

## ИСПОЛЬЗОВАНИЕ НЕЧЕТКОГО ЛОГИЧЕСКОГО ВЫВОДА ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

А.Р. Вахитов, В.А. Силич

Томский политехнический университет

E-mail: arv@tpu.ru

*Показан способ интеллектуального анализа данных, основанный на использовании нечеткого логического вывода. Обсуждаются принципы реализации способа, области применения, а также преимущества по сравнению с другими способами обработки данных. Особое внимание уделяется практическому применению данного способа в предметной области, связанной с НИРС в вузе. Сделан вывод о целесообразности использования нечеткой логики в условиях неопределенности и неполноты знаний.*

### Ключевые слова:

*Интеллектуальные информационные системы, нечеткая логика, НИРС, анализ данных.*

### Key words:

*Intellectual information system, fuzzy logic, students' research effort, data analysis.*

Знания, которыми располагает человек, в какой-то степени всегда неполны, приближены, ненадежны. Тем не менее, людям на основе таких знаний удается делать достаточно обоснованные выводы и принимать разумные решения. Следовательно, чтобы интеллектуальные информационные системы были действительно полезны, они должны учитывать неполную определенность знаний и успешно действовать в таких условиях [1, 2]. Таким образом, неполная определенность и нечеткость имеющихся знаний – скорее типичная картина при анализе и оценке положения вещей, при построении выводов и рекомендаций, чем исключение. В процессе исследований по искусственному интеллекту для решения этой проблемы разработано несколько подходов.

Одним из таких подходов является нечеткая логика Л. Заде. В его работе [3] понятие множества расширено допущением, что функция принадлежности элемента к множеству может принимать любые значения в интервале  $[0..1]$ , а не только 0 или 1. Такие множества были названы нечеткими. Также Л. Заде были предложены логические операции над нечеткими множествами и предложено понятие лингвистической переменной, в качестве зна-

чений которой выступают нечеткие множества.

Модель нечеткой логики делает возможным реализацию в системе интеллектуальных функций, основанных на анализе неполной информации о предметной области, и построение удобного пользовательского интерфейса, в котором вывод данных имеет такие сходства с результатами человеческих рассуждений, как приближенность, неуверенность и субъективность. Кроме того, благодаря непрерывности функции принадлежности появляются преимущества в скорости обработки данных.

Функциональная схема процесса нечеткого вывода представлена на рис. 1 [4].

Выполнение первого этапа нечеткого вывода – фаззификации – осуществляет фаззификатор. За процедуру непосредственно нечеткого вывода ответственна машина нечеткого логического вывода, которая производит второй этап процесса вывода на основании задаваемой нечеткой базы знаний (набора правил), а также этап композиции. Дефаззификатор выполняет последний этап нечеткого вывода – дефаззификацию [5].

Исследуемая предметная область (НИРС в вузе) система была описана в терминах нечеткой логики. В соответствии с предложенным методом

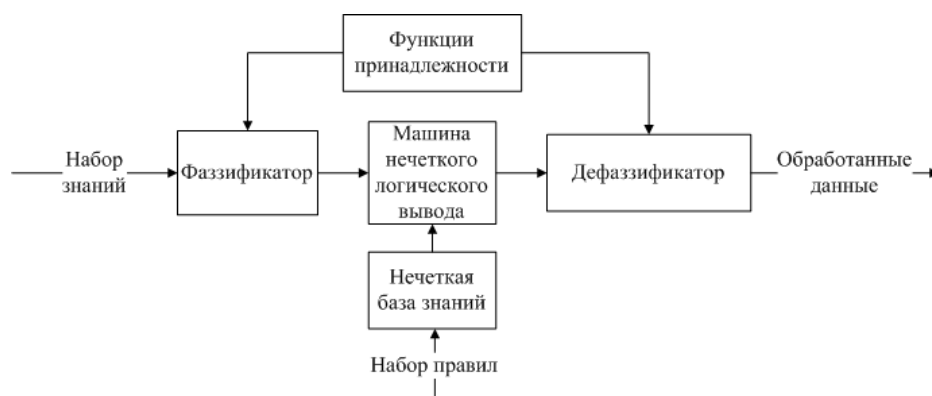


Рис. 1. Функциональная схема процесса нечеткого вывода

был осуществлен нечеткий логический вывод. Предметная область описывается следующими входными параметрами  $A_i$  и выходными параметрами  $B_j$ , таблица:

