УДК 519.816

Н.М. Кораблёв

Харьковский национальный университет радиоэлектроники, Харьков

ФОРМАЛИЗАЦИЯ НЕЧЕТКОЙ ИНФОРМАЦИИ НА ОСНОВЕ ОПРОСА ОДНОГО ЭКСПЕРТА

Рассматривается построение моделей экспертного оценивания признаков на основе семантических пространств путем опроса одного эксперта о типичных значениях термов, а также о разбиении универсального множества признака. Эти модели могут применяться для формализации информации в рамках как качественных, так и количественных признаков.

функция принадлежности, терм-множество, признак, семантическое пространство

Введение

Важным этапом обработки информации является этап ее формализации, то есть этап представления этой информации в виде, позволяющем на следующих этапах ее обработки применять аппараты известных математических теорий. Известно, что нечеткая экспертная информация трудно формализуема при использовании традиционных подходов [1, 2]. Формализация нечеткой информации на основе классических и субъективных вероятностей оказалась малоэффективной в связи с ограничительными требованиями их использования.

Модельный подход на основе аппарата теории нечетких множеств позволил устранить недостатки традиционных формализаций нечеткой экспертной информации [1 – 4]. С точки зрения этого подхода моделями экспертного оценивания признаков служат семантические пространства (СП), термы которых соответствуют уровням вербальных шкал, используемых для оценивания признаков [5, 6]. Формализация нечеткой информации на основе СП является фундаментом, на котором строятся методы обработки информации непосредственно в рамках аппарата теории нечетких множеств, которые используются для решения различных практических задач [3]. Однако, не все модели, построенные на основе СП, обладают свойствами, обеспечивающими успешность решения практических задач на основе этих моделей. Одним из таких свойств является свойство полноты, которое состоит в возможности описания каждого из элементов универсального множества в лингвистических термах этого пространства [4, 7, 8].

Субъективное и приближенное представление значений функций принадлежности (ФП) термов СП может приводить к неадекватности нечетких моделей субъективным суждениям и исходным данным. Широкое использование такого представления требует проведения исследований, направленных на повышение адекватности как самих моделей формализации, так и нечетких моделей на их основе. Критерием адекватности может служить «естественность» заключений, получаемых на основе этих моделей [9, 10].

Требования к моделям экспертного оценивания признаков, как правило, формулируются в рамках конкретной задачи, а качество построенных моделей зависит от опыта экспертов. Это определяется тем, что методы формализации ограничены как способами получения информации от экспертов, так и видом информации, а также отсутствием общих требований к ФП используемой для формализации совокупности нечетких множеств. При этом отсутствуют методы формализации экспертного оценивания качественных признаков на универсальных множествах, элементами которых являются значения интенсивности проявлений этих признаков, вследствие чего отсутствуют методы, позволяющие использовать формализации качественных признаков для определения количественных показателей их проявлений. Поэтому разработка методов формализации нечеткой экспертной информации на основе СП путем опроса одного эксперта или группы экспертов является актуальной задачей.

Статья посвящена формализации нечеткой экспертной информации или построению моделей экспертного оценивания признаков на основе СП путем опроса одного эксперта о типичных значениях термов и о разбиении универсального множества признака.

Постановка задачи. Пусть X – некоторое множество x с ФП $\mu_{\tilde{A}}(x)$ к некоторому множеству $\tilde{A}$ с терм-множеством T(X) = $\{X_1, X_2, ..., X_m\}$, где X_1 – терм, соответствующий минимальной интенсивности проявления признака, X_m – терм, соответствующий максимальной интенсивности проявления признака на универсальном множестве $U = [a, b]$. Предположим, что экспертом определены типичные для термов X_l, $l = \overline{1,m}$ интервалы (x_l^1, x_l^2), $l = \overline{1,m}$, т.е. интервалы, для всех точек которых ФП соответствующих термов равны единице.

