

1. Алгоритм PageRank вычисляет веса для всех веб-страниц (которые были проиндексированы) ещё до процедуры выполнения запроса пользователя. Алгоритм HITS применяется только к веб-страницам, полученным в результате выполнения определенного запроса пользователя.
2. Алгоритм HITS находит как «авторитеты», так и «концентраторы», PageRank – только «авторитеты».
3. Алгоритм PageRank требует нетривиальных вычислений, HITS – простой алгоритм, но очень затратный по времени вычисления.

Наиболее эффективными признаками для увеличения качества поиска являются признаки, основывающиеся на анализе ссылочной структуры веб-ресурсов, но в коллекциях, не обладающих данной структурой, можно получить улучшение качества поиска с использованием других признаков, подсчитываемых для целого документа или некоторых его атрибутов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Manning C.D., Raghavan P., Schütze H. Introduction to information retrieval // Cambridge University Press. – 2008. – 544 p.
2. Brin S., Page L. The Anatomy of a Large-Scale Hypertextual Web Search Engine, Proc. Seventh World Wide Web Conf., Elsevier Science. – New York, 1998.
3. Page L. et al. The PageRank Citation Ranking: Bringing Order to the Web, Stanford Digital Library Technologies, Working Paper 1999-0120, Stanford Univ., Palo Alto, Calif., 1998.
4. Kleinberg J.M. Authoritative Sources in a Hyperlinked Environment. Journal of the ACM 46, 5, 1999. – P. 604-632.
5. Bhart K., Henzinger M. Improved Algorithms for Topic Distillation in a Hyperlinked Environment. In Proceedings of ACM SIGIR'98 (Melbourne, Australia), 1998.

Статью рекомендовал к публикации д.т.н., профессор В.П. Карелин.

Котов Эдуард Михайлович – Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге; e-mail: kotov@tti.sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 88634371743; кафедра прикладной информатики; старший преподаватель.

Kotov Eduard Michaylovich – Taganrog Institute of Technology – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”; e-mail: kotov@tti.sfedu.ru; 44, Nekrasovsky, Taganrog, 347928, Russia; +78634371743; the department of applied information science; senior instructor.

УДК 519.14

А.В. Боженюк, Н.С. Опенько

ИССЛЕДОВАНИЕ И АНАЛИЗ МЕТОДОВ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ НЕЧЕТКОЙ ИНФОРМАЦИИ*

*Описываются методы принятия решений на основе построения и обоснования механизмов нечеткого логического вывода. Данная задача является актуальной, потому что имеет широкое практическое применение. Рассматриваются основные схемы нечеткого вывода на основе методов Мамдани и Сугено, описываются их недостатки. Рассмотрен метод выбора решений на основе истинности правила *modus ponens*. Построена модель принятия решений на основе степени истинности правила *modus ponens*, использованная*

* Работа поддержана РФФИ, проект № 11-01-00011а.

для задачи постановки конкретного медицинского диагноза в случае угрозы перелома костей человека. Представлена постановка задачи. Эффективность предложенного метода заключается в том, что он позволяет на основе текущего состояния системы обеспечить определение всех возможных исходов у конкретного больного и постановку максимально достоверного диагноза.

Нечеткий логический вывод; лингвистическая переменная; степень истинности; правило modus ponens.

A.V. Bozhenyuk, N.S. Openko

RESEARCH AND ANALYSIS OF METHODS OF DECISION-MAKING ON THE BASIS OF FUZZY INFORMATION

This article describes the methods of decision making on the basis of constructing and justifying the fuzzy inference mechanism. This problem is relevant because it has wide practical application. In the paper the basic model of fuzzy inference methods Mamdani and Sugeno. Also described a method of selection decisions based on the truth of the rules of modus ponens. In this paper, we construct a model of decision-making on the basis of degree of truth rules modus ponens, used for the problem of setting a specific medical diagnosis in the event of fracture of the person. To illustrate the problem statement is presented. The effectiveness of this method is that it allows on the basis of the current state of the system ensure that all possible outcomes of an individual patient and posing as a definite diagnosis.