**Таблица.** Входные и выходные параметры системы

Обозначение	Описание	Универсум (список возможных значений)
Входные параметры		
$A_1$	Количество результатов НИРС	Множество натуральных чисел $n$
$A_2$	Сумма денег, полученных за НИРС	Множество положительных вещественных чисел $m$
$A_3$	Число страниц, опубликованных по итогам НИРС	$n$
$A_4$	Число наград за НИРС	$n$
$A_5$	Число поощрений по итогам НИРС	$n$
$A_6$	Сумма денег, затраченных на поощрения	$m$
Выходные параметры		
$B_1$	Премирование	$m$
$B_2$	Назначение преподавателю руководства НИРС	$n$
$B_3$	Рекомендация для поступления в аспирантуру	$n$
$B_4$	Рекомендация на стажировку	$n$
$B_5$	Командировка на конференцию	$n$

Все универсумы находятся в пределах измеримого диапазона с 5 степенями градации (термами): очень низкий  $[0..x_1]$ , средний  $[x_1..x_2]$ , высокий  $[x_2..x_3]$ . Конкретные значения  $x_i$  зависят от масштабов анализа данных (на уровне конкретного студента или преподавателя, кафедры, факультета либо вуза) и особенности измерения данного параметра. Исследуемая система описывается правилами вывода:

- $L_1: (A_5 \in [0..x_1] \wedge A_6 \in [0..x_1]) \wedge (A_1 \in [x_2..x_3] \vee A_2 \in [x_2..x_3] \vee A_3 \in [x_2..x_3] \vee A_4 \in [x_2..x_3]) \rightarrow B_1 \in [x_1..x_2];$
- $L_2: (A_5 \in [0..x_1] \wedge A_6 \in [0..x_1]) \wedge (A_1 \in [x_2..x_3] \wedge A_2 \in [x_2..x_3] \wedge A_3 \in [x_2..x_3] \wedge A_4 \in [x_2..x_3]) \rightarrow B_1 \in [x_2..x_3];$
- $L_3: (A_5 \in [x_1..x_2] \wedge A_6 \in [x_1..x_2]) \wedge (A_1 \in [x_2..x_3] \vee A_2 \in [x_2..x_3] \vee A_3 \in [x_2..x_3] \vee A_4 \in [x_2..x_3]) \rightarrow B_1 \in [0..x_1];$
- $L_4: (A_5 \in [x_1..x_2] \wedge A_6 \in [x_1..x_2]) \wedge (A_1 \in [x_2..x_3] \wedge A_2 \in [x_2..x_3] \wedge A_3 \in [x_2..x_3] \wedge A_4 \in [x_2..x_3]) \rightarrow B_1 \in [x_1..x_2];$
- $L_5: (A_5 \in [x_2..x_3] \wedge A_6 \in [x_2..x_3]) \wedge (A_1 \in [x_2..x_3] \wedge A_2 \in [x_2..x_3] \wedge A_3 \in [x_2..x_3] \wedge A_4 \in [x_2..x_3]) \rightarrow B_1 \in [0..x_1];$
- $L_6: (A_2 \in [x_2..x_3] \wedge A_4 \in [x_2..x_3] \wedge A_1 \in [0..x_1] \wedge A_5 \in [0..x_1]) \rightarrow B_2 \in [x_1..x_2];$
- $L_7: (A_2 \in [x_1..x_2] \wedge A_4 \in [x_1..x_2] \wedge A_1 \in [0..x_1] \wedge A_5 \in [0..x_1]) \rightarrow B_2 \in [0..x_1];$

