

ПРЕДСТАВЛЕНИЕ НЕЧЕТКОЙ ЭКСПЕРТНОЙ ИНФОРМАЦИИ, ПОЛУЧЕННОЙ ПРИ ОЦЕНИВАНИИ КАЧЕСТВЕННЫХ ПРИЗНАКОВ

УДК 004.827

УЛЬЯНОВСКАЯ Юлия Викторовна

к.т.н., доцент, доцент кафедры информационных систем и технологий Академии таможенной службы Украины.

Научные интересы: экспертные системы, методы обработки нечетких данных

e-mail: uyy@rambler.ru

ВВЕДЕНИЕ

Одним из направлений развития информационных технологий является разработка интеллектуальных автоматизированных систем поддержки принятия решения, работа которых основана на экспертной информации. Класс задач, решение которых происходит на основе на экспертных данных, достаточно велик. К таким задачам относятся задачи планирования, оценки платежеспособности предприятий и физических лиц, оценки рисков, задачи классификации и экспертного оценивания предметов и т.д. Информация, которой описываются объекты предметных областей, является совокупностью как количественных так и качественных данных, которые в большинстве случаев носят нечеткий, лингвистический характер. При обработке экспертной информации и построении моделей принятия решений необходимо не только описывать объекты, но и сравнивать объекты между собой, знать насколько один объект превосходит другой, в каком порядке относительно друг друга они располагаются. В этом случае встает вопрос о ранжировках и построении отношений на множестве альтернатив.

АНАЛИЗ ПРЕДЫДУЩИХ ИССЛЕДОВАНИЙ

Методы, используемые в настоящее время для получения и обработки экспертных оценок, достаточно многочисленны и разнообразны. Выбор метода определяется характером анализируемой информации. Для

получения качественных оценок используются парные сравнения, множественные сравнения, методы ранжирования и т.д. Для получения количественных оценок используются непосредственная численная оценка альтернатив, метод Черчмена-Акофа и др. Применение методов анализа и обработки экспертной информации зависит от характера измерений. Некорректное использование результатов экспертного оценивания может привести к ошибочным выводам. Избежать ошибок можно лишь исследовав характер измерений и обусловленные им возможные методы преобразования полученной экспертной информации. В работах Л.Д. Мешалкина, А.И. Орлова, В.Б. Кузьмина, Б.Г. Литвака исследуется проблема допустимости использования различных результатов экспертного ранжирования [1, 2].

В экспертизе обычно участвует группа лиц, при этом не исключена возможность, что одна и та же группа специалистов может сходить в мнении относительно оценки одного объекта экспертизы и иметь расхождения относительно другого объекта. Поэтому одной из наиболее важных проблем при обработке мнений экспертов является проверка согласованности, классификация или обобщение экспертной информации [3]. Разработан ряд методов такой проверки. Классические методы математической статистики не всегда дают адекватный результат. В частности проблемы возникают при попытке усреднения мнений экспертов. Статистические методы проверки согласованности

зависят от математической природы ответов экспертов. Соответствующие статистические теории весьма трудны, если эти ответы - ранжировки или разбиения, и достаточно просты, если ответы - результаты независимых парных сравнений. Анализируя выше сказанное, можно сделать вывод, что не смотря на широкое развитие методов экспертных оценок, ряд вопросов, связанных с нечеткими экспертными оценками остается открытыми. Не достаточно разработаны методы сравнительной оценки методов экспертного оценивания, обработки и получения группового мнения, а также методы ранжирования нечетких альтернатив.

ПОСТАНОВКА ЗАДАЧИ

Целью данной работы является разработка способа обработки и представления нечеткой экспертной информации с помощью методов нечеткой логики с целью дальнейшей разработки методов получения обобщенного группового мнения экспертов.

