УДК 510.22

МНОГОКРИТЕРИАЛЬНОЕ ПРИНЯТИЕ РЕШЕНИЙ ПО ПРОТИВОРЕЧИВЫМ ДАННЫМ: ПОДХОД ТЕОРИИ МУЛЬТИМНОЖЕСТВ*

А.Б.Петровский
Институт системного анализа РАН, pab@isa.ru

В работе рассматриваются новые методы классификации и упорядочения многопризнаковых объектов, которые характеризуются многими разнородными (количественными и качественными) признаками и могут существовать в нескольких экземплярах с отличающимися, в частности, противоречивыми значениями признаков. Методы основаны на теории метрических пространств мультимножеств и применены для решения практических задач принятия решений: конкурсного отбора проектов и построения рейтинга компаний, оцененных несколькими экспертами по многим критериям.

1. Введение

В различных прикладных задачах многокритериального принятия решений, распознавания образов, классификации, обработки разнородной информации, теории кодирования, других предметных областях часто возникает необходимость сгруппировать или упорядочить анализируемые объекты, основываясь на их свойствах, выраженных признаками (атрибутами) объектов. Вместе с тем имеется достаточно широкий круг задач, где изучаемые объекты характеризуются многими разнородными признаками, которые могут быть и количественными, и качественными, и, кроме того, одни и те же объекты могут существовать в нескольких экземплярах с отличающимися значениями признаков, свертка которых или невозможна, или математически некорректна. В качестве примеров таких задач укажем классификацию и ранжирование объектов, оцененных несколькими экспертами по многим качественным критериям, распознавание графических символов, обработку текстовых документов. Множественность и повторяемость факторов, описывающих объекты, усложняет и затрудняет решение таких задач. Главные трудности связаны с необходимостью учитывать одновременно большое количество вербальных и числовых данных и обрабатывать эти данные, не прибегая к дополнительным преобразованиям типа усреднения, смешивания, взвешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных.

Удобной математической моделью для представления многопризнаковых объектов является мультимножество или множество с повторяющимися элементами. Кратность элементов – существенная особенность мультимножества, позволяющая отличать его от множества и рассматривать мультимножество как качественно новое математическое понятие. В работе предложены методы классификации и упорядочения совокупности многопризнаковых объектов, которые базируются на теории метрических пространств мультимножеств. Метод классификации объектов позволяет строить обобщенное решающее правило для их отбора, которое аппроксимирует различные правила экспертной сортировки объектов. Метод упорядочения объектов основан на оценке их близости по отношению к некоторому «идеальному» объекту в многопризнаковом пространстве. Эти методы допускают использование различных, в том числе и противоречивых, данных для описания объектов.

* Работа выполнена при поддержке программ фундаментальных исследований РАН «Математическое моделирование и интеллектуальные системы», «Фундаментальные основы информационных технологий и систем», Российского фонда фундаментальных исследований (проекты 02-01-01077, 04-01-00290), гранта НШ1964.2003.1 Президента Российской Федерации по поддержке ведущих научных школ.

2. Мультимножества и операции над ними

Дадим краткий обзор теории мультимножеств и метрических пространств мультимножеств [9-12]. Мультимножеством $\boldsymbol{A}$, порожденным обычным множеством $U=\{x_1,x_2,\ldots\}$, все элементы которого различны, называется совокупность групп элементов вида $\boldsymbol{A}=\{k_A(x)\bullet x|x\in U, k_A(x)\in \mathbb{Z}_+\}$. Здесь $k_A{:}U\to\mathbb{Z}_+=\{0,1,2,\ldots\}$ называется функцией числа экземпляров мультимножества, определяющей кратность вхождения элемента $x_i\in U$ в мультимножество $\boldsymbol{A}$, что обозначено символом •. Если $k_A(x)=\chi_A(x)$, где $\chi_A(x)=1$ при $x\in\boldsymbol{A}$ и $\chi_A(x)=0$ при $x\notin\boldsymbol{A}$, то мультимножество $\boldsymbol{A}$ становится обычным множеством. Если все мультимножества семейства $\mathcal{A}=\{\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots\}$ образуются из элементов множества G, то G называется доменом для семейства $\mathcal{A}$, а множество $\mathrm{Supp}\boldsymbol{A}=\{x|x\in G, \chi_{\mathrm{Supp}A}(x)=\chi_A(x)\}$ – опорным множеством или носителем мультимножества $\boldsymbol{A}$. Мощность мультимножества $|\boldsymbol{A}|=\Sigma_x k_A(x)$ определяется как общее число экземпляров всех его элементов; размерность мультимножества $/\boldsymbol{A}/=\Sigma_x\chi_A(x)=|\mathrm{Supp}\boldsymbol{A}|$ – как общее число различных элементов. Максимальное значение функции кратности $\mathrm{hgt}\boldsymbol{A}=\max\limits_{x\in G} k_A(x)$ называется высотой, а элемент $x_{A*}=\arg\max\limits_{x\in G} k_A(x)$ – пиком мультимножества $\boldsymbol{A}$. Мультимножество называется пустым $\varnothing$, если $k_\varnothing(x)=0$, максимальным $\boldsymbol{Z}$, если $k_Z(x)=\max\limits_{A\in\mathcal{A}} k_A(x)$, постоянным $\boldsymbol{C}_{[h]}$, если $n_{C[h]}(x)=h=\mathrm{const}$, $\forall x\in U$. Таким образом, пустое мультимножество $\varnothing$ есть постоянное мультимножество высоты $\mathrm{hgt}\varnothing=0$, любое обычное множество B является постоянным мультимножеством высоты $\mathrm{hgt}B=1$.

