Определение аналога показателя надежности для моделей экспертного оценивания характеристик

О. М. Полещук[1], Е. Г. Комаров[2]
Московский государственный технический университет
им. Н.Э. Баумана, Мытищинский филиал
[1]olga.m.pol@yandex.ru, [2]komarov@mgul.ac.ru

Ashraf Darwish
Helwan University Cairo, Egypt
ashraf darwish.eg@ieee.org

Аннотация. **Разработан метод построения моделей экспертного оценивания характеристик в виде полных ортогональных семантических пространств. Функции принадлежности соседних термов имеют области неопределенности, на основе которых определяются аналоги ошибок первого и второго рода и аналог показателя надежности. Определенный показатель надежности связан обратно пропорционально со степенью нечеткости моделей экспертного оценивания характеристик.**

Ключевые слова: модель экспертного оценивания; полное ортогональное семантическое пространство; степень нечеткости; аналог показателя надежности

I. Введение

Нечеткой переменной называется тройка $\left(X,U,\tilde{A}\right)$, где X – название переменной; U – область ее определения (универсальное множество); $\tilde{A}$ – нечеткое множество универсального множества, описывающее возможные значения нечеткой переменной [1].

На основе понятия нечеткой переменной вводится понятие лингвистической переменной.

Лингвистической переменной называется пятерка $\left\{X,T\left(X\right),U,V,S\right\}$, где X – название переменной; $T\left(X\right)$ – термы переменной X, то есть множество названий значений переменной X. Каждое из этих значений – нечеткая переменная со значением из универсального множества U. V – синтаксическое правило, порождающее названия значений лингвистической переменной X. S – семантическое правило, которое ставит в соответствие каждой нечеткой переменной с названием из $T\left(X\right)$ нечеткое подмножество множества U.

Семантическим пространством называется лингвистическая переменная с фиксированным терм-множеством $\left\{X,T\left(X\right),U,S\right\}$.

С позиции аппарата теории нечетких множеств моделями экспертного оценивания характеристик служат семантические пространства, имеющие широкий спектр практических применений: экспертные системы, интеллектуальные системы поддержки принятия решений, анализ данных и управление сложными процессами [2].

Проведенные теоретические исследования свойств семантических пространств, направленные на повышение адекватности моделей экспертного оценивания характеристик и их полезности для решения практических задач, позволили обоснованно сформулировать требования к функциям принадлежности $\mu_l\left(x\right), l=\overline{1,m}$ их терм-множеств [2].

1. Для каждого понятия $X_l, l=\overline{1,m}$ существует $\widehat{U}_l \neq \varnothing$ где $\widehat{U}_l = \left\{x \in U : \mu_l\left(x\right)=1\right\}$ есть точка или отрезок.

2. $\mu_l\left(x\right), l=\overline{1,m}$ не убывает слева от $\widehat{U}_l$ и не возрастает справа от $\widehat{U}_l$.

3. $\mu_l\left(x\right), l=\overline{1,m}$ имеют не более двух точек разрыва первого рода.

4. Для каждого $x \in U$ $\sum_{l=1}^{m}\mu_l(x)=1$.

Семантические пространства, функции принадлежности которых удовлетворяют сформулированным требованиям, получили название полных ортогональных семантических пространств.

Первое требование означает, что у каждого понятия (терма) существует хотя бы один эталон – типичный представитель (степень оттеночной уверенности эксперта в принадлежности типичного представителя соответствующему понятию равна единице). Если эталонов несколько, то они все расположены рядом, а не разбросаны по универсальному множеству.

Второе требование означает, что если объекты «близки» в смысле метрики в универсальном множестве, то они также «близки» в смысле принадлежности к некоторому понятию.

Третье требование обеспечивает возможность обработки нечеткой и четкой информации одновременно с единых позиций, поскольку в качестве функций принадлежности могут использоваться обычные характеристические функции.

Четвертое требование обеспечивает для каждого объекта из универсального множества существование хотя бы одного понятия, которое описывает этот объект с ненулевой степенью принадлежности. Кроме этого, оно обеспечивает разделимость понятий, образующих семантическое пространство, и исключает использование синонимии или семантически близких терминов.

