УДК 517.9

АГРЕГИРОВАНИЕ ЛИНГВИСТИЧЕСКОЙ ИНФОРМАЦИИ

Т.Н. Недикова, И.С. Халяпина

В статье представлены результаты исследования свойств операторов агрегирования лингвистической информации, позволяющих формировать обобщенные оценки многомерных альтернатив

Ключевые слова: операторы агрегирования, лингвистическая переменная

Обработка нечеткой информации является одним из направлений современной информационной технологии и служит средством для разработки интеллектуальных информационных систем и в том числе систем принятия решений. Одним их основных компонентов таких систем является подсистема оценки альтернатив – допустимых вариантов решений, объектов. В условиях многокритериальности оценочная модель позволяет получить обобщенную оценку, позволяющую упорядочить альтернативы по предпочтению, разбить их на упорядоченные по качеству группы или выбрать наилучшую. В случае количественных оценок альтернатив по критериям для формирования обобщенной оценки используются различные свертки с соответствующим аксиоматическим обоснованием. Однако большинство ситуаций характеризуются неопределенностью, имеющей характер нечеткости, неточности, что делает получение количественных оценок практически невозможным, с другой стороны существуют показатели, которые являются качественными по своей природе. В этих случаях актуально использование моделей приближенной оценки, информационной основой которых являются экспертные знания. Формализация нечетких словесных понятий, которыми оперирует эксперт, обеспечивается введением нечеткой и лингвистической переменной. Нечеткая переменная представляет собой нечеткое подмножество некоторого универсального множества, которое описывает ограничение на возможные значения данной переменной. Лингвистическая переменная определяется через три компоненты – множество ее значений, представляющих собой наименования нечетких переменных, областью определения каждой из которых является одно и то же универсальное множество; синтаксическую процедуру, предназначенную для образования новых значений и семантическую процедуру для отображения нового лингвистического значения в нечеткую переменную.

В данной статье представлены результаты исследований свойств операторов агрегирования лингвистической информации в приложении к построению моделей приближенной оценки альтернатив.

Рассмотрим лингвистическую шкалу S, которая представляет собой конечное вполне упорядоченное множество термов $\{S_i\}$, $i \in \{0,..., T\}$, удовлетворяющих условиям: 1) если $i < j$, то S_i предшествует S_j; 2) отрицание терма определяется $N(S_i) = S_{T-i+1}$; 3) дизъюнкция термов определяется $S_i \vee S_j = S_k$, где $k = \max(i, j)$; 4) конъюнкция термов определяется $S_i \wedge S_j = S_k$, где $k = \min(i, j)$.

Рассмотрим задачу: пусть задано множество альтернатив $X = \{x\}$, для описания которых используется набор критериев $K = \{K_1,\dots, K_n\}$. Каждая альтернатива x оценивается по каждому из критериев K_i в некоторой лингвистической шкале S, так что в результате этого оценивания каждой альтернативе $x \in X$ ставится в соответствие векторная оценка $A(x) = (a_1(x),\dots, a_n(x)) = (a_1,\dots, a_n) = A$, где $a_i = a_i(x)$ – оценка альтернативы x по критерию K_i. Весовые коэффициенты критериев w_i $(i = 1,\dots,n)$ могут быть как числовыми, так и лингвистическими из шкалы S. Необходимо на основе векторных оценок альтернатив выбрать лучшую альтернативу. В основу решения данной задачи положим подход, основанный на построении обобщенных оценок альтернатив с помощью операторов агрегирования.

Наиболее общий подход к агрегированию лингвистической информации заключается в аксиоматическом определении оператора агрегирования, при этом можно выделить три основные стратегии свертки индивидуальных экспертных оценок: конъюнктивную, дизъюнктивную и компромиссную стратегии [1].

