

УДК 004.827:902.6]:069.51

О.В. ЛИПСКАЯ

Одесский национальный политехнический университет, Украина

## МЕТОД ПОИСКА И ГРУППИРОВАНИЯ ОБЪЕКТОВ БАЗ ДАННЫХ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ ВРЕМЕННЫХ ХАРАКТЕРИСТИК

Приведен метод поиска и группирования объектов баз данных в условиях неопределенности временных характеристик с использованием интегрирования. Временные характеристики представлены в виде лингвистических переменных из теории нечетких множеств и описываются трапециевидной функцией принадлежности. Решение задачи поиска осуществляется на основе понятия мощности нечеткого множества. Сделан переход от понятия мощности временной характеристики к площади фигуры, ограниченной функцией принадлежности лингвистической переменной, которой представлена временная характеристика. Решение задачи поиска временных характеристик сводится к интегрированию площадей фигур, полученных в результате пересечения эталонной и текущей анализируемой временной характеристики.

**Ключевые слова:** временная характеристика, нечеткость, неопределенность, лингвистическая переменная, нечеткая логика, мощность множества, интегрирование.

### Введение

Временные характеристики экспонатов археологического музея сформулированы в различных терминах и с разной степенью подробности. Кроме того, в силу специфики своей предметной области, они не строго формализованы и носят нечеткий характер. Так, например, временная характеристика может быть представлена таким терминами как «половина века», «треть века», «часть века» и т. д., которые являются неточными. Это затрудняет дальнейший анализ временных характеристик, в частности затрудняется выполнение операции поиска экспоната по временным характеристикам.

Для представления временных характеристик используется механизм нечеткой логики, предложенный Л. Заде [1]. С его помощью характеристику можно представить в виде лингвистической переменной со своим терм-множеством, состоящим из нечетких переменных [2]. Нечеткие переменные задаются своей областью определения и лингвистической функцией принадлежности. Использование базового понятия из теории нечетких множеств – лингвистической переменной – позволяет осуществить дальнейшую обработку временных характеристик, используя операции, применяемые над нечеткими множествами.

Для выполнения операции поиска необходимо проанализировать временные характеристики с заданной эталонной временной характеристикой.

В статье представлен метод решения задачи поиска временных характеристик по эталонной характеристике, основанный на интегрировании.

### 1. Задание временных характеристик

#### 1.1 Типы временных характеристик

Временные характеристики экспонатов представляют собой лингвистические переменные, состоящие из терм-множества, представленного в виде нечетких переменных. Например, лингвистическая переменная «четверть века» задается следующими нечеткими переменными: «половина века» = («первая», «вторая», «третья», «последняя»).

Характеристическая функция принадлежности  $\mu$  имеет трапециевидную форму (рис. 1):

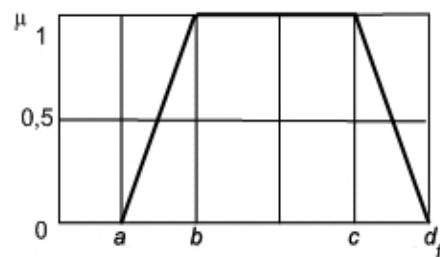


Рис. 1. Трапециевидная функция принадлежности

Функция принадлежности задается следующей формулой:

$$\mu = \begin{cases} 0, & t \leq a; \\ \frac{t-a}{b-a}, & a < t < b; \\ 1, & b \leq t \leq c; \\ \frac{d-t}{d-c}, & c < t < d, \\ 0, & t \geq d, \end{cases} \quad (1)$$

где  $(a, b, c, d)$  – вектор параметров нечеткой переменной, причем выполняется следующее условие:

$$a \leq b \leq c \leq d. \quad (2)$$

Исходя из графика трапециевидной функции принадлежности видно, что чем меньше разница между  $b$  и  $a$ , а также разница между  $c$  и  $d$ , тем ближе нечеткая переменная к четкой. В случае, если

$$\begin{aligned} b &= a; \\ d &= c, \end{aligned} \quad (3)$$

нечеткая переменная становится четкой, и ее функция принадлежности принимает прямоугольный вид (рис. 2):