- $L_7: (A_5 \in [x_2..x_3] \vee (A_1 \in [x_2..x_3] \wedge A_2 \in [x_2..x_3] \wedge A_4 \in [x_2..x_3])) \rightarrow B_3 \in [x_2..x_3];$
- $L_8: (A_5 \in [x_1..x_2] \vee (A_1 \in [x_1..x_2] \wedge A_2 \in [x_1..x_2] \wedge A_4 \in [x_1..x_2])) \rightarrow B_3 \in [x_1..x_2];$
- $L_9: (A_1 \in [x_2..x_3] \vee A_2 \in [x_2..x_3] \vee A_3 \in [x_2..x_3]) \rightarrow B_4 \in [x_1..x_2];$
- $L_{10}: (A_1 \in [x_2..x_3] \wedge A_2 \in [x_2..x_3] \wedge A_3 \in [x_2..x_3]) \rightarrow B_4 \in [x_2..x_3];$
- $L_{11}: (A_1 \in [x_2..x_3] \wedge A_3 \in [x_2..x_3]) \rightarrow B_5 \in [x_1..x_2];$
- $L_{12}: (A_1 \in [x_2..x_3] \vee A_3 \in [x_2..x_3]) \rightarrow B_5 \in [x_2..x_3];$
- $L_{13}: (A_1 \in [x_2..x_3] \wedge A_6 \in [x_2..x_3]) \rightarrow A_5 \in [x_2..x_3];$
- $L_{14}: (A_1 \in [x_2..x_3] \wedge A_6 \in [0..x_1]) \rightarrow A_5 \in [x_1..x_2];$
- $L_{15}: (A_1 \in [0..x_1] \vee A_6 \in [0..x_1]) \rightarrow A_5 \in [0..x_1].$

Следует отметить, что наряду с возможностью вывода выходных параметров на основе входных в системе также возможен вывод неизвестных входных параметров из известных входных параметров (правила вывода  $L_{13}-L_{15}$ ).

Рассмотрим алгоритм нечеткого вывода на конкретном примере. У одного из студентов необходимо доопределить значение  $A_5$ , зная значения  $A_1$  и  $A_6$ , используя затем полученные параметры для генерирования решения о том, заслуживает ли студент дополнительных поощрений.

Универсум значения числа поощрений  $A_5$  для этого студента находится в отрезке  $[0..6]$ . Начальное множество термов – низкое, среднее, высокое. Функции принадлежности  $\mu(A_5)$  показаны на рис. 2.

Универсум значения числа результатов  $A_1$  для этого студента находится в отрезке  $[0..20]$ . Начальное множество термов – {малое, среднее, большое}. Функции принадлежности  $\mu(A_1)$  приведены на рис. 3.

Универсум значения суммы денег  $A_6$  для этого студента находится в отрезке  $[0..12000]$ . Начальное множество термов – {малая, средняя, большая}. Функции принадлежности  $\mu(A_6)$  имеют следующий вид, рис. 4.

Нечеткий логический вывод производится в несколько этапов:

1. Этап фаззификации.

На основе значений  $A_1=15$  и  $A_6=7000$  была осуществлена фаззификация, в результате которой получены следующие степени уверенности в значениях входных переменных:

- Число достижений  $A_1$  большое – 0,65; среднее – 0,7; малое – 0,35.
- Сумма денег  $A_6$  большая – 1; средняя – 0,5; малая – 0.

2. Этап нечеткого вывода.

На данном этапе вычислены степени уверенности посылок правил  $L_{13}-L_{15}$ , представляющих из себя нечеткие импликации:

- $L_{13}: \min(A_1 \in [x_2..x_3] \wedge A_6 \in [x_2..x_3]) = \min(0,65; 1) = 0,65.$
- $L_{14}: \min(A_1 \in [x_2..x_3] \wedge A_6 \in [0..x_1]) = \min(0,65; 0) = 0.$
- $L_{15}: \max(A_1 \in [0..x_1] \vee A_6 \in [0..x_1]) = \max(0,35; 0) = 0,35.$

3. Этап композиции.