В качестве моделей экспертного оценивания характеристик в статье выбираются полные ортогональные семантические пространства.

II. Построение моделей экспертного оценивания характеристик

В настоящей работе будет рассматриваться задача построения моделей экспертного оценивания качественных характеристик на основе информации группы экспертов. Будем предполагать, что для оценки некоторой характеристики X используется шкала с уровнями (термами) $X_l, l=\overline{1,m}$ на отрезке [0,1]. Ноль соответствует полному отсутствию проявления характеристики X, единичная точка соответствует полному присутствию проявления характеристики X.

Каждому из k экспертов предлагается на отрезке $[0,1]$ отметить интервалы, которые они считают типичными для каждого из термов $X_l, l=\overline{1,m}$. Предположим, что i -ым экспертом определены типичные для термов $X_l, l=\overline{1,m}$ интервалы $\left(x_{li}^1, x_{li}^2\right), l=\overline{1,m}, i=\overline{1,k}$, то есть интервалы, для которых степень уверенности эксперта в их принадлежности к соответствующим термам равна единице (и для которых функции принадлежности соответствующих термов $\mu_{li}(x), l=\overline{1,m}, i=\overline{1,k}$ равны единице). Для одних термов типичными могут быть интервалы, для других термов типичными могут быть точки (по одной для каждого терма). Тогда

$$\mu_{1i}(x) \equiv \left(0, x_{1i}^2, 0, x_{2i}^1 - x_{1i}^2\right),$$

$$\mu_{li}(x) \equiv \left(x_{li}^1, x_{li}^2, x_{li}^1 - x_{l-1i}^2, x_{l+1i}^1 - x_{l-1i}^2\right), l=\overline{2,m-1},$$

$$\mu_{mi}(x) \equiv \left(x_{mi}^1, 1, x_{mi}^1 - x_{m-1i}^2, 0\right), i=\overline{1,k}.$$

Первыми двумя параметрами функций принадлежности являются абсциссы вершин верхних оснований трапеций, которые являются графиками функций, а последние два параметра являются длинами соответственно левого и правого крыльев трапеций [3].

Пусть $X_i = \left\{\mu_{li}, l=\overline{1,m}\right\}, i=\overline{1,k}$ – модели экспертных оценок характеристики X, $\mu_{li}(x) \equiv \left(a_1^{li}, a_2^{li}, a_L^{li}, a_R^{li}\right), l=\overline{1,m}, i=\overline{1,k}$. Обозначим за $X = \left\{f_l(x), l=\overline{1,m}\right\}$, $f_l(x) \equiv \left(a_1^l, a_2^l, a_L^l, a_R^l\right), l=\overline{1,m}$ - модель групповой экспертной оценки характеристики X.

Определим потерю информации между групповой экспертной оценкой $X = \left\{f_l(x), l=\overline{1,m}\right\}$ и оценкой i -го эксперта $X_i = \left\{\mu_{li}, l=\overline{1,m}\right\}, i=\overline{1,k}$:

$$d\left(X_i, X\right) = \frac{1}{2}\sum_{l=1}^{m}\int_0^1 \left|\mu_{li}(x) - f_l(x)\right| dx.$$

Потерей информации при построении групповой экспертной оценки назовем среднее значение потерь информации между всеми экспертными оценками и групповой экспертной оценкой [4]:

$$\sigma = \frac{1}{k}\sum_{i=1}^{k} d\left(X_i, X\right), i=\overline{1,k}.$$

Введем новые параметры функций принадлежности индивидуальных и групповой экспертных оценок, которые являются абсциссами точек изломов их графиков:

$$a_{li1} = a_1^{li} - a_L^{li}, a_{li2} = a_1^{li}, a_{li3} = a_2^{li}, a_{li4} = a_2^{li} + a_R^{li},$$

$$a_{l1} = a_1^l - a_L^l, a_{l2} = a_1^l, a_{l3} = a_2^l, a_{l4} = a_2^l + a_R^l.$$

Так как $a_{1i1} = 0, a_{1i2} = 0, i=\overline{1,k}$, то полагается $a_{11} = 0, a_{12} = 0$. Так как $a_{mi3} = 1, a_{mi4} = 1, i=\overline{1,k}$, то полагается $a_{m3} = 1, a_{m4} = 1$.

