенштейна для редактирования входного запроса, необходимо использовать конечный автомат – логическое устройство, принимающее некоторое слово только в том случае, если расстояние Левенштейна между входным словом и данным не превышает заданного значения [20].

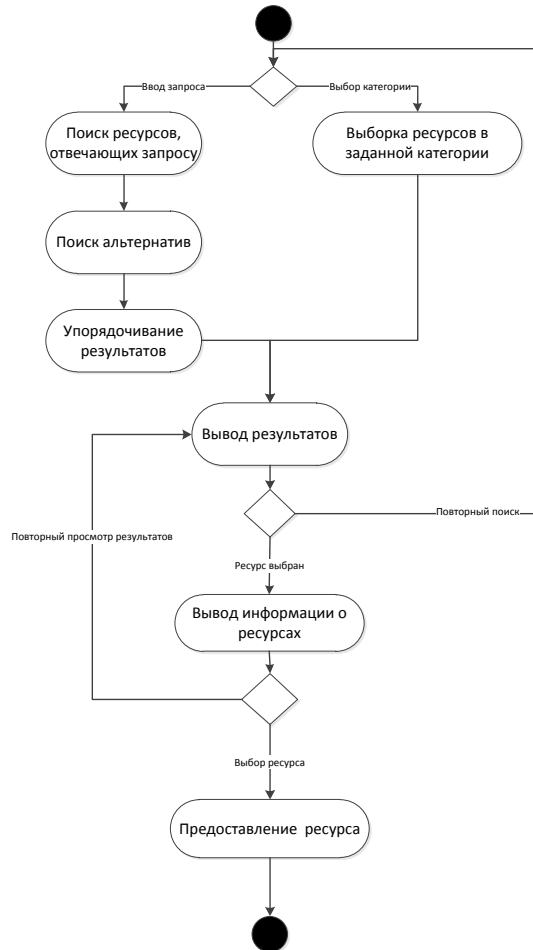


Рис. 2. Диаграмма состояний

Формально задачу поиска слова в словаре можно представить следующим образом:

- V – входная строка;
- R – выходная строка;
- S – словарь.

Выбрать из заданного словаря подмножество W всех слов, мера отличия которых k от поискового запроса не превышает некоторого порогового значения n.

$$W = \{W_i \mid W_i \in S \ \& \ k(W_i, V) \leq n\}$$

Мера отличия двух слов вычисляется через расстояние Левенштейна, т.е. как минимальное число вставок, замен и удалений символов. Цена вставки, замены и удаления символа в данной задаче принимается равной 1. Таким образом, конечный автомат Левенштейна позволяет найти слово, подходящее по словарю и являющееся самым близким к введенному.

Из начального состояния при вводе запроса пользователя система переходит в состояние поиска образовательных ресурсов, отвечающих запросу, а затем поиска альтернативных образовательных ресурсов.

Если обучающийся выбрал категорию образовательного ресурса, то система делает выборку образовательных ресурсов заданной категории. После того, как поиск завершен, система выводит результаты и ждет отклика пользователя.

Выводы. Использование предложенного подхода при реализации веб-сервиса выбора образовательных ресурсов для реализации индивидуальных образовательных траекторий обучающихся на основе коллаборативной фильтрации позволит предложить более эффективные технологии образовательной деятельности, что повысит качество образования и даст возможность студенту удовлетворить свои потребности в самообразовании и самосовершенствовании.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Bouguettaya A Sheng., Q.Z., and Daniel F. Eds., Web Services Foundations. – Springer, 2014.
2. Sheng Q.Z., Qiao X., Vasilakos A.V., Szabo C., Bourne S., and Xu X. Web Services Composition: A Decade's Overview // Information Sciences. – 2014. – Vol. 280. – P. 218-238.
3. Abdelwahab A. Feature optimization approach for improving the collaborative filtering performance using particle swarm optimization // Journal of Computational Information Systems 8.1. – 2012. – P. 435-450.
4. Liu H. A new user similarity model to improve the accuracy of collaborative filtering, Knowledge-Based Systems. – 2013.
5. Lina Yao, Quan Z. Sheng, Anne. H.H. Ngu, Jian Yu, Aviv Segev. Unified Collaborative and Content-Based Web Service Recommendation // IEEE Transactions on Services Computing. – May-June 2015. – Vol. 8, No. 3. – P. 453-466.
6. Quan Z. Sheng, Zakaria Maamar, Lina Yao, Claudia Szabo, and Scott Bourne. Behavior Modeling and Automated Verification of Web Services // Information Sciences. – 2014. – Vol. 258. – P. 416-433.
7. Zakaria Maamar, Nora Faci, K. Boukadi, Quan Z. Sheng, Lina Yao. Commitments to Regulate Social Web Services Operation // IEEE Transactions on Services Computing (TSC). – 2014. – Vol. 7, No. 2. – P. 154-167.
8. Szu-Yin Lin. A trustworthy QoS-based collaborative filtering approach for web service discovery // Journal of Systems and Software. – July 2014. – Vol. 93. – P. 217-228.
9. Lina Yao, Sheng, Q.Z., Segev, A., Jian Yu. Recommending Web Services via Combining Collaborative Filtering with Content-based Features // 2013 IEEE International Conference on Web Services (ICWS), Santa Clara. – 2013. – P. 42-49.
10. Atisha Sachan. A Survey on Recommender Systems based on Collaborative Filtering Technique // International journal of Innovations in Engineering and technology (IJET). – 2013.
11. Липко Ю.Ю. Исследование подходов трансформации требований при формализации предметной области // Информатизация и связь. – 2014. – № 2. – С. 75-76.
12. Rogozov Ю.И., Свиридов А.С., Липко Ю.Ю. Использование гиперметода разработки методов создания прототипов на примере построения структурно-независимых баз данных // Известия ЮФУ. Технические науки. – 2011. – № 1 (114). – С. 130-135.
13. Липко Ю.Ю. Алгоритм формализации требований при разработке информационных систем // Известия ЮФУ. Технические науки. – 2014. – № 6 (155). – С. 153-158.
14. Липко Ю.Ю., Свиридов А.С., Беликов А.Н. Approach to automatized formation of configurable information systems structure // Сборник материалов VIII Международной конференции «8th International Conference on Application of Information and Communication Technologies» (AICT-2014). – 2014. – С. 115-120.
15. Липко Ю.Ю. Model transformation process requirements in a formal model in the development of information systems // Innovative technologies and Didactics in Teaching: collected papers. – Berlin: MVB Marketing- und Verlagsservice des Buchhandels GmbH, 2014. – P. 54-59.
16. Rogozov Ю.И., Свиридов А.С. Концепция построения методологических информационных систем // Информатизация и связь. – 2014. – № 2. – С. 11-14.