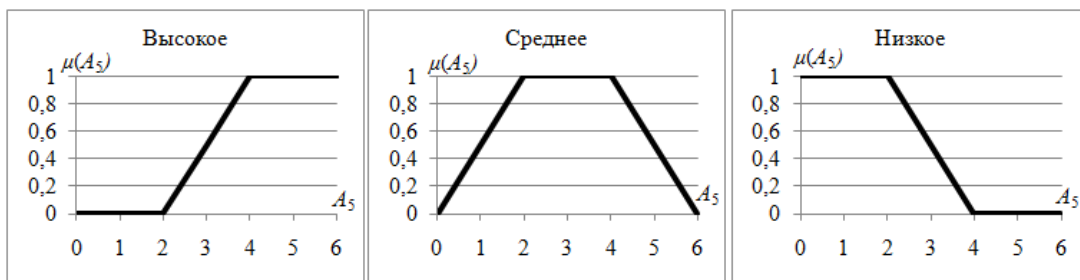


Рис. 2. Функции принадлежности  $\mu(A_5)$

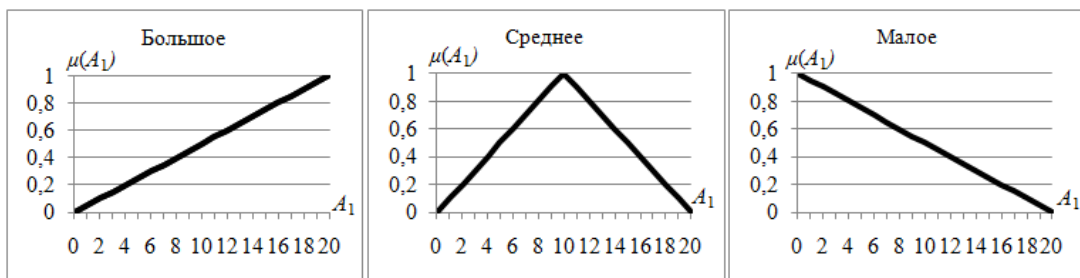


Рис. 3. Функции принадлежности  $\mu(A_1)$

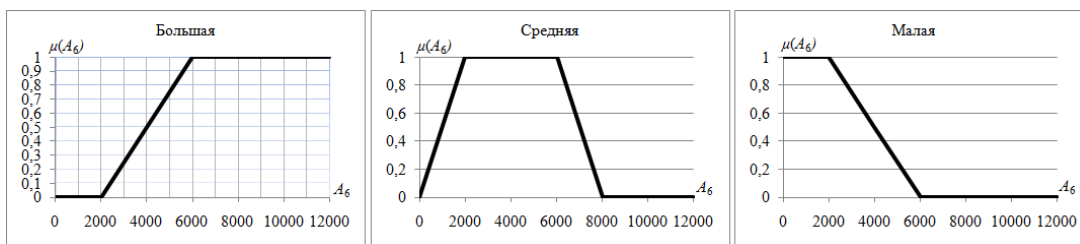


Рис. 4. Функции принадлежности  $\mu(A_6)$

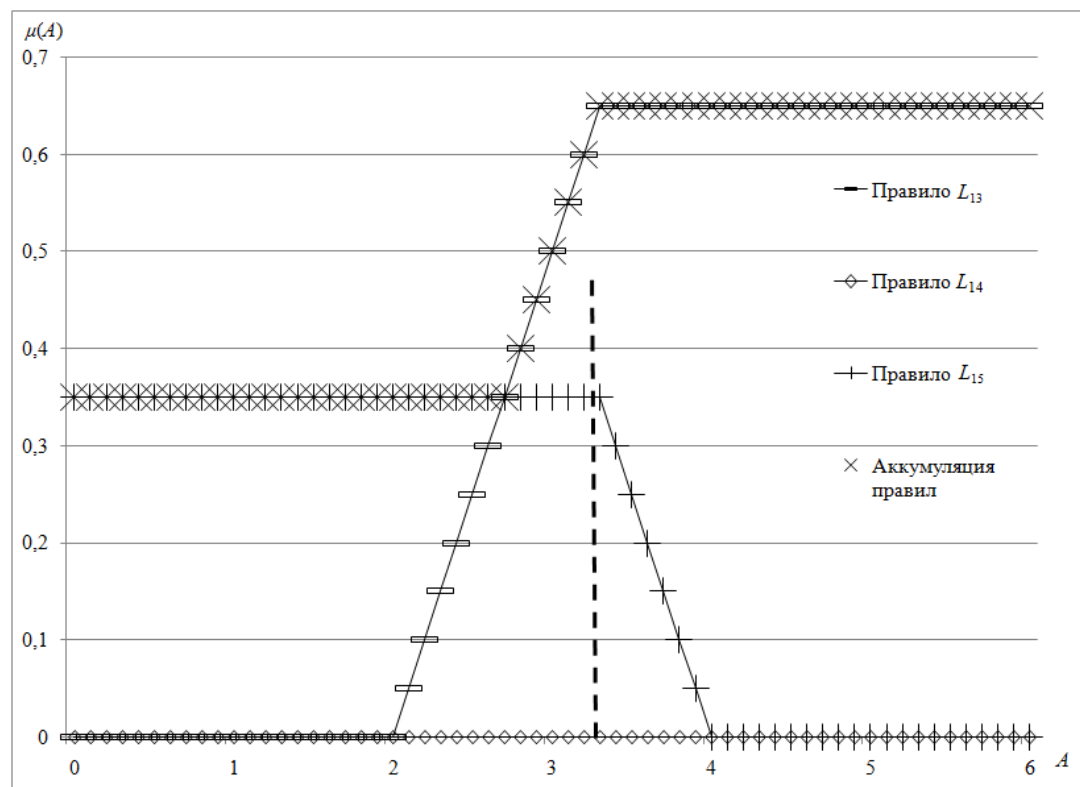


Рис. 5. Аккумуляция правил

Степень уверенности заключения задается функцией принадлежности соответствующего термина. Поэтому с использованием определения нечеткой импликации как минимума левой и правой частей получены новые нечеткие переменные, соответствующую степени уверенности в значении выходных данных при применении к заданным входным соответствующего правила.

Затем была проведена аккумуляция – объединение результатов применения всех правил, рис. 5.

В результате была получена функция принадлежности для числа поощрений  $A_5$ , которая говорит о степени уверенности в значении искомого параметра на основе входных параметров и правил нечеткого логического вывода.

#### 4. Этап дефазификации.

Для преобразования нечеткого набора значений к точным был использован метод первого мак-

симула, в результате чего было определено, что число поощрений находится в диапазоне «среднее» и равно примерно 3.

Затем полученные данные были использованы для определения выходных параметров  $B_i$ . Зная, что  $A_1=15$ ,  $A_5=3$ ,  $A_6=7000$ , согласно правилу нечеткого логического вывода  $L_3$ :

$$(A_5 \in [x_1..x_2] \wedge A_6 \in [x_1..x_2]) \wedge (A_1 \in [x_2..x_3] \vee A_2 \in [x_2..x_3] \vee A_3 \in [x_2..x_3] \vee A_4 \in [x_2..x_3]) \rightarrow B_1 \in [0..x_1]$$

было определено, что с данными показателями НИРС этот студент заслуживает премирования в размере [0..2000].

Таким образом, использование нечеткого логического вывода делает возможным получение новых знаний на основе анализа существующих данных даже в условиях неполноты и приближенности сведений об исследуемой предметной области.

#### СПИСОК ЛИТЕРАТУРЫ

1. Brachman R., Sefridge P. Knowledge representation support for data archeology // Intelligent and Cooperative Information Systems. – 1993. – № 2. – P. 113–120.