Необходимо путем опроса одного эксперта о типичных значениях термов или о разбиении универсального множества признака построить модели экспертного оценивания признаков и их полезности для решения практических задач на основе СП, которые удовлетворяли бы требованиям к ФП $\mu_l(x)$, $l = \overline{1,m}$

их терм-множеств [4]: 1) существует $\hat{U}_l \neq \emptyset$, где $\hat{U}_l = \{x \in U : \mu_l(x) = 1\}$ – точка или отрезок; 2) если $\hat{U}_l = \{x \in U : \mu_l(x) = 1\}$, тогда $\mu_l(x)$, $l = \overline{1,m}$ не убывает слева от $\hat{U}_l$, не возрастает справа от $\hat{U}_l$; 3) ФП $\mu_l(x)$, $l = \overline{1,m}$ имеют не более двух точек разрыва первого рода; 4) для каждого $x \in U$ существует l, $l = \overline{1,m} : \mu_l(x) \neq 0$; 5) $\forall x \in U \quad \sum_{l=1}^{m} \mu_l(x) = 1$.

ФП, которые удовлетворяют этим требованиям, обладают свойством полноты [4].

Если эксперт по каким-либо причинам затрудняется определить типичные интервалы для термов СП, то он может разбить универсальное множество на непересекающиеся интервалы, каждый из которых отображается на один из термов. Для этого случая также необходимо построить ФП $\mu_l(x)$, $l = \overline{1,m}$, которые обладали бы свойством полноты.

Математической основой построения методов обработки и анализа нечеткой информации с использованием нечетких множеств является алгебра нечетких чисел. Поэтому рассмотрим сначала построение совокупности нечетких чисел, используемых для формализации лингвистических значений признаков.

1. Построение совокупности нечетких чисел для формализации лингвистических значений признаков

Известно [5], что нечеткое число $\tilde{A}$ с ФП $\mu_{\tilde{A}}(x)$ называется нормальным, если $\max_x \mu_{\tilde{A}}(x) = 1$, $x \in R$. Нечеткое число $\tilde{A}$ с ФП $\mu_{\tilde{A}}(x)$ называется унимодальным, если существует единственная точка $x \in R$, для которой выполняется равенство $\mu_{\tilde{A}}(x) = 1$. Нечеткое число $\tilde{A}$ с ФП $\mu_{\tilde{A}}(x)$ называется многомодальным, если точка $x \in R$: $\mu_{\tilde{A}}(x) = 1$ не является единственной, и толерантным, если существуют интервалы, для всех точек которых выполняется равенство $\mu_{\tilde{A}}(x) = 1$. Этот интервал называется интервалом толерантности нечеткого числа $\tilde{A}$. Рассмотрим толерантные и унимодальные (L – R)-числа с ФП [5]

$$\mu_{\tilde{A}}(x) = \begin{cases} L((a_1 - x)/L), \ 0 \le (a_1 - x)/L \le 1, a_L > 0; \\ R((x - a_2)/a_R), \ 0 \le (x - a_2)/a_R \le 1, a_R > 0; \\ 1, \ (a_1 - x)/a_L < 0 \cap (x - a_2)/a_L < 0; \\ 0, \ (a_1 - x)/a_L > 1 \cap (x - a_2)/a_R > 1 \end{cases} \quad (1)$$

и следующими условиями на функции L и R:

1) L(0) = R(0) = 1; L(1) = R(1) = 0;

2) L(x) и R(x) – монотонно убывающие функции на множестве [0,1].

Нечеткое число $\tilde{A}$ записывается в виде $\tilde{A} \equiv (a_1, a_2, a_L, a_R)$ (или $\mu_{\tilde{A}}(x) \equiv (a_1, a_2, a_L, a_R)$), где a_1, a_2, a_L, a_R являются параметрами толерантного (L – R)-числа $\tilde{A}$. Отрезок $[a_1, a_2]$ называется интервалом толерантности, а a_L и a_R – соответственно левым и правым коэффициентами нечеткости. Функция $L((a_1 - x)/L)$ является левой границей ФП толерантного (L – R)-числа, а функция $R((x - a_2)/a_R)$ является правой границей ФП толерантного (L – R)-числа. При $a_L = 0$ предполагается, что $L((a_1 - x)/L) = 0$, при $a_R = 0$ предполагается, что $R((x - a_2)/a_R) = 0$.