17. Tao Yue, Lionel C. Briand, Labiche Y. A systematic review of transformation approaches between user requirements and analysis models // Requirements Engineering. – 2011. – P. 75-99.
18. Fatna Belqasmi and Roch Glitho, Concordia University Chunyan Fu, Tekelec «RESTful Web Services for Service Provisioning in NextGeneration Networks: A Survey». IEEE IEEE Communications Magazine. December. – 2011.
19. Official Congressional Biography, Allard Kenneth Lowenstein, published by Office of the Clerk, U.S. House of Representatives, accessed March 26, 2011.
20. Гасфилд Д. Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология: пер. с англ. И.В. Романовского. – СПб.: Невский Диалект, 2003. – 654 с.

REFERENCES

1. Bouguettaya A Sheng., Q.Z., and Daniel F. Eds., Web Services Foundations. Springer, 2014.
2. Sheng Q.Z., Qiao X., Vasilakos A.V., Szabo C., Bourne S., and Xu X. Web Services Composition: A Decade's Overview, *Information Sciences*, 2014, Vol. 280, pp. 218-238.
3. Abdelwahab A. Feature optimization approach for improving the collaborative filtering performance using particle swarm optimization, *Journal of Computational Information Systems* 8.1, 2012, pp. 435-450.
4. Liu H. A new user similarity model to improve the accuracy of collaborative filtering, *Knowledge-Based Systems*, 2013.
5. Lina Yao, Quan Z. Sheng, Anne. H.H. Ngu, Jian Yu, Aviv Segev. Unified Collaborative and Content-Based Web Service Recommendation, *IEEE Transactions on Services Computing*, May-June 2015, Vol. 8, No. 3, pp. 453-466.
6. Quan Z. Sheng, Zakaria Maamar, Lina Yao, Claudia Szabo, and Scott Bourne. Behavior Modeling and Automated Verification of Web Services, *Information Sciences*, 2014, Vol. 258, pp. 416-433.
7. Zakaria Maamar, Nora Faci, K. Boukadi, Quan Z. Sheng, Lina Yao. Commitments to Regulate Social Web Services Operation, *IEEE Transactions on Services Computing (TSC)*, 2014, Vol. 7, No. 2, pp. 154-167.
8. Szu-Yin Lin. A trustworthy QoS-based collaborative filtering approach for web service discovery, *Journal of Systems and Software*, July 2014, Vol. 93, pp. 217-228.
9. Lina Yao, Sheng, Q.Z., Segev, A., Jian Yu. Recommending Web Services via Combining Collaborative Filtering with Content-based Features, *2013 IEEE International Conference on Web Services (ICWS), Santa Clara*, 2013, pp. 42-49.
10. Atisha Sachan. A Survey on Recommender Systems based on Collaborative Filtering Techniques, *International journal of Innovations in Engineering and technology (IJJET)*, 2013.
11. Lipko Yu.Yu. Issledovanie podkhodov transformatsii trebovaniy pri formalizatsii predmetnoy oblasti [The study approaches the transformation of the requirements for the formalization of the subject area], *Informatizatsiya i svyaz' [Informatization and communication]*, 2014, No. 2, pp. 75-76.
12. Rogozov Yu.I., Sviridov A.S., Lipko Yu.Yu. Ispol'zovanie gipermetoda razrabotki metodov sozdaniya prototipov na primere postroeniya strukturno-nezavisimyykh baz dannykh [Development of method of use hypermethod prototyping for constructing structural-independent database], *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2011, No. 1 (114), pp. 130-135.
13. Lipko Yu.Yu. Algoritm formalizatsii trebovaniy pri razrabotke informatsionnykh sistem [The algorithm of formalization of requirements when developing information systems], *Izvestiya YuFU. Tekhnicheskie nauki [Izvestiya SFedU. Engineering Sciences]*, 2014, No. 6 (155), pp. 153-158.
14. Lipko Yu.Yu., Sviridov A.S., Belikov A.N. Approach to automatized formation of configurable information systems structure, *Sbornik materialov VIII Mezhdunarodnoy konferentsii [The collection of materials of VIII International conference] «8th International Conference on Application of Information and Communication Technologies» (AICT-2014)*, 2014, pp. 115-120.
15. Lipko Yu.Yu. Model transformation process requirements in a formal model in the development of information systems, *Innovative technologies and Didactics in Teaching: collected papers*. Berlin: MVB Marketing- und Verlagservice des Buchhandels GmbH, 2014, pp. 54-59.