Описание задачи. Пусть имеется n экспертов $X = \{x_i\} = \{x_1, x_2, \dots, x_n\}$. Экспертам предлагается оценить множество $A = \{a_j\} = \{a_1, a_2, \dots, a_m\}$ альтернатив определяя их значениями терм - множества $T^i = \{t^i_1, t^i_2, \dots, t^i_{k_j}\}$. Определим оптимальное множество лингвистических значений альтернатив следующим образом. Пусть относительно альтернативы a_j имеем терм-множество $T^j_i = \{t^j_{i1}, t^j_{i2}, \dots, t^j_{ik_j}\}$ лингвистических значений, сформированного экспертом x_i . Тогда $T^j = \bigcap_i T^j_i$

$i=1, \dots, n$. Эксперт при этом указывает число μ из интервала $[0;1]$, характеризующее степень соответствия альтернативы a_j выбранному терму t^j_k . В этом случае получим дискретную функцию принадлежности, построенную прямым методом. Мнения экспертов при этом могут совпадать либо быть не согласованными. Возникает вопрос об обработке полученных данных.

Сформируем мнения экспертов относительно каждой альтернативы a_j в матрицу $n \times k_j$.

		t^j_1	t^j_2	...	$t^j_{k_j}$	
$M^i_{jT^j} =$	x_1	$\mu^1_1(a_j)$	$\mu^1_2(a_j)$...	$\mu^1_{k_j}(a_j)$	(1)
	x_2	$\mu^2_1(a_j)$	$\mu^2_2(a_j)$...	$\mu^2_{k_j}(a_j)$	
	
	x_n	$\mu^n_1(a_j)$	$\mu^n_2(a_j)$...	$\mu^n_{k_j}(a_j)$	

РЕШЕНИЕ ЗАДАЧИ

Для простоты обозначим $\mu^j_k(a_j) = \mu^j_{jk}$ - значение функции принадлежности альтернативы a_j терму t^j_k в соответствии с мнением x_i - го эксперта.

При таком подходе необходимо учитывать компетентность экспертов, однако в прямых методах построения функции принадлежности для группы экспертов деликатно обходится вопрос о том, кто же назначает весовые коэффициенты оценкам экспертов. Одним из методов решения этого вопроса можно использовать частоту правильных оценок эксперта, основанную на статистических данных. Пусть x_i эксперт принимал участие в N экспертизах, причем в M из них его результаты были правильными. Тогда вес x_i эксперта обозначим через $w_i = M/N$, при этом $0 \leq w_i \leq 1$.

Таким образом, в соответствии с обозначениями нечеткой алгебры каждая альтернатива будет задана выражением

$$a_j = (T^j, \mu_T(a_j), w^j) \tag{2}$$

где w^j - средний вес экспертного мнения относительно альтернативы a_j . Проиллюстрируем сказанное графически (рис. 1):

Введем несколько условий для выражения (1):

$$\exists k' : k \leq k' \leq k_j, \sum_{i=1}^n \mu^i_{jk'} \neq 0, \sum_{i=1}^n \mu^i_{jk} = 0, k \neq k' \tag{3}$$

$$\exists t^j_k : \max_{i,h} (\mu^i_{jk} - \mu^h_{jk}) \leq 0.5 \tag{4}$$

Если выполняются условия (3) и (4) то спроектировав оценки на ось (μ, T) получим достаточно хорошо сгруппированные мнения и функции принадлежности имеют один центр кластеризации. В этом случае будем называть оценку экспертов согласованной. Если при тех же условиях центров несколько, то оценку экспертов будем называть условно согласованной. Мнения экспертов будем считать не согласованными, если имеется несколько ненулевых столбцов.

Построенное выше пространство (рис. 1) назовем экспертным пространством и будем обозначать его через A^n_m , каждая точка которого a^j_i - оценка x_i -м экспертом a_j -й альтернативы описывается вектором $a^j_i = (t^j_k, \mu^j_{jk}, w^j_i)$.