Унимодальное (*L-R*)-число $\tilde{A}$ имеет ФП толерантного (*L-R*)-числа при условии $a_1 = a_2$. Символически унимодальное (*L-R*)-число $\tilde{A}$ записывается в виде $\tilde{A} \equiv (a_1, a_L, a_R)$.

Обозначим через Λ совокупность всех толерантных и унимодальных чисел с условиями на функции L и R. Элементы совокупности Λ будет называть Λ – числами, которые используем для построения ФП терм-множеств на основе опроса одного эксперта.

2. Построение функций принадлежности на основе опроса эксперта о типичных значениях термов

Построим СП для признака X с терм-множеством T(X) = $\{X_1, X_2, ..., X_m\}$. Будем считать, что нечеткие числа, соответствующие термам СП, являются Λ-числами. Пусть $\mu_l(x)$, $l = \overline{1,m}$ – ФП термов X_l, $l = \overline{1,m}$. Тогда при четном m

$$\mu_1(x) = \begin{cases} 1, & a \le x \le x_1^2; \\ L\left(\left(x - x_1^2\right) / \left(x_2^1 - x_1^2\right)\right), & x_1^2 < x \le x_2^1; \\ 0, & x_2^1 < x \le b. \end{cases} \quad (2)$$

$$\mu_2(x) = \begin{cases} 0, & a \le x \le x_1^2; \\ L\left(\left(x_2^1 - x\right) / \left(x_2^1 - x_1^2\right)\right), & x_1^2 < x \le x_2^1; \\ 1, & x_2^1 < x \le x_2^2; \\ R\left(\left(x - x_2^2\right) / \left(x_3^1 - x_2^2\right)\right), & x_2^2 < x \le x_3^1; \\ 0, & x_3^1 < x \le b. \end{cases} \quad (3)$$

$$\mu_{m-1}(x) = \begin{cases} 0, & a \le x \le x_{m-2}^2; \\ R\left(\left(x_{m-1}^1 - x\right) / \left(x_{m-1}^1 - x_{m-2}^2\right)\right), & x_{m-2}^2 < x \le x_{m-1}^1; \\ 1, & x_{m-1}^1 < x \le x_{m-1}^2; \\ L\left(\left(x - x_{m-1}^2\right) / \left(x_m^1 - x_{m-1}^2\right)\right), & x_{m-1}^2 < x \le x_m^1; \\ 0, & x_m^1 < x \le b. \end{cases} \quad (4)$$

. . . .

$$\mu_m(x)=\begin{cases}0, & a\le x\le x_{m-1}^2;\\ L\left(\left(x_m^1-x\right)/\left(x_m^1-x_{m-1}^2\right)\right), & x_{m-1}^2<x\le x_m^1;\\ 1, & x_m^1<x\le b.\end{cases}\qquad(5)$$

При нечетном m у двух последних ФП L(x) и R(x) меняются местами.

Построенное СП отражает знания эксперта, принимающего участие в опросе и его субъективное мнение. Если возможно получение дополнительной информации относительно значений ФП в точках универсального множества, лежащих между типичными интервалами соседних термов, то можно уточнить вид ФП.

