

Сеченов Михаил Дмитриевич

Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: michael_setch@mail.ru.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 88634371625.

Щеглов Сергей Николаевич

E-mail: leo@tsure.ru.

Setchenov Mixail Dmitrievich

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: michael_setch@mail.ru.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: 88634371625.

Shcheglov Sergej Nikolaevich

E-mail: leo@tsure.ru.

УДК 004.827

С.В. Скороход

**К ВОПРОСУ О ПРИМЕНЕНИИ ЛИНГВИСТИЧЕСКИХ ПЕРЕМЕННЫХ
ДЛЯ ПОСТРОЕНИЯ ЭКСПЕРТНЫХ ОЦЕНОК***

Рассматривается задача построения значения нечёткой переменной методом экспертного опроса. Анализируется подход, основанный на использовании лингвистической переменной. Продемонстрированы недостатки существующих методов, использующих нечёткие множества второго уровня. Предлагается подход, позволяющий осуществить переход от нечёткого множества второго уровня к значению нечёткой переменной, лишенный указанных недостатков.

Нечёткая переменная лингвистическая переменная экспертная оценка.

S.V. Skorokhod

**APPLICATION OF LINGUISTIC VARIABLES FOR EXPERT ESTIMATIONS
CONSTRUCTION**

The problem of a fuzzy variable value construction is considered by a method of expert inquiry. The approach based on use by a linguistic variable is analyzed. Lacks of the existing methods using fuzzy sets of the second level are shown. The approach, permitting to realize passage from an fuzzy set of the second level to a value of the indistinct variable, deprived the indicated shortages is offered.

Fuzzy variable linguistic variable expert estimation.

Постановка задачи. В системах принятия решений, основанных на обработке нечёткой информации, часто решается следующая задача. Имеется некоторая нечёткая переменная [1] \tilde{K} , заданная на числовой базовой шкале B . Требуется при помощи опроса нескольких экспертов построить её значение. Данная задача сводится к построению функции принадлежности $\mu_{\tilde{K}}(x)$ нечёткого множества (НМ):

* Работа выполнена при финансовой поддержке Международного научного фонда экономических исследований академика Н.П. Федоренко. Проект № 2009-103

$$K = \int_{x \in B} \mu_K(x) / x \quad (1)$$

с учётом мнения каждого из экспертов.

Существует множество прямых и косвенных методов её решения, обзор которых приведён в [1]. Общим их недостатком является большая трудоёмкость процедуры экспертного опроса, что делает их крайне неудобными для применения в автоматизированных системах. В данной работе предлагается подход, основанный на использовании лингвистической переменной [2], включающей в себя набор терм-множеств. Функции принадлежности терм-множеств строятся однократно любым из известных методов на подготовительном этапе, после чего они используются для построения $\mu_K(x)$, что существенно уменьшает трудоёмкость работы экспертов.

Нечёткая и лингвистическая переменные. Нечёткая переменная \tilde{K} определяется тройкой элементов $\tilde{K} = \langle N_{\tilde{K}}, B, K \rangle$, где $N_{\tilde{K}}$ – её название, B – базовая шкала, K – НМ (1), задающее её значение. Нечёткие переменные используются для описания каких-либо субъективных понятий, оцениваемых некоторым лицом, действующим в процессе принятия решения. Значением переменной является НМ (1), отражающее субъективные предпочтения этого лица. В общем случае, \tilde{K} имеет столько значений, сколько людей привлечено к оценке.

Для примера рассмотрим понятие «высокая сложность», оценка которого производится двумя экспертами. Нечёткая переменная будет иметь вид: $\tilde{K} = \langle \text{«высокая»}, [0;1], K \rangle$. Учитывая справедливость известного утверждения «сколько людей, столько и мнений», с большой степенью уверенности можно предположить, что мнения экспертов, совпадая, в целом, между собой, разойдутся в деталях. В результате получим два значения K_1 и K_2 , функции принадлежности которых изображены на рис. 1.