Введем следующие основные операции над мультимножествами [10-12]:

объединение $\bigcup\limits_{i\in I}\boldsymbol{A}_i=\{k_{\bigcup_i A_i}(x)\bullet x\,|\,k_{\bigcup_i A_i}(x)=\max\limits_{i\in I}k_{A_i}(x), \forall x\in G\}$;

пересечение $\bigcap\limits_{i\in I}\boldsymbol{A}_i=\{k_{\bigcap_i A_i}(x)\bullet x\,|\,k_{\bigcap_i A_i}(x)=\min\limits_{i\in I}k_{A_i}(x), \forall x\in G\}$;

арифметическое сложение $\sum\limits_{i\in I}\boldsymbol{A}_i=\{k_{\sum_i A_i}(x)\bullet x\,|\,k_{\sum_i A_i}(x)=\sum\limits_{i\in I}k_{A_i}(x), \forall x\in G\}$;

арифметическое вычитание $\boldsymbol{A}-\boldsymbol{B}=\{k_{A-B}(x)\bullet x\,|\,k_{A-B}(x)=k_A(x)-k_{A\cap B}(x)\}$;

симметрическая разность $\boldsymbol{A}\Delta\boldsymbol{B}=\{k_{A\Delta B}(x)\bullet x\,|\,k_{A\Delta B}(x)=|k_A(x)-k_B(x)|\}$;

дополнение $\overline{\boldsymbol{A}}=\boldsymbol{Z}-\boldsymbol{A}=\{k_{\overline{A}}(x)\bullet x\,|\,k_{\overline{A}}(x)=k_Z(x)-k_A(x)\}$;

умножение на число (репродукция) $h\bullet\boldsymbol{A}=\{k_{h\bullet A}(x)\bullet x\,|\,k_{h\bullet A}(x)=h\cdot k_A(x), h\in\mathbb{N}\}$;

арифметическое умножение $\prod\limits_{i\in I}\boldsymbol{A}_i=\{k_{\prod_i A_i}(x)\bullet x\,|\,k_{\prod_i A_i}(x)=\prod\limits_{i\in I}k_{A_i}(x), \forall x\in G\}$;

арифметическая n-ая степень $\boldsymbol{A}^n=\{k_{A^n}(x)\bullet x\,|\,k_{A^n}(x)=(k_A(x))^n\}$;

прямое произведение

$$\boldsymbol{A}_1\times\ldots\times\boldsymbol{A}_n=\{k_{A_1\times\ldots\times A_n}\bullet\langle x_{p_1},\ldots,x_{p_n}\rangle\,|\,k_{A_1\times\ldots\times A_n}=\prod_{i=1}^{n}k_{A_i}(x_{p_i}), x_{p_i}\in\boldsymbol{A}_i, p_i\in I_i, i=1,\ldots,n\};$$

прямая n-ая степень $(\times\boldsymbol{A})^n=\{k_{(\times A)^n}\bullet\langle x_1,\ldots,x_n\rangle\,|\,k_{(\times A)^n}=\prod\limits_{i=1}^{n}k_A(x_i), x_i\in\boldsymbol{A}\}$,

где $\langle x_1,\ldots,x_n\rangle$ – n-элементный кортеж. Заметим, что репродукцию мультимножества можно представить как сумму h мультимножеств $\boldsymbol{A}$: $h\bullet\boldsymbol{A}=\boldsymbol{A}+\ldots+\boldsymbol{A}$, или как произведение постоянного мультимножества $\boldsymbol{C}_{[h]}$ и мультимножества $\boldsymbol{A}$: $h\bullet\boldsymbol{A}=\boldsymbol{C}_{[h]}\bullet\boldsymbol{A}$.