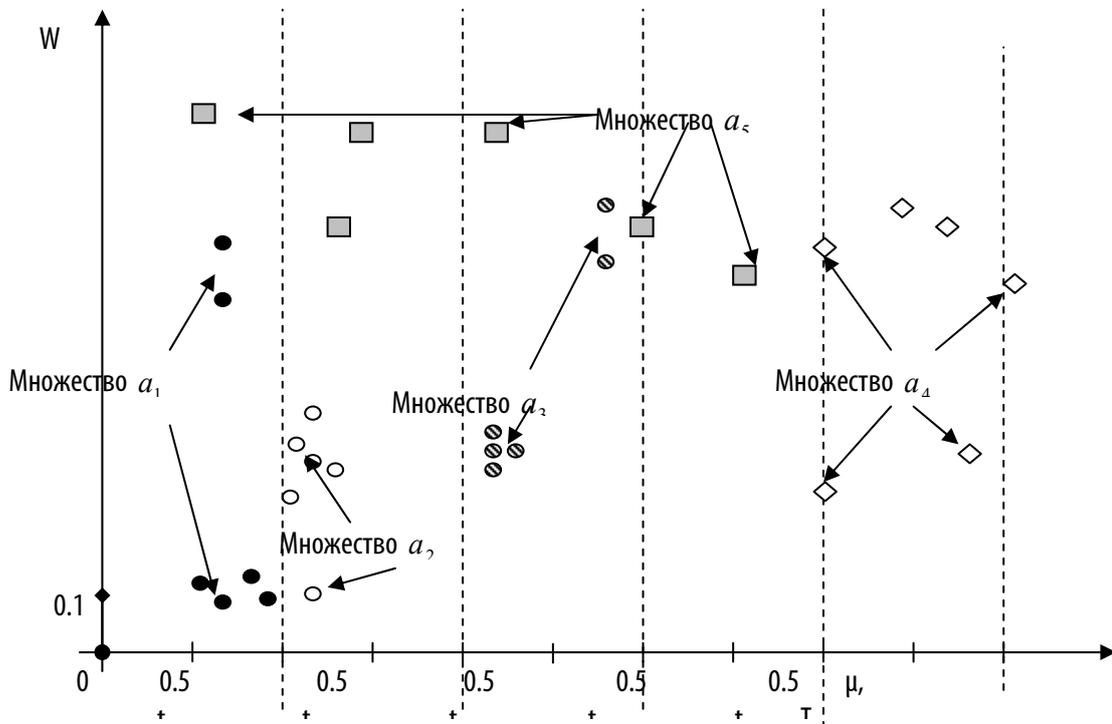


Рисунок 1 – Графическое представление оценок шестью экспертами пяти альтернатив

Рассмотрим способы обработки экспертных мнений и получения группового мнения. Отметим, что в процессе оценки эксперт может испытывать затруднения, которые выражаются в невозможности отнести оценку альтернативы к одному из существующих термов и необходимости размещать ее на границе двух термов или выносить за пределы крайних термов. Пусть в первом случае оценка лежит на границе k -го и $k+1$ терма. Тогда $\mu_{jk}^i = 0.5$ и $\mu_{jk+1}^i = 0.5$. Во втором случае методы нечеткой алгебры позволяют расширить множество лингвистических значений признака с помощью квантификаторов следующим образом.

Для получения обобщенной оценки a_j альтернативы проведем обработку данных (1) по одному из методов кластеризации с заданием количества кластеров, соответствующих количеству термов, в которых лежат значения альтернативы. Поскольку каждой альтернативе соответствует своя матрица, после ее обработки предложенным выше способом получим представление альтернативы в виде

$$a_j = \{ \langle \mu_{j1} / t_1 \rangle, \langle \mu_{j2} / t_2 \rangle, \dots, \langle \mu_{jk_j} / t_{k_j} \rangle \} \quad (5)$$

где μ_{jk_j} – центр кластера, определяемый для каждого терма из множеств лингвистических значений альтернативы.

Пусть лингвистическая переменная x задана множеством $\langle T, U, G, M \rangle$, где $T = \{t_1, t_2, \dots, t_n\}$ – терм множество лингвистических значений признака, U – универсальное множество, G – синтаксические правила порождающие новые термы с использованием квантификаторов «не», «очень» и «более-менее», M – семантические правила, которые заданы в виде таблицы (1):

Таблица 1 –

Семантические правила для образования новых термов

Квантификатор	Функция принадлежности ($u \in U$)
не t	$1 - \mu_t(u)$
очень t	$(\mu_t(u))^2$
более-менее t	$\sqrt{\mu_t(u)}$

Рассмотрим для примера альтернативу a_5 . Пусть эксперты оценивают ее следующим образом (табл. 2).