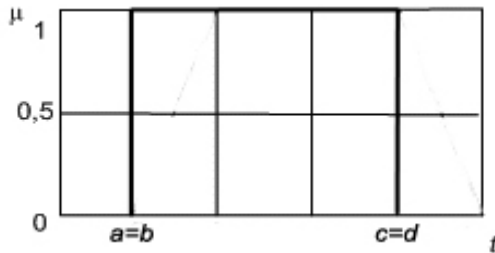


Рис. 2. Прямоугольная функция принадлежности

Вторым частным случаем является временная характеристика, заданная с максимальной нечеткостью. Ее функция принадлежности имеет треугольный вид (рис. 3) и соответствует следующему условию:

$$c = b. \quad (4)$$

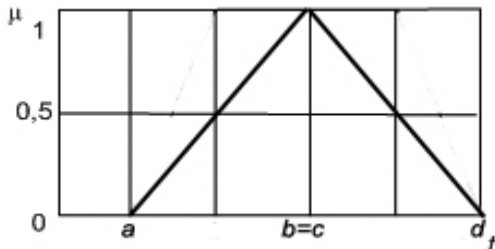


Рис.3. Треугольная функция принадлежности

### 1.2. Вычисление мощности временной характеристики

Мощностью временной характеристики  $M$  будем называть кардинальное число значений, на котором определена лингвистическая переменная этой временной характеристики:

$$M = \left\{ \sum \mu(t) / \mu(t) > 0 \right\}, \quad (5)$$

где  $\mu(t)$  – значение функции принадлежности в точке  $t$ .

Нахождение мощности временной характеристики сведем к задаче нахождения определенного интеграла как площади фигуры, ограниченной

функцией принадлежности временной характеристики, осью абсцисс и прямыми:  $x = a$  и  $x = b$ . Т.е.:

$$M = \int_a^b \frac{t-a}{b-a} dt + \int_b^c 1 dt + \int_c^d \frac{d-t}{d-c} dt. \quad (6)$$

Проведем ряд преобразований:

$$\begin{aligned} M &= \int_a^b \frac{t-a}{b-a} dt + \int_b^c 1 dt + \int_c^d \frac{d-t}{d-c} dt = \\ &= \frac{1}{b-a} \left( \frac{t^2}{2} - at \right) \Big|_a^b + t \Big|_b^c + \frac{1}{d-c} * \\ &* \left( dt - \frac{t^2}{2} \right) \Big|_c^d = \frac{t^2 - 2at}{2(b-a)} \Big|_a^b + t \Big|_b^c + \\ &+ \frac{2dt - t^2}{2(d-c)} \Big|_c^d = \frac{b-a}{2} + b - \\ &- c + \frac{d-c}{2} = \frac{c+d-a-b}{2}. \end{aligned} \quad (7)$$

Таким образом, мощность временной характеристики будет рассчитываться следующим образом:

$$M = \frac{c+d-a-b}{2}. \quad (8)$$

Следовательно, исходя из (3) и (7), для четкой временной характеристики мощность  $M$ :

$$M = d. \quad (9)$$

Соответственно, применяя (4) и (7), для временной характеристики с треугольной функцией принадлежности:

$$M = \frac{d-a}{2}. \quad (10)$$

## 2. Поиск по временным характеристикам

Для поиска экспоната по его временной характеристике необходимо проанализировать ее соответствие эталонной. Очевидно, что эталон задается в тех же терминах, что и временные характеристики. Анализ проводится в два этапа:

- по заданному требованию определяется нечеткая лингвистическая переменная, соответствующая эталонной датировке;
- выполнение логической операции «Пересечение» из теории нечетких множеств, т.е. получение области пересечения (ОП), над нечеткими переменными эталонной (ЭХ) и текущей анализируемой (ТХ) временных характеристик и вычисление степени соответствия временной характеристики экспоната требованию.