Потерей информации в рамках границ l -го и $l+1$ -го термов между групповой экспертной оценкой и оценкой i -го эксперта будем называть полусумму интеграла от модуля разности соответствующих правых границ функций принадлежности l-го терма групповой экспертной оценки и оценки i -го эксперта и интеграла от модуля разности соответствующих левых границ функций принадлежности $l+1$ -го терма групповой экспертной оценки и оценки i -го эксперта.

Рассмотрим различные случаи расположения границ функций принадлежности соседних термов оценки i -го эксперта и границ функций принадлежности этих же термов групповой оценки. В зависимости от расположения границ функций принадлежности найдем потери информации.

Если $a_{l3} > a_{li3}, a_{l4} > a_{li4}$, то потеря информации в рамках границ l-го и $l+1$-го термов равна площади трапеции с основаниями $a_{l3} - a_{li3}, a_{l4} - a_{li4}$ и единичной высотой, то есть $\frac{1}{2}\left(a_{l3} - a_{li3} + a_{l4} - a_{li4}\right)$.

Если $a_{l3} < a_{li3}, a_{l4} < a_{li4}$, то потеря информации в рамках границ l -го и $l+1$ -го термов равна площади трапеции с основаниями $a_{li3} - a_{l3}, a_{li4} - a_{l4}$ и единичной высотой, то есть $\frac{1}{2}\left(-a_{l3} + a_{li3} - a_{l4} + a_{li4}\right)$.

Если $a_{l3} \le a_{li3}, a_{l4} \ge a_{li4}$, то потеря информации в рамках границ l-го и $l+1$-го термов равна сумме площадей двух треугольников. Один треугольник имеет основание $-a_{l3} + a_{li3}$, другой треугольник тоже имеет основание $a_{l4} - a_{li4}$. Найдем высоты этих треугольников. Так как треугольники с основаниями $-a_{l3} + a_{li3}$ и $a_{l4} - a_{li4}$ подобны, то имеем

$$\begin{cases} \dfrac{h_1}{h_2} = \dfrac{a_{li3} - a_{l3}}{a_{l4} - a_{li4}}, \\ h_1 + h_2 = 1 \end{cases}$$

где h_1 и h_2 – высоты соответствующих треугольников.

Откуда получаем, что высота треугольников с основанием $-a_{l3} + a_{li3}$ равна

$$h_1 = \frac{a_{li3}a_{l4} - a_{li3}a_{l3} - a_{l3}a_{li4} + a_{l3}a_{li3} - a_{l3}a_{l4}}{(a_{l4} - a_{l3})(a_{l4} - a_{l3} - a_{li4} + a_{li3})} + \\ + \frac{a_{l3}^2 + a_{l3}a_{li4} - a_{l3}a_{li3}}{(a_{l4} - a_{l3})(a_{l4} - a_{l3} - a_{li4} + a_{li3})} = \\ = \frac{a_{il3} - a_{l3}}{a_{l4} - a_{l3} - a_{il4} + a_{il3}}.$$

Высота треугольников с основанием $a_{l4} - a_{li4}$ равна

$$h_2 = 1 - \frac{a_{li3} - a_{l3}}{a_{l4} - a_{l3} - a_{li4} + a_{li3}} = \frac{a_{l4} - a_{li4}}{a_{l4} - a_{l3} - a_{li4} + a_{li3}}.$$

В этом случае потеря информации равна

$$\frac{(a_{l3} - a_{li3})^2 + (a_{l4} - a_{li4})^2}{2(a_{l4} - a_{li4} + a_{li3} - a_{l3})}.$$

Если $a_{l3} > a_{li3}, a_{l4} < a_{li4}$, то потеря информации равна

$$-\frac{(a_{l3} - a_{li3})^2 + (a_{l4} - a_{li4})^2}{2(a_{l4} - a_{li4} + a_{li3} - a_{l3})}.$$