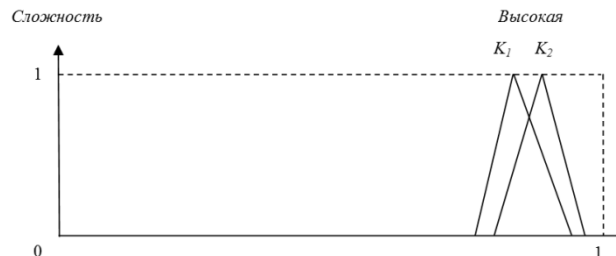


Рис. 1. Значения нечёткой переменной

Лингвистическая переменная L определяется тройкой $L = \langle N_L, B, T \rangle$, где N_L – её название, B – базовая шкала, $T = \{t_1..t_m\}$ – терм-множество, задающее набор её значений. При этом каждый элемент терм-множества $t_i, i = \overline{1, m}$ является нечёткой переменной на шкале B . Лингвистическая переменная является аппаратом для представления в системах принятия решений качественных категорий и их вербальных значений, а её терм-множество должно быть функционально полным и покрывать всё базовое множество B .

В качестве примера возьмём категорию «сложность задачи». Предположим, что в результате опроса сформировано функционально полное множество терминов, которыми эксперты характеризуют её уровень: «минимальная» (t_1), «низкая» (t_2), «лёгкая» (t_3), «средняя» (t_4), «повышенная» (t_5), «высокая» (t_6), «максималь-

ная» (t_7). Тогда соответствующая лингвистическая переменная $L = \langle \langle \text{Сложность} \rangle, [0; 1], T \rangle$, где $T = \{t_1, \dots, t_7\}$, а функции принадлежности термов изображены на рис. 2. Их анализ позволяет сделать вывод о том, что данное термножество удовлетворяет всем требованиям.

Необходимо отметить, что построение термножества и его функций принадлежности является важной задачей, существенно влияющей на качество процесса принятия решений. Для этого используется тщательный опрос экспертов [2], который выполняется однократно на этапе создания или настройки системы, использующей лингвистические переменные. В дальнейшем функции принадлежности термов могут редактироваться, но при подстройке системы, а не при выборе конкретного решения.

Построение экспертной оценки. Рассмотрим теперь этап принятия решения. Пусть имеется конкретная задача, сложность которой должна быть оценена одним или несколькими экспертами. Другими словами, требуется, с учётом их мнений, построить функцию принадлежности (1), имея предварительно созданную лингвистическую переменную с набором термов, изображённых на рис. 2. Для этого экспертам предлагается выбрать подходящий терм, соответствующий сложности рассматриваемой задачи.

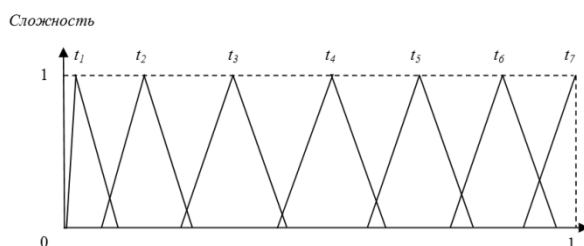


Рис. 2. Функции принадлежности термов

Если эксперт один, и он сделал однозначный выбор в пользу термина t^* , то в качестве (1) берётся функция принадлежности этого термина:

$$K = t^* = \int_{x \in [0;1]} \mu_{t^*}(x) / x. \quad (2)$$

В случае, если он сомневается, какому терму отдать предпочтение, ему следует оценить степени $\mu_D(t_i) \in [0;1], i = \overline{1, m}$ соответствия термов сложности задачи. Результатом является НМ второго уровня [2] на множестве термов:

$$D = \{ \langle \mu_D(t_i) / t_i \in T \rangle \}. \quad (3)$$

Заметим, что (2) является частным случаем (3), когда

$$(\mu_D(t^*) = 1) \ \& \ (\forall t \in T)(t \neq t^* \rightarrow \mu_D(t) = 0).$$

Если экспертов несколько (p человек), то в результате опроса получим p НМ второго уровня (3):

$$D_j = \{ \langle \mu_{D_j}(t_i) / t_i \in T \rangle \}, j = \overline{1, p}.$$

Вычислив среднее арифметическое функций принадлежности соответствующих термов по формуле:

$$\mu_D(t_i) = \frac{\sum_{j=1}^p \mu_{D_j}(t_i)}{p}, i = \overline{1, m},$$

получим единственное НМ второго уровня вида (3).

Итак, в результате опроса, как одного, так и нескольких экспертов, всегда получается НМ второго уровня (3). Рассмотрим вопрос о построении на его основе искомой функции принадлежности (1).