Рассмотрим типичные интервалы соседних термов X_l и X_{l+1}, $l=\overline{1,m-1}$ – (x_l^1,x_l^2) и (x_{l+1}^1,x_{l+1}^2). Если предположить, что правая граница ФП нечеткого числа, соответствующего терму X_l, и левая граница ФП нечеткого числа, соответствующего терму X_{l+1}, являются линейными функциями, то в явном аналитическом виде они выглядят следующим образом:

$$y=1-\left(x-x_l^2\right)/\left(x_{l+1}^1-x_l^2\right);\ y=1+\left(x-x_{l+1}^1\right)/\left(x_{l+1}^1-x_l^2\right).\qquad(6)$$

Пусть $b\in(x_l^2,x_{l+1}^1)$. Если эксперт считает, что $\mu_{l+1}(b)=1+\left(b-x_{l+1}^1\right)/\left(x_{l+1}^1-x_l^2\right)$, то границы выбираются линейными функциями. Если эксперт считает, что $\mu_{l+1}(b)>1+\left(b-x_{l+1}^1\right)/\left(x_{l+1}^1-x_l^2\right)$, то границы выбираются по типу функций, изображенных на рис. 1, а. Если эксперт считает, что $\mu_{l+1}(b)<1+\left(b-x_{l+1}^1\right)/\left(x_{l+1}^1-x_l^2\right)$, то границы выбираются по типу функций, изображенных на рис. 1, б.

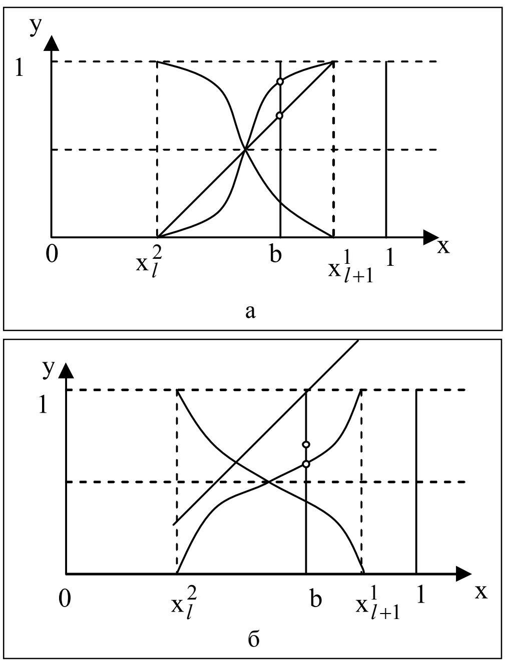

Рис. 1. Границы функций принадлежности соседних термов СП

3. Построение функций принадлежности на основе опроса эксперта о разбиении универсального множества признака

Пусть эксперт может разбить универсальное множество на непересекающиеся интервалы, каждый из которых отображается на один из термов. Будем считать, что нечеткие числа, соответствующие термам СП, являются Λ-числами. Пусть $\mu_l(x)$, $l=\overline{1,m}$ – ФП термов X_l, $l=\overline{1,m}$. Обозначим длину интервала, соответствующего терму X_l, $l=\overline{1,m}$, через c_l, $l=\overline{1,m}$, $\sum_{l=1}^{m}c_l=b-a$, $U=[a,b]$. Построим ФП $\mu_l(x)$, $l=\overline{1,m}$ в виде криволинейных трапеций со средними линиями, равными c_l, $l=\overline{1,m}$:

1. Если $c_m\le c_{m-1}$, то

$$\mu_m(x)=\begin{cases}0, & a\le x\le b-3c_m/2;\\ L\left(\left(b-c_m/2-x\right)/c_m\right), & b-{}^{3c_m}\!/_2<x\le\\ & \le b-c_m/2;\\ 1, & b-c_m/2<x\le b.\end{cases}\qquad(7)$$

2. Если $c_m>c_{m-1}$, то

$$\mu_m(x)=\begin{cases}0, & a\le x\le b-c_m-c_{m-1}/2;\\ L\left(\left(b-c_m+c_{m-1}/2-x\right)/c_{m-1}\right), & \\ & b-c_m-c_{m-1}/2<x\le b-c_m+c_{m-1}/2;\\ 1, & 1-c_m+c_{m-1}/2<x\le b.\end{cases}\qquad(8)$$