Fuzzy inference; linguistic variable; the degree of truth; modus ponens rule.

Введение. При использовании информации, получаемой от эксперта и хранящейся в БД системы принятия решения (ПР), возникают трудности. Это, в первую очередь, связано с тем, что экспертная информация может содержать неточности и неопределенности, которые присущи в описании правил вывода посредством высказываний естественного языка. Другая трудность заключается в использовании этих правил, когда наблюдаемые факты не могут быть подходящим образом идентифицированы с посылкой правила вывода. Кроме того, согласно принципу несовместимости, сложность системы управления и точности, с которой ее можно проанализировать традиционными математическими методами, находятся в состоянии взаимного противоречия [1].

Теория нечетких множеств и основанная на ней логика [2] позволяют, используя естественный язык со всем набором имеющихся в нем средств для выражения человеческих способов рассуждений и принятия решений с помощью качественных представлений, понятий и оценок, описывать неточные категории, представления и знания, оперировать ими и делать соответствующие заключения и выводы. Использование нечеткого описания, наряду с четким представлением информации, позволяет всесторонне и компактно описывать общую смысловую постановку задач управления и принятия решений, возникающих в различных прикладных областях. Решение задач, в которых исходная информация включает в себя нечеткие и четкие характеристики, является в настоящее время интересной и актуальной проблемой.

Анализ нечетких моделей. В связи с этим в настоящее время большое внимание уделяется использованию нечетких методов и алгоритмов [3]. В этом случае модель ПР строится в виде логико-лингвистического описания взаимосвязей входных управляющих и выходных управляемых параметров. Формирование лингвистических моделей осуществляется на естественном или близком к нему языке в виде совокупности правил типа <если ... то ...>, которые образуют основу базы знаний процесса ПР.

Здесь основную роль играет дедуктивное правило modus ponens, имеющее вид

$$\begin{aligned}
 P_1 &: < \text{если } A \text{ есть } a \text{ то } B \text{ есть } b >; \\
 P_2 &: < A \text{ есть } a > - \text{истинно}; \\
 \hline
 &< B \text{ есть } b > - \text{истинно}.
 \end{aligned}
 \tag{1}$$

Согласно ему, если существует информация в виде высказывания P_1 и наблюдается факт в виде высказывания P_2 , то делается вывод $<B \text{ есть } b>$. Если же высказывание P_1 не является посылкой высказывания P_2 (например, имеет вид $<A \text{ есть } a'>$), тогда правило modus ponens не может быть применено. Однако в работе [1] Л.Заде расширил правило modus ponens, в котором, если понятия a, b, a' высказываний P_1 и P_2 моделируются нечеткими множествами, то нечеткое заключение $<B \text{ есть } b'>$ может быть выведено.

Здесь главными задачами являются: задача построения и обоснования механизма нечеткого логического вывода, согласно которому по нечетким знаниям типа P_1 и входным параметрам P_2 делается вывод о нечетких значениях выходного параметра управления; задача “дефаззификации”, т.е. задача преобразования полученного нечеткого множества в конкретное значение выходного параметра управления.

Обозначим через A и B множества значений входных и выходных параметров процесса ПР. Пусть β_A и β_B – лингвистические переменные, определенные на множествах A и B с базовыми значениями $T_A = \{\alpha_{A_j}\}, j = \overline{1, m}$, и $T_B = \{\alpha_{B_i}\}, i = \overline{1, n}$ соответственно. Здесь α_{A_j} и α_{B_i} – нечеткие переменные. Нечеткую информацию, полученную от эксперта, представим системой нечетких условных высказываний:

$$\tilde{L} = \{ \tilde{L}_i : < \text{если } \tilde{A}_i \text{ то } \tilde{B}_i > \}, i = \overline{1, n}.
 \tag{2}$$

Здесь \tilde{A}_i и \tilde{B}_i – нечеткие высказывания соответственно, $< \beta_A \text{ есть } \alpha_{A_i} >$ и $< \beta_B \text{ есть } \alpha_{B_i} >$, $\alpha_{A_i} \in T_A$ и $\alpha_{B_i} \in T_B$.