При этом степень соответствия  $S$  вычисляется по формуле:

$$S = \frac{M_{\text{ОП}}}{M_{\text{ТХ}}} * 100\%, \quad (11)$$

где  $M_{\text{ОП}}$  – мощность области пересечения;  $M_{\text{ТХ}}$  – мощность текущей временной характеристики.

Мощность текущей временной характеристики вычисляется по формулам (8) - (10).

Мощность области пересечения можно представить как сумму площадей фигур, ограниченных функциями принадлежности текущей и эталонной временных характеристик в области пересечения.

Если участок области пересечения ограничен прямой  $\mu = \frac{t-a}{b-a}$ , то его мощность подсчитывается следующим образом:

$$M = \int_x^y \frac{t-a}{b-a} dt = \frac{(y-x)(y+x-2a)}{2(b-a)}, \quad (12)$$

где  $x$  – нижний предел интегрирования;  $y$  – верхний предел интегрирования.

В случае, когда функция принадлежности достигает своего максимального значения, равного 1, на рассматриваемом интервале, то (12) принимает вид:

$$M = \frac{b-a}{2}, \quad (13)$$

Если участок области пересечения ограничен прямой  $\mu = \frac{d-t}{d-c}$ , то его мощность подсчитывается так:

$$M = \int_k^m \frac{d-t}{d-c} dt = \frac{(m-k)(2d-m-k)}{2(d-c)}, \quad (14)$$

где  $m$  – нижний предел интегрирования;  $k$  – верхний предел интегрирования.

В случае, когда функция принадлежности достигает своего максимального значения, равного 1, то (14) выглядит следующим образом:

$$M = \frac{d-c}{2}. \quad (15)$$

Если участок области пересечения ограничен прямой  $\mu=1$ , то его мощность подсчитывается по следующей формуле:

$$M = \int_p^r 1 dt = r - p, \quad (16)$$

где  $r$  – нижний предел интегрирования;  $p$  – верхний предел интегрирования.

Применяя формулы (11) - (13) можно вычислить мощность области пересечения любой сложности.

Рассмотрим пример решения задачи поиска временных характеристик методом интегрирования.

### 3. Пример решения задачи поиска временных характеристик методом интегрирования

Пусть эталонная временная характеристика: V-VI вв. Текущая анализируемая характеристика: IV-V вв. Проанализируем степень соответствия текущей характеристики (ТХ) к эталонной (ЭХ).

Приведем значения векторов параметров характеристик в табл. 1.

Таблица 1  
Вектор параметров ТХ и ЭХ

	ТХ	ЭХ
a	300	400
b	315	415
c	485	585
d	500	600

Функции принадлежности ЭХ и ТХ приведены на рис. 4.

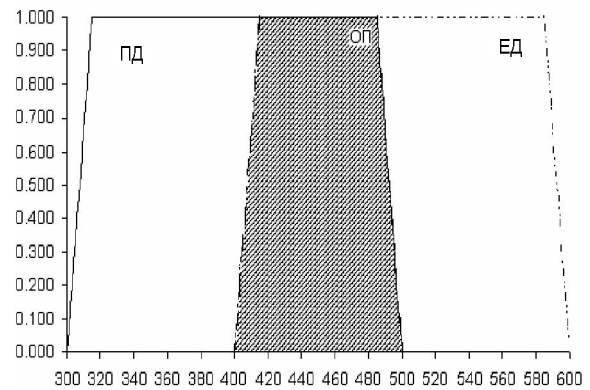


Рис.4. Функции принадлежности ЭХ и ТХ

Вычислим степень соответствия  $S$ , применяя (5) и (11):

$$S = \frac{M_{\text{ОП}}}{M_{\text{ТХ}}} * 100\% = \frac{85}{185} * 100\% = 45.95\%,$$

Теперь вычислим  $S$ , применяя метод интегрирования. Воспользуемся формулами (8), (11), (13), (15) и (16). Пределы интегрирования приведены в табл. 2.