16. Rogozov Yu.I., Sviridov A.S. Kontseptsiya postroeniya metodologicheskikh informatsionnykh sistem [The concept of methodological information systems], *Informatizatsiya i svyaz'* [Informatization and Communication], 2014, No. 2, pp. 11-14.
17. Tao Yue, Lionel C. Briand, Labiche Y. A systematic review of transformation approaches between user requirements and analysis models, *Requirements Engineering*, 2011, pp. 75-99.
18. Fatna Belqasmi and Roch Glitho, Concordia University Chunyan Fu, Tekelec «RESTful Web Services for Service Provisioning in NextGeneration Networks: A Survey». IEEE IEEE Communications Magazine. December, 2011.
19. Official Congressional Biography, Allard Kenneth Lowenstein, published by Office of the Clerk, U.S. House of Representatives, accessed March 26, 2011.
20. Gasfild D. Stroki, derev'ya i posledovatel'nosti v algoritmakh: Informatika i vychislitel'naya biologiya [Strings, trees and sequences in algorithms: computer science and computational biology]: translation from English I.V. Romanovskogo. St. Petersburg: Nevskiy Dialekt, 2003, 654 p.

Статью рекомендовал к опубликованию д.т.н., профессор В.А. Обуховец.

Липко Юлия Юрьевна – Южный федеральный университет; e-mail: jullipko@gmail.com; 347900, г. Таганрог, ул. Чехова, 49, кв. 21; кафедра системного анализа и телекоммуникаций; доцент.

Lipko Yulia Ur'evna – Southern Federal University; e-mail: jullipko@gmail.com; 49, Chekhov street, ap. 21, Taganrog, 347900, Russia; the department of SAiT; associate professor.

УДК 004.891.2

DOI 10.18522/2311-3103-2016-8-122137

А.Н. Целых, В.С. Васильев, Л.А. Целых

ИЗВЛЕЧЕНИЕ АТТРИБУТОВ ИЗ ГРАФОВ С ИСПОЛЬЗОВАНИЕМ ФУНКЦИОНАЛОВ МЕХАНИСТИЧЕСКОЙ АНАЛОГИИ ДЛЯ ЭКСПЕРТНЫХ СИСТЕМ*

В данной работе мы представляем новый подход для развития базы знаний продукционной экспертной системы (ЭС) на основе свойств причинно-следственных отношений между атрибутами логического вывода. Предлагается использовать нечеткие когнитивные карты, представляющие собой нечеткие ориентированные взвешенные знаковые графы с циклами обратной связи, как базу для генерации набора антецедентов и консеквентов нечетких правил ЭС, основанной на механизме нечеткого логического вывода. Полученная на этой основе база знаний использует системные качества для повышения интеллектуальности логического вывода. Современные графы имеют высокую степень размерности и многочисленные циклы, что значительно затрудняет эвристическое определение атрибутов логического вывода. Проблема коррекции коэффициентов передачи обратных связей разрешается путем приведения матрицы смежности (насколько это возможно) к верхнему (или нижнему) треугольному виду с использованием квадратичных функционалов механистической аналогии (барьерного штрафа и «инверсных»). Первый тип функционалов является знакоопределенным, выпуклым, поэтому процесс минимизации имеет единственное решение, достижимое любыми методами 0-го и 1-го порядка. Функционалы второго типа являются неограниченными, но, поскольку производится постоянное отображение действительно-значного решения на целочисленные перестановки, то задача также имеет решение, достижимое теми же методами. Представлен алгоритм минимизации функционала методом попарных сравнений. Показано, что алгоритм является вычислительно эффективным, соответствующим $O(n^3)$ затрат. Проведена серия вычислительных экспериментов для анализа и сравнения качества получаемых решений, а также обоснования выбора лучшего функционала. Эксперимент показал эффективность для матриц размером

* Данная работа произведена при поддержке гранта Российского Фонда фундаментальных исследований № 16-01-00098.