В общем случае механизм логического вывода включает четыре этапа [3, 4]: введение нечеткости (фаззификация); нечеткий вывод; композиция; приведение к четкости (дефаззификация). Существуют различные модели нечеткого вывода, основными из которых являются модели Мамдани и Сугено [5–7].

Традиционно нечеткий логический вывод для системы вида (2) представляется в виде максиминной композиции в качестве композиционного правила вывода и трактовке условного высказывания в виде операции импликация по Мамдани. При таком подходе определяется функция принадлежности $\mu_B^*(b) = \bigvee_{i=1, n} (\mu_{A_i}(a^*) \& \mu_{B_i}(b))$.

Здесь μ_{A_i} и μ_{B_i} – функции принадлежности, соответствующие нечетким переменным $\alpha_{A_i} \in T_A$ и $\alpha_{B_i} \in T_B$.

При решении задачи “дефаззификации” на основе анализа $\mu_B^*(b)$ определяется конкретное четкое значение b_0 выходного параметра B . Существуют различные методы дефаззификации, в основе которых можно выделить два подхода. Первый подход определяет значение b_0 , исходя из анализа всей функции принадлежности μ_B^* . Второй подход использует экстремальные значения функции принадлежности μ_B^* . Типичным представителем первого подхода является метод нахождения центра тяжести фигуры b_0 , ограниченной функцией принадлежности $\mu_B^*(b)$.

Рассмотренный нечеткий логический вывод обладает рядом существенных недостатков [8, 9], а именно: диапазон изменения выходного параметра составляет часть от общего диапазона управления (в некоторых случаях всего лишь 33,3 %); зависимость $b_0(a)$ является нелинейной.

Типичным представителем второго подхода является метод середины максимума. В частности, если экстремум функции μ_B наблюдается во всех точках некоторого интервала $[b_1, b_2] \in B$, естественно определить значение $b_0 = \frac{b_1 + b_2}{2}$. Использование на этапе дефаззификации метода середины максимума является проблематичным, так как функция μ_B не обладает свойством квазивогнутости.

Принятие решений на основе степени истинности правила modus ponens. В работах [10–12] было введено понятие степени истинности нечеткого правила modus ponens для схемы вывода:

$$\begin{aligned} & \tilde{L}; \\ & \underline{A^* : - истинно}; \\ & \underline{B : - истинно}. \end{aligned} \quad (3)$$

В качестве решений предлагается выбирать такие значения, для которых введенная степень истинности $T_{m.p.}$, определяемая, согласно

$$T_{m.p.}(a^*, b) = \underset{i=1, n}{\&}(T(A^*/\tilde{A}_i) \rightarrow T(B/\tilde{B}_i)) \cdot \quad (4)$$

достигает наибольшего значения.

В выражении (4) величины $T(A^*/\tilde{A}_i)$ и $T(B/\tilde{B}_i)$ есть степени истинности четких высказываний $A^* : \langle \beta_A \text{ есть } a^* \rangle$ и $B : \langle \beta_B \text{ есть } b \rangle$ относительно нечетких \tilde{A}_i и \tilde{B}_i соответственно, которые определяются как $T(A^*/\tilde{A}_i) = \mu_{A_i}(a^*)$ и $T(B/\tilde{B}_i) = \mu_{B_i}(b)$, операция $\&$ – некоторая t -норма, а операция \rightarrow есть операция нечеткой импликации [13].

В работе [10] было показано, что если функции принадлежности базовых значений выходной лингвистической переменной β_B являются квазивогнутыми непрерывными функциями, операция нечеткой импликации обладает свойствами непрерывности и $0 \rightarrow \forall = 1$ (ложь влечет все), а в качестве t -нормы использовать операцию \min , то истинность нечеткого правила modus ponens для схемы вывода (3) также является квазивогнутой непрерывной функцией. То есть принимает свое максимально возможное значение либо в одной точке, либо на некотором интервале. Поэтому на этапе дефаззификации естественно использовать метод середины максимума.