Традиционный метод, описанный в [2,3], основан на использовании α -уровней термов из T , где в качестве α -уровня для терма t_i используется величина $\mu_D(t_i)$. Тогда (1) вычисляется по формуле:

$$\mu_K(x) = \max_{t_i \in T} \min(\mu_D(t_i), \mu_{t_i}(x)). \tag{4}$$

Проиллюстрируем этот способ на примере. Пусть для набора термов (см. рис. 2) сформировано НМ $D = \{<0.6/t_4>, <0.4/t_5>\}$. Тогда (4) будет иметь вид, изображённый на рис. 3 штриховкой.

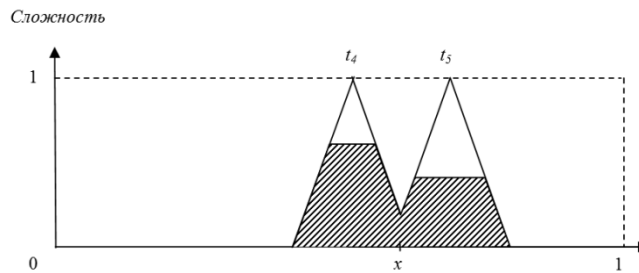


Рис. 3. Традиционный метод

Достоинством этого метода является его применимость для функций принадлежности термов любого вида. Но у него есть и существенный недостаток. В нашем примере наблюдается разброс мнений экспертов, которые почти равномерно распределились между t_4 и t_5 . Логично предположить, что искомая функция принадлежности должна иметь пик между пиками этих термов, немного ближе к t_4 , т.е. в точке x . Однако в этой точке она имеет локальный минимум, что противоречит здравому смыслу.

Рассмотрим другой метод, лишённый такого недостатка. Для этого наложим ограничения на форму функций принадлежности термов и предположим, что все они имеют треугольный вид, что соответствует изображениям (см. рис. 1-3). Поскольку эти функции определены на числовой базовой шкале (отрезок $[0; 1]$), мы имеем дело с треугольными нечёткими числами (НЧ) [4].

Треугольное НЧ $C = \langle c^l, c', c^r \rangle$ задаётся тройкой чисел, которые обозначают минимальное, наиболее правдоподобное (максимум функции принадлежности) и максимальное значения его области определения. Над треугольными НЧ определены операции сложения, вычитания, умножения и деления, результаты которых также являются треугольными НЧ, которые вычисляются по формулам:

$$C_1 + C_2 = \langle c_1^l + c_2^l, c_1' + c_2', c_1^r + c_2^r \rangle,$$

$$C_1 - C_2 = \langle c_1^l - c_2^l, c_1^i - c_2^i, c_1^r - c_2^r \rangle,$$

$$C_1 \cdot C_2 = \langle c_1^l \cdot c_2^l, c_1^i \cdot c_2^i, c_1^r \cdot c_2^r \rangle,$$

$$C_1 / C_2 = \langle c_1^l / c_2^l, c_1^i / c_2^i, c_1^r / c_2^r \rangle.$$

Полагая, что терм $t_i = \langle t_i^l, t_i^i, t_i^r \rangle, i = \overline{1, m}$, используем операцию умножения треугольных НЧ для вычисления (1) как средневзвешенного термов из T , в котором весовыми коэффициентами термов выступают значения $\mu_D(t_i)$ из (3). Тогда K в (1) имеет вид треугольного НЧ:

$$K = \langle \frac{\sum_{i=1}^m \mu_D(t_i) \cdot t_i^l}{\sum_{i=1}^m \mu_D(t_i)}, \frac{\sum_{i=1}^m \mu_D(t_i) \cdot t_i^i}{\sum_{i=1}^m \mu_D(t_i)}, \frac{\sum_{i=1}^m \mu_D(t_i) \cdot t_i^r}{\sum_{i=1}^m \mu_D(t_i)} \rangle.$$

Проиллюстрируем этот метод на предыдущем примере. Полученное НЧ K изображено на рис. 4 и, на мой взгляд, гораздо более точно отражает колебания экспертов между термами t_4 и t_5 нежели функция принадлежности (см. рис. 3).