Таблица 2  
Значения пределов интегрирования

x	400
y	415
k	485
m	500
r	415
p	485

Применяя формулы:

$$S = \frac{M_{\text{ОП}}}{M_{\text{ТХ}}} * 100\% = \frac{7,5 + 7,5 + 70}{185} * 100\% = 45,95\%.$$

Таким образом, применяя метод интегрирования для решения задачи поиска временных характеристик, был получен тот же результат, что и при применении классического подсчета мощности множеств.

### Заключение

Метод поиска и группирования объектов баз данных в условиях неопределенности временных характеристик свелся к вычислению площадей фигур, ограниченных характеристической функцией принадлежности текущей анализируемой и эталонной временных характеристик.

Были выведены формулы, позволяющие вычислять мощность любой временной характеристики,

а также области пересечения, сформированной при пересечении функций принадлежности эталонной и текущей анализируемой характеристик.

В дальнейшем, полученные формулы можно применять для анализа различных временных характеристик, обладающих нечеткостью.

### Литература

1. Zadeh L.A. *Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic* / L.A. Zadeh // *Fuzzy Sets Syst.* – 1997. – Vol. 90. – № 2. – P. 35-40.

2. Крисилова Г.Ф. *Решение задач хранения и поиска временных характеристик экспонатов археологического музея* / Г.Ф. Крисилова, О.В. Липская // *Труды одесского политехнического университета.* – 2007. – № 59. – С. 110-114.

*Поступила в редакцию 3.02.2010*

**Рецензент:** д-р техн. наук, проф., проф. кафедры Ф.В. Новиков, Харьковский национальный экономический университет, Харьков.

### МЕТОД ПОШУКУ ТА ГРУПУВАННЯ ОБ'ЄКТІВ БАЗ ДАНИХ В УМОВАХ НЕВИЗНАЧЕНОСТІ ЧАСОВИХ ХАРАКТЕРИСТИК

*О.В. Липська*

Наведено метод пошуку та групування об'єктів баз даних в умовах невизначеності часових характеристик, в якому застосовано інтегрування. Часові характеристики представлені у вигляді лінгвістичних змінних з теорії нечітких множин і описуються трапецієподібною функцією приналежності. Рішення задачі пошуку здійснюється на основі поняття потужності нечіткого множини. Зроблено перехід від поняття потужності часової характеристики до площі фігури, обмеженої функцією приналежності лінгвістичної змінної, якою подана часова характеристика. Рішення задачі пошуку часових характеристик зводиться до інтегрування площ фігур, отриманих в результаті перетину еталонної та поточної аналізованої часової характеристики.

**Ключові слова:** часова характеристика, нечіткість, невизначеність, лінгвістична змінна, нечітка логіка, потужність множини, інтегрування.

### THE METHOD OF FINDING AND GROUPING OBJECTS STORED IN DATABASES IN CONDITIONS OF TIMELINE CHARACTERISTICS INDETERMINACY

*O.V. Lipska*

The method of finding and grouping objects that are stored in databases in conditions of timeline characteristics indeterminacy is based on integration method. The timeline characteristics are presented in the form of linguistic variables of the fuzzy sets theory and are described as a trapezoidal membership function. Decision along the search problem is based on the concept of fuzzy power set. It is made a transition from the concept of timeline characteristics' power to shape of the area bounded by the membership function of linguistic variable, which describes the timeline characteristic. Solution of search timeline characteristics is reduced to integrating the area of figures, which are the result of intersection two timeline characteristics such as the currently analyzed characteristic and the model characteristic.

**Keywords:** timeline characteristics, ambiguity, uncertainty, linguistic variable, fuzzy logic, power sets, integration.

**Липская Ольга Васильевна** – аспирант кафедры Системного программного обеспечения Института компьютерных систем Одесского национального политехнического университета, Одесса, Украина, e-mail: o.lipskaya@gmail.com.