Модель принятия решений на основе степени истинности правила modus ponens была использована для задачи постановки конкретного медицинского диагноза в случае угрозы перелома костей человека.

В качестве первой входной лингвистической переменной используется плотность кости человека (\tilde{A}), где \tilde{A}_1 – высокая плотность кости, \tilde{A}_2 – средняя плотность, \tilde{A}_3 – низкая плотность. В качестве второй входной лингвистической переменной принимается параметр степени нестабильности суставов (\tilde{B}), где \tilde{B}_1 –

легкая степень нестабильности, \tilde{B}_2 – средняя степень, \tilde{B}_3 – тяжелая степень. Следовательно, множества значений входных переменных имеют следующий вид: $\tilde{A} = \{\tilde{A}_1, \tilde{A}_2, \tilde{A}_3\}$ и $\tilde{B} = \{\tilde{B}_1, \tilde{B}_2, \tilde{B}_3\}$.

В качестве выходной лингвистической переменной рассматривается параметр вида перелома (\tilde{C}), где \tilde{C}_1 – линейный перелом, \tilde{C}_2 – оскольчатый перелом, \tilde{C}_3 – фрагментарно-оскольчатый перелом, \tilde{C}_4 – нет перелома. Следовательно, множество значений выходных переменных равно:

$$\tilde{C} = \{\tilde{C}_1, \tilde{C}_2, \tilde{C}_3, \tilde{C}_4\}.$$

В рассматриваемой системе не может быть однозначного соотношения входных и выходных переменных. Следовательно, в этом случае система нечетких высказываний будет содержать нечеткие правила следующего вида, например:

ПРАВИЛО_1: ЕСЛИ « \tilde{A}_1 & \tilde{B}_1 » ТО « $1 \tilde{C}_4 \vee 0,9 \tilde{C}_2$ »;

ПРАВИЛО_2: ЕСЛИ « \tilde{A}_1 & \tilde{B}_2 » ТО « $0,8 \tilde{C}_4 \vee 1 \tilde{C}_1 \vee 0,7 \tilde{C}_2$ »;

ПРАВИЛО_3: ЕСЛИ « \tilde{A}_1 & \tilde{B}_3 » ТО « $0,8 \tilde{C}_1 \vee 1 \tilde{C}_2 \vee 0,7 \tilde{C}_3$ ».

Выводы. Рассмотрены основные схемы нечеткого вывода на основе методов Мамдани и Сугено, описаны их недостатки. Рассмотрен метод выбора решений на основе истинности правила *modus ponens*. Построена модель принятия решений на основе степени истинности правила *modus ponens*, использованная для задачи постановки конкретного медицинского диагноза в случае угрозы перелома костей человека. На основе анализа текущего состояния определяется значение выходной переменной, нахождение которой позволяет обеспечить оптимальное решение, а именно, определить все возможные исходы у конкретного больного и поставить максимально достоверный диагноз.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Zadeh L.A.* Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976.
2. *Zadeh L.A.* Fuzzy sets // *Inform. Control.* – 1965. – Vol. 8. – P. 338-353.
3. Прикладные нечеткие системы / Под ред. *Т.Тэрано, К.Асаи, М.Сугэно.* – М.: Мир, 1993.
4. *Борисов А.Н., Алексеев А.В., Крумберг О.А.* и др. Модели принятия решений на основе лингвистической переменной. – Рига: Зинатне, 1982. – 256 с.
5. *Mamdani E.H. and Assilian S.* An experiment in linguistic synthesis with a fuzzy logic controller // *Int. J. Man-Mach. Stud.* – 1975. – Vol. 7. – P. 1-13.
6. *Takagi T. and Sugeno M.* Fuzzy identification of systems and its applications to modeling and control // *IEEE Tran. Syst., Man, Cybern.* – 1985. – Vol. SMC 15. – P. 116-132.
7. *Дьяконов В., Круглов В.* Математические пакеты расширения MATLAB. Специальный справочник. – СПб.: Питер, 2001. – 480 с.
8. *Лохин В.М., Макаров И.М., Манько С.В., Романов М.П.* Методические основы аналитического конструирования регуляторов нечеткого управления // *Известия РАН. ТиСУ.* – 2000. – № 1. – С. 56-69.
9. *Guney K., Sarikaya N.* Comparison of Mamdani and Sugeno Fuzzy Inference System Models for Resonant Frequency Calculation of Rectangular Microstrip Antennas // *Progress In Electromagnetics Research B.* – 2009. – Vol. 12. – P. 81-104.
10. *Берштейн Л.С., Боженик А.В.* Нечеткие модели принятия решений: дедукция, индукция, аналогия. – Таганрог: ТРТУ, 2001. – 110 с.
11. *Берштейн Л.С., Боженик А.В.* Моделирование процесса определения предпочтительных параметров на основе нечеткого логического вывода // *Электронное моделирование.* – 1989. – № 3. – С. 98-101.