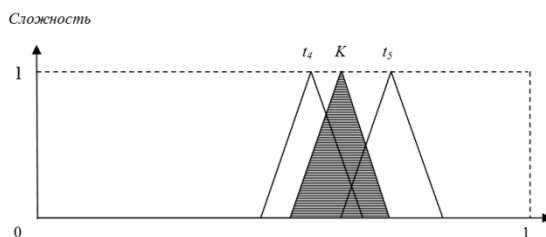


Рис. 4. Метод средневзвешенного термов

Выводы. В работе рассмотрена проблема построения нечеткой экспертной оценки, возникающая в задачах принятия решений. Для минимизации экспертного опроса предложено использовать лингвистическую переменную и её термножество. Проведён анализ двух способов построения экспертной оценки: традиционного, с использованием функций принадлежности произвольного вида, и основанного на использовании треугольных НЧ. Показано, что второй способ даёт более правдоподобные, с точки зрения здравого смысла, результаты.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Нечёткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А. Поспелова. – М.: Наука. Гл. ред. Физ.-мат. лит.– 1986.– 312 с.
2. Мелихов А.Н., Берштейн Л.С., Коровин С.Я. Ситуационные советующие системы с нечёткой логикой. – М.: Наука. Гл. ред. Физ.-мат. лит. – 1990. – 272 с.
3. Скороход С.В. Моделирование целей управления в условиях неопределённости // Материалы международной научной конференции «Гатищевские чтения: актуальные проблемы науки и практики». – Тольятти: Изд-во ВУИТ.– 2004. – С. 253-258.
4. Скороход С.В. Применение нечётких чисел для оценки квалификации персонала // Известия ТРТУ. – 2005. – № 3 (47). – С. 214-216.

Скороход Сергей Васильевич

Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: sss64@mail.ru.

347923, г. Таганрог, ул. Инструментальная, 19/2, кв. 48.

Тел.: 88634648891.

Skorokhod Sergey Vasilievitch

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: sss64@mail.ru.

19/2, ap. 48. Instrumentalnaya street, Taganrog, 347923, Russia.

Phone: +78634648891.

УДК 519.712.2

Л.А. Гладков, Н.В. Гладкова

**НОВЫЕ ПОДХОДЫ К ПОСТРОЕНИЮ СИСТЕМ АНАЛИЗА
И ИЗВЛЕЧЕНИЯ ЗНАНИЙ НА ОСНОВЕ ГИБРИДНЫХ МЕТОДОВ***

Приводятся основные отличия методов Data Mining от традиционных методов анализа, обсуждаются их преимущества и недостатки и приводятся предложения по их решению на основе использования гибридных интеллектуальных технологий и методов вычислительного интеллекта. Рассмотрены основные аспекты применения нечетких генетических алгоритмов для решения задач извлечения знаний. Описаны основные компоненты организации и процесса взаимодействия генетического алгоритма и нечеткого логического контроллера. Приводятся основные определения и принципы использования методов эволюционного проектирования и моделирования, многоагентных систем и нечетких математических моделей при создании гибридных компонентов интеллектуальных систем. Обсуждаются преимущества и недостатки традиционных методов и приводятся предложения по их решению на основе использования гибридных интеллектуальных технологий и методов вычислительного интеллекта. Приводится обоснование актуальности разработки новых гибридных методов анализа и извлечения данных.

Анализ и извлечение знаний; нечеткий генетический алгоритм; нечеткий логический контроллер; фаззификация; дефаззификация; мультиагентная система; эволюционное проектирование; модели эволюции.

L.A. Gladkov, N.V. Gladkova

**NEW APPROACHES TO CONSTRUCTION OF DATA MINING SYSTEMS
ON THE BASIS OF HYBRID METHODS**

In work the basic differences of methods Data Mining from traditional methods of the analysis are resulted. Also advantages and lacks of methods Data Mining are discussed and offers under their decision on the basis of use of hybrid intellectual technologies and methods of computing intelligence are resulted. Also the basic aspects of application of fuzzy genetic algorithms for the decision of problems of extraction of knowledge are considered. The basic components of the organization and process of interaction of genetic algorithm and the fuzzy logic controller are described. The basic definitions and principles of use of methods of evolutionary designing and modeling, multiagent systems and fuzzy mathematical models at creation of hybrid components of intellectual systems are resulted. Also advantages and lacks of traditional methods are discussed

* Работа выполнена при поддержке: РФФИ (гранты № 08-01-00473), г/б № 2.1.2.1652.