Таким образом, общая потеря информации равна

$$\sigma = \frac{1}{k}\sum_{l=1}^{m-1}\sum_{i=1}^{k}\left[\frac{1}{2}\delta_{li}^1 (a_{l3} - a_{li3} + a_{l4} - a_{li4})\right] + \\ + \frac{1}{k}\sum_{l=1}^{m-1}\sum_{i=1}^{k}\left[\delta_{li}^2\left(\frac{(a_{l3} - a_{li3})^2 + (a_{l4} - a_{li4})^2}{2(a_{l4} - a_{li4} + a_{li3} - a_{l3})}\right)\right].$$

$$\delta_{li}^1 = \begin{cases} 1, & a_{l3} \ge a_{li3},\ a_{l4} \ge a_{li4} \\ -1, & a_{l3} \le a_{li3},\ a_{l4} \le a_{li4} \\ 0, & a_{l3} > a_{li3}, a_{l4} < a_{li4}\ \text{или}\ a_{l3} < a_{li3}, a_{l4} > a_{li4} \end{cases};$$

$$\delta_{li}^2 = \begin{cases} 1, & a_{l3} < a_{li3}, a_{l4} > a_{li4} \\ -1, & a_{l3} > a_{li3}, a_{l4} < a_{li4} \\ 0, & a_{l3} \ge a_{li3}, a_{l4} \ge a_{li4}\ \text{или}\ a_{l3} \le a_{li3}, a_{l4} \le a_{li4} \end{cases}.$$

Неизвестные параметры $a_{l3}, a_{l4}, l = \overline{1, m-1}$ являются решениями оптимизационной задачи

$$\sigma = \frac{1}{2k}\sum_{l=1}^{m-1}\sum_{i=1}^{k}\left[\delta_{li}^1 (a_{l3} - a_{li3} + a_{l4} - a_{li4})\right] + \\ + \frac{1}{2k}\sum_{l=1}^{m-1}\sum_{i=1}^{k}\left[\delta_{il}^2\left(\frac{(a_{l3} - a_{li3})^2 + (a_{l4} - a_{li4})^2}{a_{l4} - a_{li4} + a_{li3} - a_{l3}}\right)\right] \to min.$$

Решения находятся в рамках известных методов [5].

Построенная модель групповой экспертной оценки качественной характеристики сохраняет максимум информации, заложенной в индивидуальных экспертных оценках. Согласованность индивидуальных экспертных оценок предлагается проверять, используя показатель, приведенный ниже [2]

$$\kappa = \frac{1}{m}\sum_{l=1}^{m}\frac{\int_0^1 min(\mu_{li}(x), \mu_{l2}(x), ..., \mu_{lk}(x))dx}{\int_0^1 max(\mu_{l1}(x), \mu_{l2}(x), ..., \mu_{lk}(x))dx}.$$

Этот показатель меняется от нуля до единицы, и чем он ближе к единице, тем более согласованы индивидуальные экспертные оценки.

III. Определение аналога показателя надежности моделей экспертного оценивания характеристик

Пусть $X = \{\mu_l(x), l = \overline{1, m}\}$ – модель группового экспертного оценивания качественной характеристики X, $\mu_l(x) \equiv (a_1^l, a_2^l, a_L^l, a_R^l)$. Для функций принадлежности $\mu_l(x), \mu_{l+1}(x), l = \overline{1, m-1}$ соседних термов модели существуют зоны неопределенности ($x \in U : 0 < \mu_l(x) < 1, 0 < \mu_{l+1}(x) < 1$), мощность которых равна $a_R^l, l = \overline{1, m-1}$. Опираясь на свойства функций принадлежности полных ортогональных семантических пространств и понятие геометрической вероятности, аналогами вероятностей ошибок первого и второго рода для терма X_l (соответственно аналогами вероятностей ошибок первого и второго рода для терма X_{l+1}) можно считать $\alpha_l = \beta_l = \dfrac{a_R^l}{2}$.