12. *Bershtein L.S., Bozhenyuk A.V., Rozenberg I.N.* Decision Making on the Basis of Monotonic Expert Information // 6th European Congress on Intelligent Techniques & Soft Computing. Aachen, Germany, September 7-10, 1998. – Vol. 2. – P. 1136-1140.
13. *Iancu I.* Propagation of uncertainty and imprecision in knowledge-based systems // Fuzzy Sets and Systems. – 1998. – № 94. – P. 29-43.

Статью рекомендовал к опубликованию д.т.н., профессор В.П. Карелин.

Боженюк Александр Витальевич – Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге; e-mail: avb002@yandex.ru; 347928, г. Таганрог, пер. Некрасовский, 44, тел.: 88634371743; кафедра прикладной информатики; д.т.н.; профессор.

Опенько Наталья Сергеевна – e-mail: openko.natalya@mail.ru; кафедра прикладной информатики; аспирантка.

Bozhenyuk Alexander Vitalievich – Taganrog Institute of Technology – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”; e-mail: avb002@yandex.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +78634371743; the department of applied information science; professor.

Openko Natalia Sergeevna – e-mail: openko.natalya@mail.ru; the department of applied information science; postgraduate student

УДК 519.688:[519.17+681.518]

С.Л. Беляков, М.Н. Савельева

КЛАСТЕРИЗАЦИЯ ПРИ ЛОГИСТИЧЕСКОМ РОУТИНГЕ*

Анализируется модель роутинга, использующая для накопления и обработки опыта картографическую базу данных ГИС. Особенность анализируемой модели состоит в использовании электронных географических карт, описывающих среду транспортировки через практически опробованные маршруты экспертом-логистом. На основании проверенного пути производится разбиение карты на кластеры. Новизна заключается в том, что предложенный алгоритм поиска наилучшего маршрута реализован на нечетком темпоральном гиперграфе. Ребрами данного гиперграфа являются полученные кластеры. Это позволяет строить более достоверные решения в условиях неопределенности.

ГИС; логистический опыт; кластеризация; маршрутизация; нечеткий темпоральный гиперграф.

S.L. Beliacov, M.N. Savelyeva

CLUSTERIZATION IN LOGISTIC ROUTING

The article covers analysis of the routing model that uses for accumulation and processing experience cartographic database of GIS. Feature of the analyzed model is the use of electronic maps describing the impact of transportation routes through practically tested expert logistician. The partition map into clusters is based on proven route. The novelty lies in the fact that the proposed algorithm for finding the best route is realized on a fuzzy temporal hypergraph. Edges of the hypergraph are obtained clusters. This allows to build more reliable solutions in the face of uncertainty.

GIS; logistic experience; clusterization; routing; fuzzy temporal hypergraph.

* Работа выполнена при финансовой поддержке грантов РФФИ 11-01-00011-а, 10-01-00029-а.