Построим ФП терма X_{m-1}:

1. Если $c_{m-1}\ge\max(c_m,c_{m-2})$, то

$$\mu_{m-1}(x)=\begin{cases}0, & 0\le x\le\sum_{l=1}^{m-3}c_l+{}^{c_{m-2}}\!/_2;\\ R\left(\left(\sum_{l=1}^{m-2}c_l+{}^{c_{m-2}}\!/_2-x\right)\Big/c_{m-1}\right), & \sum_{l=1}^{m-3}c_l+{}^{c_{m-2}}\!/_2<\\ & <x\le\sum_{l=1}^{m-2}c_l+\dfrac{c_{m-2}}{2};\\ 1, & \sum_{l=1}^{m-2}c_l+{}^{c_{m-2}}\!/_2<x\le b-{}^{3c_m}\!/_2;\\ L\left(\left(x-b+\dfrac{3c_m}{2}\right)\Big/c_m\right), & 1-{}^{3c_m}\!/_2<x\le b-{}^{c_m}\!/_2;\\ 0, & b-{}^{c_m}\!/_2<x\le b.\end{cases}\qquad(9)$$

2. Если $c_m<c_{m-1}<c_{m-2}$, то

$$\mu_{m-1}(x)=\begin{cases}0, & a\le x\le 1-c_m-3c_{m-1}/2;\\ R\left(\left(b-c_m-c_{m-1}/2\right)/c_{m-1}\right), & b-c_m-3c_{m-1}/2<\\ & <x\le b-c_m-c_{m-1}/2;\\ 1, & b-c_m-c_{m-1}/2<x\le b-3c_m/2;\\ L\left(\left(x-b+3c_m/2\right)/c_m\right), & b-3c_m/2<x\le b-c_m/2;\\ 0, & b-c_m/2<x\le b.\end{cases}\qquad(10)$$

3. Если $c_{m-2} < c_{m-1} < c_m$, то

$$\mu_{m-1}(x)=\begin{cases} 0, & 0 \le x \le \sum\limits_{l=1}^{m-3} c_l + c_{m-2}/2; \\ R\left(\left(\sum\limits_{l=1}^{m-2} c_l + c_{m-2}/2 - x\right)\Big/ c_{m-2}\right), & \sum\limits_{l=1}^{m-3} c_l + c_{m-2}/2 < \\ & < x \le \sum\limits_{l=1}^{m-2} c_l + c_{m-2}/2; \\ 1, & \sum\limits_{l=1}^{m-2} c_l + c_{m-1}/2 < x \le b - c_m - c_{m-1}/2; \\ L((x-b+c_m+c_{m-1}/2)/c_{m-1}), & b-c_m-c_{m-1}/2 < \\ & < x \le b - c_m + c_{m-1}/2 \\ 0, & b-c_m+c_{m-1}/2 < x \le b. \end{cases} \quad (11)$$

4. Если $c_{m-1} \le \min(c_m, c_{m-2})$, то

$$\mu_{m-1}(x)=\begin{cases} 0, & 0 \le x \le b - c_m - 3c_{m-1}/2; \\ R((-c_m - c_{m-1}/2 - x)/c_{m-1}), & b - c_m - 3c_{m-1}/2 < \\ & < x \le b - c_m - c_{m-1}/2; \\ L((x-b+c_m+c_{m-1}/2)/c_{m-1}), & b - c_m - c_{m-1}/2 < \\ & < x \le b - c_m + c_{m-1}/2; \\ 0, & b - c_m + c_{m-1}/2 < x \le b. \end{cases} \quad (12)$$

Аналогично $\mu_{m-1}(x)$ строятся ФП $\mu_l(x)$, $l = \overline{2, m-2}$.