Тогда аналогами вероятностей ошибок первого и второго рода моделей экспертного оценивания характеристик будем называть

$$\alpha = \beta = \{\alpha_l\}_{l=\overline{1, m-1}},$$

а аналогом показателя надежности

$$V = (1 - \alpha)^2.$$

Таким образом, аналогом показателя надежности модели экспертного оценивания характеристики $X=\left\{\mu_l(x), l=\overline{1,m}\right\}$, $\mu_l(x)\equiv\left(a_1^l,a_2^l,a_L^l,a_R^l\right)$ называется

$$V=\left(1-\max_{l=\overline{1,m-1}}\left\{\frac{a_R^l}{2}\right\}\right)^2.$$

В [6] дано определение степени нечеткости полного ортогонального семантического пространства $\zeta=\frac{1}{|U|}\int_U f(\mu_{l_1}(x)-\mu_{l_2}(x))dx$,

где $\mu_{l_1}(x)=\max_{1\le l\le m}\mu_l(x)$, $\mu_{l_2}(x)=\max_{\substack{1\le l\le m\\ l\ne l_1}}\mu_l(x)$, f – убывает и $f(0)=1, f(1)=0$.

Если $f(x)=1-x$, то

$\zeta=\frac{1}{|U|}\int_U\left(1-\left(\mu_{l_1}(x)-\mu_{l_2}(x)\right)\right)dx=\frac{|\overline{U}|}{2|U|}$, где $\overline{U}=U-\bigcup_{l=1}^{m}\widehat{U}_l$.

Аналог показателя надежности связан со степенью нечеткости следующим образом (при $f(x)=1-x$): если степень нечеткости модели минимальна, то есть равна нулю, то аналог показателя надежности максимален и равен единице; если степень нечеткости максимальна, то есть равна 0.5, то аналог показателя надежности минимален и равен 0.25. Определенный показатель позволяет существенно расширить информацию, полученную на основе степени нечеткости, так как при одинаковой степени нечеткости модели экспертного оценивания характеристик могут иметь разные значения аналога показателя надежности. Это следует из определений степени нечеткости и аналога показателя надежности. При $f(x)=1-x$ степень нечеткости зависит от суммы длин интервалов, которые являются зонами неопределенности и не зависит от того, равные по длине эти интервалы или один из интервалов существенно превосходит по длине остальные. При определении аналога показателя надежности учитывается максимальная зона неопределенности, что позволяет получить не усредненную, а реальную картину нечеткости экспертных критериев при использовании их в задачах принятия решений.

IV. Заключение

В статье определен аналог показателя надежности для моделей группового экспертного оценивания качественных характеристик. Модели экспертного оценивания являются полными ортогональными семантическими пространствами, для которых ранее было введено понятие степени нечеткости. Степень нечеткости модели являлась показателем ее качества и позволяла оценить степень трудности, которую испытывают эксперты при оценивании характеристик в условиях неопределенности. Аналог показателя надежности модели экспертного оценивания расширяет информацию об ее качестве и позволяет выбрать лучшую модель при равных степенях нечеткости.

Список литературы

[1] Zadeh L.A.The Concept of a linguistic variable and its application to approximate reasoning // Information Sciences. 1975. Vol. 8. Pp. 199-249.

[2] Poleshchuk O., Komarov E. Expert Fuzzy Information Processing, Studies in Fuzziness and Soft Computing, 2011. T. 268. C. 1-239.

[3] Полещук О.М. Математическая модель обработки экспертных оценок // Вестник Московского государственного университета леса – Лесной вестник. 2005. № 6 (42). С. 161-164.

[4] Poleshchuk O.M., Komarov E.G., Darwish A. Comparative analysis of expert criteria on the basis of complete orthogonal semantic spaces // Proceedings of the 19th International Conference on Soft Computing and measurements (SCM) . 2016. Pp. 369-373.

[5] Coleman T.F., Li Y. A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables // SIAM J. Optim. 1996. Vol. 6. № 4. Pp. 1040-1058.

[6] A. Ryjov, A. Belenki, R. Hooper, V. Pouchkarev, A. Fattah, L. Zadeh, Development of Intelligent for Monitoring and Evaluation of Peaceful Nuclear Activities // Proceedings of the IAEA. 1998. P. 122.