2. Борисов А.Н., Алексеев А.В., Меркурьева Г.В. и др. Обработка нечеткой информации в системах принятия решений. – М.: Радио и связь, 1989. – 304 с.

3. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976. – 165 с.
4. Корниенко А.В. Интеллектуальные информационные системы в экономике. – Томск: Изд-во ТПУ, 2008. – 177 с.
5. Кузнецов С.Д. Неопределенная информация и трехзначная логика // СУБД. – 1997. – № 5. – С. 65–67.

Поступила 08.09.2010 г.

УДК 004.89

## ВЫБОР КЛАССА МАТЕМАТИЧЕСКОЙ МОДЕЛИ СИСТЕМЫ НА ОСНОВЕ МЕТОДА СААТИ И ИНТЕГРАЛЬНЫХ КРИТЕРИЕВ

А.Р. Вахитов, В.А. Силич

Томский политехнический университет

E-mail: arv@tpu.ru

*Показано обоснование выбора класса математической модели информационной системы, содержащей сведения о НИРС в вузе. Произведено описание альтернативных классов моделей, основных критериев выбора и применения метода Саати и интегральных критериев для выбора наиболее подходящего варианта из имеющихся альтернатив. Сделан вывод о целесообразности использования нечеткой логики в качестве класса математической модели для исследуемой предметной области.*

#### Ключевые слова:

Математическая модель, НИРС, метод Саати, интегральные критерии.

#### Key words:

Mathematical model, students' research effort, Saati's method, integral criterias.

Создание математического обеспечения информационной системы предполагает обоснование выбора класса математической модели из множества  $X$  альтернативных вариантов  $x_i$ , а также непосредственное описание предметной области в терминах выбранного класса. К числу основных логических моделей, для которых разработаны формальные методы логического вывода, относятся [1, 2]:

- $x_1$  – исчисление высказываний;
- $x_2$  – исчисление предикатов;
- $x_3$  – семантические сети;
- $x_4$  – дескриптивная логика;
- $x_5$  – нечеткая логика.

Исследуемой предметной областью является НИРС в вузе. Обоснование выбора класса математической модели является важным этапом при разработке системы, так как здесь должны учитывать-