Рассмотрим оператор агрегирования общего вида $\alpha(x) = Agg(W, A)$, где x – альтернатива из заданного множества, $\alpha(x)$ – ее обобщенная оценка, $A = (a_1,\dots, a_n)$ – вектор частных оценок альтернативы, $W = (w_1,\dots, w_n)$ – вектор весовых коэффициентов. Для случая, когда частные оценки a_i принимают свои значения в некоторой лингвистической шкале S в [2] предложено семейство лингвистических операторов, основанных на OWA-операторе и понятии выпуклой комбинации лингвистических термов. N – местный OWA-оператор, ассоциированный с вектором весов $W = (w_1,\dots, w_n)$, таких что $\forall\ i\ (w_i > 0)$ и $\sum_{i=1}^{n} w_i = 1$, есть отображение F: $[0,1]^{n+1} \to [0,1]$, при-

---

Недикова Татьяна Николаевна – ВГТУ, канд. техн. наук,
e-mail: admin.voronezh@vsk.ru
Халяпина Ирина Сергеевна – ВГУ, магистр,
e-mail: xalyapina@mail.ru

чем $F(W,A)=\sum_{i=1}^{n} w_i b_i$, где $B=(b_1,\ldots,b_n)$ – вектор, полученный из $A=(a_1,\ldots,a_n)$ упорядочением элементов по невозрастанию.

К основным свойствам OWA-оператора относятся [2]: коммутативность, монотонность, идемпотентность.

Заметим, что различные OWA-операторы отличаются друг от друга наборами весов, при этом можно выделить три важнейших частных случая:

$F(W',A)=\max\{a_1,\ldots,a_n\}$ для $W'=(1,0,\ldots,0)$,

$F(W'',A)=\min\{a_1,\ldots,a_n\}$ для $W'=(0,\ldots,0,1)$,

$$F(W''',A)=\frac{1}{n}\sum_{i=1}^{n} a_i \text{ для } W'''=\left(\frac{1}{n},\ldots,\frac{1}{n}\right).$$

Анализ этих простейших операторов позволяет установить, что $F(W',A)$ реализует дизъюнктивное агрегирование, $F(W'',A)$ – конъюнктивное, а $F(W''',A)$ – компромиссное. Для произвольного OWA-оператора справедливо неравенство

$$F(W',A)\le F(W,A)\le F(W'',A),$$

которое означает, что $F(W,A)$ в общем случае является оператором осреднения.

Для классификации OWA-операторов по отношению к связкам и и или вводятся специальные величины [2]: orness $(W)=\frac{1}{n-1}\sum_{i=1}^{n}(n-i)w_i$ характеризует близость к дизъюнкции, и andness $(W)=1$-orness (W) – к конъюнкции.

Из определения следует, что orness $(W)\in[0,1]$, при этом orness $(W')=1$, orness $(W'')=0$, orness $(W''')=0{,}5$. Для произвольного OWA-оператора примем следующее определение: если orness $(W)>0{,}5$, то соответствующий оператор будем называть квазидизъюнкцией, иначе квазиконъюнкцией. Данные операторы актуальны для тех случаев, когда эксперт затрудняется с полной уверенностью идентифицировать тип операции, лежащей в основе формирования обобщенной оценки.

Некоторые оценочные модели ориентированы на наличие компенсационных свойств оператора агрегирования, когда малые значения оценок альтернативы по одному критерию компенсируются большими значениями оценок по другому (другим) критериям. Анализ семантики операторов показывает, что оператор min не обладает компенсационными свойствами, в то время как оператор max соответствует случаю полной компенсации. Учитывая введенные характеристики операторов orness (W) и andness (W), можно считать, что для произвольного оператора $F(W,A)$ чем ближе величина orness (W) к 1, тем в большей степени данный оператор обладает компенсационными свойствами.

Для OWA-оператора $F(W,A)$ можно определить двойственный оператор $\overline{F}(Q,A)$ с вектором весов $Q=(q_1,\ldots,q_n)$, где $q_i=w_{n-i+1}$, причем orness (Q) = andness (W) и наоборот. Двойственный оператор к двойственному совпадает с исходным.

Заметим, что перераспределяя веса определенным образом, можно изменять стратегию агрегирования. Справедливо следующее утверждение [2]: пусть W^1 и W^2 - векторы весов соответствующих OWA-операторов, такие что $W^1=(w_1^1,\ldots,w_n^1)$ и $W^2=(w_1^2,\ldots,w_n^2)$, причем эти векторы отличаются только двумя координатами, т.е. существуют индексы k и r ($k>r$) такие, что для заданной константы $t>0$ отличающиеся координаты определяются следующим образом: $w_r^2=w_r^1+t$, $w_k^2=w_k^1-t$, а для $i\notin\{k,r\}$ координаты совпадают, т.е. $w_i^2=w_i^1$. Тогда orness $(W^1)<$ orness (W^2).