Таблиця 2 –
Экспертные данные для альтернативы a_5

Эксперт	Экспертные данные		
	t_k^s	μ_{5k}^i	w_i
x_1	1	0.5	0.9
x_2	2	0.35	0.7
x_3	2	0.4	0.85
x_4	3	0.2	0.85
x_5	3	0.5	0.7
x_5	4	0.5	0.7
x_6	4	0.6	0.6

Проведем кластеризацию в пакете MATLAB. Поскольку альтернатива принимает значимые значения только для четырех из пяти термов, зададим в параметрах кластеризации четыре кластера. Каждый центр кластера будет являться функцией принадлежности альтернативы соответствующему терму. Результаты кластеризации приведены на рис. 2.

В соответствии с выражением (5) и результатами кластеризации, приведенными на рис. 2, альтернатива a_5 будет задана выражением (6):

$$a_5 = \{ \langle 0.5 / t_1 \rangle, \langle 0.37498 / t_2 \rangle, \langle 0.34927 / t_3 \rangle, \langle 0.5499 / t_4 \rangle, \langle 0 / t_5 \rangle \}. \quad (6)$$

Проведя аналогичную обработку для каждой из альтернатив, заданных в виде (5) можем сгруппировать их в матрицу $A_{обп}$:

$$A_{обп} = \begin{matrix} & \begin{matrix} t_1 & t_2 & \dots & t_k \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ \dots \\ a_n \end{matrix} & \begin{matrix} \mu_{11} & \mu_{12} & \dots & \mu_{1k} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2k} \\ \dots & \dots & \dots & \dots \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nk} \end{matrix} \end{matrix}, \quad (7)$$

где μ_{ij} – функция принадлежности t_j терму a_i альтернативы. Выражение (7) является матричным способом представления нечеткой экспертной информации,

полученной в результате обработки мнений n экспертов.

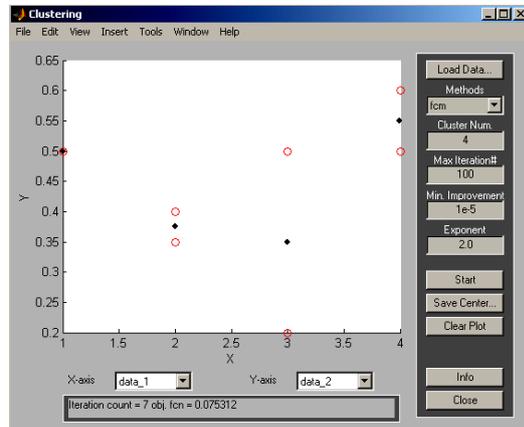


Рисунок 2 – Кластеризация с заданием количества кластеров

ВЫВОДЫ

Существующие на сегодняшний день методы обработки экспертной информации не всегда дают адекватный результат. Проблемы возникают при обработке нечетких экспертных мнений, различающихся между собой. Для решения этих проблем в работе был использован аппарат нечетких множеств. На его основе предложен способ определения оптимального множества лингвистических значений альтернатив $A = \{a_i\} = \{a_1, a_2, \dots, a_m\}$ при оценке их n экспертами $X = \{x_i\} = \{x_1, x_2, \dots, x_n\}$ значениями терм – множества $T = \{t_1, t_2, \dots, t_k\}$. В статье разработан метод обработки и представления нечеткой экспертной информации с помощью методов нечеткой логики с целью дальнейшей разработки методов получения обобщенного группового мнения экспертов, который позволяет, в отличие от существующих методов, учитывать мнение каждого эксперта, избегая при этом искусственного усреднения результатов.

ЛИТЕРАТУРА:

1. Статистические методы анализа экспертных оценок // Сб. науч. ст. Ученые записки по статистике. – 1997. – Т.29.
2. Литвак Б.Г. Экспертная информация: Методы получения и анализа / Б.Г. Литвак. – М.: Радио и связь, 1982. – 184 с.
3. Орлов А.И. Прикладная статистика: Учебник. / А.И. Орлов. – М.: Экзамен, 2004. – 656 с.