Построим ФП для терма X_1 при условии четного числа термов:

1. Если $c_1 \le c_2$, то

$$\mu_1(x)=\begin{cases} 1, & a \le x \le c_1/2; \\ L((x - c_1/2)/c_1), & c_1/2 < x \le 3c_1/2; \\ 0, & 3c_1/2 < x \le b. \end{cases} \quad (13)$$

2. Если $c_1 > c_2$, то

$$\mu_1(x)=\begin{cases} 1, & 0 \le x \le c_1 - c_2/2; \\ L((x - c_1 + c_2/2)/c_2), & c_1 - c_2/2 < x \le c_1 + c_2/2; \\ 0, & c_1 + c_2/2 < x \le b. \end{cases} \quad (14)$$

Если число термов нечетное, то получим:

1. Если $c_1 \le c_2$, то

$$\mu_1(x)=\begin{cases} 1, & a \le x \le c_1/2; \\ R((x - c_1/2)/c_1), & c_1/2 < x \le 3c_1/2; \\ 0, & 3c_1/2 < x \le b. \end{cases} \quad (15)$$

2. Если $c_1 > c_2$, то

$$\mu_1(x)=\begin{cases} 1, & 0 \le x \le c_1 - c_2/2; \\ R((x - c_1 + c_2/2)/c_2), & c_1 - c_2/2 < x \le c_1 + c_2/2; \\ 0, & c_1 + c_2/2 < x \le b. \end{cases} \quad (16)$$

Если возможно получение дополнительной информации относительно значений ФП в точках универсального множества, лежащих между типичными интервалами соседних термов, то можно уточнить вид ФП.

Выводы

Разработаны методы формализации нечеткой информации на основе семантических пространств путем прямого экспертного опроса. Эти методы могут применяться для формализации информации в рамках как качественных, так и количественных признаков. Первый метод опирается на опрос единственного эксперта о типичных представителях лингвистических значений признака. Второй метод опирается на опрос единственного эксперта о разбиении универсального множества признака на непересекающиеся интервалы, каждый из которых отображается на один из термов.

Список литературы

1. *Андрейчиков А.В., Андрейчикова О.Н. Интеллектуальные информационные системы. – М.: Финансы и статистика, 2004. – 424 с.*

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. *Прикладные нечеткие системы / Под ред. Т. Тэрано, К. Асни, М. Сугэно. Пер. с япон. – М.: Мир, 1993. – 368 с.*

4. *Рыжов А.П. Элементы теории нечетких множеств и измерения нечеткости. – М.: Диалог-МГУ, 1998. – 116 с.*

5. *Нечеткие множества в моделях управления и искусственного интеллекта / А.Н. Аверкин, И.З. Батыршин, А.Ф. Блишун и др. – М.: Наука. Гл. ред. физ-мат. лит., 1986. – 312 с.*

6. *Ежкова И.В. Семантически – инвариантная формализация лингвистических оценок // В кн. Семиотические аспекты формализации интеллектуальной деятельности. – М.: МДНТП, 1983. – С. 48-51.*

7. *Полещук О.М. Методы представления экспертной информации в виде совокупности терм-множеств полных ортогональных семантических пространств // Вестник МГУЛ. – 2002. – № 5 (25). – С. 198-216.*

8. *Полещук О.М. О развитии систем обработки нечеткой информации на базе полных ортогональных семантических пространств // Вестник МГУЛ. – 2003. – № 1 (26). – С. 112-117.*

9. *Рыжов А.П. Оценка степени нечеткости и ее применение в системах искусственного интеллекта // Интеллектуальные системы. – 1996. – Т. 1. – Вып. 1 – 4. – С. 95-102.*

10. *Рыжов А.П. Об агрегировании информации в нечетких иерархических системах // Интеллектуальные системы. – 2001. – Т. 6. – Вып. 1 – 4. – С. 341-364.*

Поступила в редколлегию 9.11.2007

Рецензент: д-р техн. наук, проф. Е.И. Кучеренко, Харьковский национальный университет радиоэлектроники, Харьков.