Таким образом, изменяя вектор весов, можно увеличивать или уменьшать величину orness (W), тем самым изменяя компенсационные свойства соответствующих операторов, поэтому она важна для целенаправленного построения операторов с различной интенсивностью компенсационных свойств.

Рассмотренный выше OWA-оператор введен для случая, когда компоненты векторов W и A являются количественными из [0,1]. Предположим теперь, что $w_i\in[0,1]$, $a_i\in S$ для всех $i=1,\ldots,n$. Соответствующий этому случаю оператор агрегирования называется лингвистическим OWA-оператором (LOWA-оператор) и определяется правилом:

$$\Phi(W,A)=\Phi_W(A)=C^n\{(w_k,b_k),k=1,\ldots,n\}=$$
$$=w_1\otimes b_1\oplus(1-w_1)\otimes C^{n-1}\{(в_j,b_j),j=2,\ldots,n\},$$

где $в_j=\frac{w_j}{\sum_{h=2}^{n} w_h}$, $j=2,\ldots n$, C^n, C^{n-1} - выпуклые комбинации термов.

При $n=2$, $b_1=S_j, b_2=S_i$

$$C^2\{(w_1,b_1),(w_2,b_2)\}=w_1\otimes S_j\oplus w_2\otimes S_i=S_k,$$

где $k=\min\{T, \text{round}(w_1(j-i))\}$, round - оператор округления.

Рассмотрим свойства LOWA-оператора.

1. $\Phi_W(A)$ является оператором осреднения.

Заметим, что если k - результирующий индекс в выпуклой комбинации двух лингвистических оценок a_i и a_j, причем $i\le j$, то $i\le k\le j$. Это означает, что $a_i\le\Phi_W(a_i,a_j)\le a_j$. В случае n аргументов $\min\{a_1,\ldots,a_n\}\le\Phi_W(A)\le\max\{a_1,\ldots,a_n\}$, что характеризует LOWA-оператор как оператор осреднения.

2. LOWA - коммутативный оператор, то есть $\Phi_W(A)=\Phi_W(\pi(A))$, где $\pi(A)$ – произвольная перестановка компонент вектора A. Свойство коммутативности обеспечивается тем, что перед агрегированием осуществляется упорядочение аргументов.

3. LOWA - идемпотентный оператор.

При $n=2$ справедливость данного свойства вытекает из того факта, что при $i=j$ индекс $k=\min\{T, i+\text{round}(w_1(j-i))\}=\min\{T, i\}=i$, т.е. если $a_i=a_j=a$, то $\Phi_W(a_i,a_j)=\Phi_W(a,a)=a$. Учитывая рекурсивное определение выпуклой комбинации, можно заключить о справедливости данного свойства для n аргументов.

4. LOWA-оператор не является ассоциативным.

Рассмотрим пример. Пусть лингвистическая шкала определена в виде S = {N, VL, L, M, H, VH, P}. Зададим вектор весов W = (0,3;0,7) и проверим справедливость равенства

$\Phi_W(\Phi_W(VL,M),VH) = \Phi_W(VL,\Phi_W(M,VH))$.

Вычислим $\Phi_W(VL,M) = L = S_2$, так как j = 3, i =1, k = min{6, 1+ round(0,3×(3-1))} = min {6,2} = 2. Далее $\Phi_W(L,VH) = M = S_3$, так как j = 5, i = 2, k = min{6, 2+ round(0,3×(5-2))} = min {6,3} = 3 . Таким образом, $\Phi_W(\Phi_W(VL,M),VH) = S_3$. С другой стороны, $\Phi_W(VL,\Phi_W(M,VH)) = S_2$. Окончательно,

$\Phi_W(\Phi_W(VL,M),VH) \neq \Phi_W(VL,\Phi_W(M,VH))$.

5. $\Phi_W(A)$ монотонно возрастает по каждому аргументу. Пусть $B' = (b_1',...,b_n')$, $B'' = (b_1'',...,b_n'')$ – два упорядоченных по невозрастанию вектора лингвистических оценок для $A' = (a_1',...,a_n')$ и $A'' = (a_1'',...,a_n'')$, таких что $\forall i\,(b_i' \geq b_2'')$, тогда $\Phi_W(A') \geq \Phi_W(A'')$ при фиксированном векторе весов W.

Данное свойство доказывается индукцией по числу аргументов.

Пусть n = 2, $b_1' = S_i, b_2' = S_j, b_1'' = S_p, b_2'' = S_q$. Ясно, что $j \geq p$, $i \geq q$.

Для $w_1 \in [0,1]$, имеем $j \cdot w_1 \geq p \cdot w_1$ и $i \cdot (1-w_1) \geq q \cdot (1- w_1)$, тогда $j \cdot w_1 + i \cdot (1-w_1) \geq p \cdot w_1 + q \cdot (1- w_1)$. Округляя, получим

round $(j \cdot w_1 + i \cdot (1-w_1)) \geq$ round $(p \cdot w_1 + q \cdot (1- w_1))$,

round $(i + w_1 \cdot (j-i)) \geq$ round $(q + w_1 \cdot (p-q))$.

Так как $i \in Z^+$ и $w_1(j-i) > 0$, то

$i + \text{round}(w_1(j-i)) \geq q + \text{round}(w_1(p-q))$,

откуда $\Phi_W(a_1',a_2') \geq \Phi_W(a_1'',a_2'')$, т.е. для n = 2 свойство доказано.

Предположим, что свойство является истинным для (n-1), т.е. $\Phi_W(a_1',...,a_{n-1}') \geq \Phi_W(a_1'',...,a_{n-1}'')$, тогда для n аргументов имеем

$$\Phi_W(A') = w_1 \otimes b_1' \oplus (1-w_1) \otimes C^{n-1}\{(в_j, b_j'), j = 2,...,n\},$$

$$\Phi_W(A'') = w_1 \otimes b_1'' \oplus (1-w_1) \otimes C^{n-1}\{(в_j, b_j''), j = 2,...,n\}.$$

Так как

$$C^{n-1}\{(в_j, b_j'), j = 2,...,n\} = \Phi_W(a_2',...,a_n'),$$

$$C^{n-1}\{(в_j, b_j''), j = 2,...,n\} = \Phi_W(a_2'',...,a_n''),$$

то в соответствии с гипотезой $\Phi_W(a_2',...,a_n') \geq \Phi_W(a_2'',...,a_n'')$. Пусть $\Phi_W(a_2',...,a_n') = S_i$, $\Phi_W(a_2'',...,a_n'') = S_j$. Так как $a_1' \geq a_1''$, а $a_1' \geq S_i$ и $a_1'' \geq S_j$, то $\Phi_W(A') = \Phi_W(a_1',S_i)$, $\Phi_W(A'') = \Phi_W(a_1'',S_j)$.

Так как для n = 2 свойство справедливо, т.е. $\Phi_W(a_1',S_i) \geq \Phi_W(a_1'',S_j)$,то $\Phi_W(A') \geq \Phi_W(A'')$, что и требовалось доказать.

Таким образом, для лингвистических операторов выполняется ряд важных алгебраических свойств, которые обеспечивают их эффективное использование в задаче выбора лучшей альтернативы.

С использованием данного класса операторов был разработан программный комплекс VALUE, к основным функциональным возможностям которого относятся: формирование лингвистических шкал; формализованная процедура оценки альтернатив в этих шкалах; выбор стратегии агрегирования; формирование вектора весовых коэффициентов на основе лингвистических квантоворов, формализующих принцип «нечеткого большинства»; вычисление обобщенных оценок и ранжирование альтернатив заданного множества. Практическое значение полученных результатов заключается в возможности обрабатывать экспертную информацию, представленную в лингвистической форме. Программа VALUE может использоваться для определения рейтинга объектов произвольных множеств в технической, экономической, социальной сферах.

Литература

1. Леденева Т.М. Обработка нечеткой информации. Воронеж: Из-дво ВГУ, 2006. 240 с.

2. Herrera F., Herrera-Viedma E. Aggregation operators for linguistic weighted information // IEEE Transactions on Systems, Man, and Cybernetics, 1997, №27, P. 646-656.

Воронежский государственный технический университет
Воронежский государственный университет

AGGREGATION OF LINGUISTIC INFORMATION

T.N. Nedikova, I.S. Xalyapina

The article features the results of research into characteristics of linguistic data aggregating operators enabling to form generalized assessment of multidimensional alternatives

Key words: aggregation operators, linguistic variable