Носители операций над мультимножествами определяются следующими выражениями:

$$\mathrm{Supp}(\bigcup_{i\in I}\boldsymbol{A}_i)=\bigcup_{i\in I}(\mathrm{Supp}\,\boldsymbol{A}_i)=\mathrm{Supp}(\sum_{i\in I}\boldsymbol{A}_i),\quad \mathrm{Supp}(\bigcap_{i\in I}\boldsymbol{A}_i)=\bigcap_{i\in I}(\mathrm{Supp}\,\boldsymbol{A}_i)=\mathrm{Supp}(\prod_{i\in I}\boldsymbol{A}_i),$$

$$\mathrm{Supp}(\boldsymbol{A}\Delta\boldsymbol{B})=(\mathrm{Supp}(\boldsymbol{A}-\boldsymbol{B}))\cup(\mathrm{Supp}(\boldsymbol{B}-\boldsymbol{A})),\quad \mathrm{Supp}(h\bullet\boldsymbol{A})=\mathrm{Supp}\boldsymbol{A}=\mathrm{Supp}(\boldsymbol{A}^n),$$

$$\mathrm{Supp}(\boldsymbol{A}_1\times\ldots\times\boldsymbol{A}_n)=(\mathrm{Supp}\boldsymbol{A}_1)\times\ldots\times(\mathrm{Supp}\boldsymbol{A}_n),\quad \mathrm{Supp}(\times\boldsymbol{A})^n=(\times\mathrm{Supp}\boldsymbol{A})^n.$$

В теории множеств операции арифметического сложения, умножения на число, арифметического умножения и возведения в степень множеств в общем случае не определяются. Аналогами этих операций могут служить соответственно покомпонентное сложение и умножение на скаляр векторов $\mathbf{a}+\mathbf{b}=(a_1+b_1,\dots,a_n+b_n)$, $h\cdot\mathbf{a}=(ha_1,\dots,ha_n)$ и матриц $\mathrm{A}+\mathrm{B}=\|a_{ij}+b_{ij}\|_{m\times n}$, $h\cdot\mathrm{A}=\|h\ a_{ij}\|_{m\times n}$, поэлементное умножение матриц $\mathrm{A}\cdot\mathrm{B}=\|a_{ij}\cdot b_{ij}\|_{m\times n}$. Последняя операция, введенная в алгебраической теории распознавания образов [3], отличается от традиционной операции умножения матриц. При переходе к множествам арифметическое умножение и возведение в степень мультимножеств вырождаются в пересечение множеств, а арифметическое сложение множеств и умножение множества на число будут неосуществимы.

Семейство мультимножеств с введенными на нем операциями объединения, пересечения, сложения и дополнения представляет собой алгебру мультимножеств $L(\boldsymbol{Z})$, где максимальное мультимножество $\boldsymbol{Z}$ является единицей алгебры, а пустое мультимножество $\varnothing$ – нулем. Действительная неотрицательная функция $m(\boldsymbol{A})$, определенная на алгебре $L(\boldsymbol{Z})$ и удовлетворяющая условию сильной аддитивности $m(\sum_i\boldsymbol{A}_i)=\sum_i m(\boldsymbol{A}_i)$, называется мерой мультимножества. Мера мультимножества $m(\boldsymbol{A})$ обладает следующим свойствами: $m(\varnothing)=0$; слабая аддитивность $m(\bigcup_i\boldsymbol{A}_i)=\sum_i m(\boldsymbol{A}_i)$ для $\boldsymbol{A}_i\bigcap\boldsymbol{A}_j=\varnothing$; монотонность $m(\boldsymbol{A})\leq m(\boldsymbol{B})\Leftrightarrow\boldsymbol{A}\subseteq\boldsymbol{B}$; симметричность $m(\boldsymbol{A})+m(\overline{\boldsymbol{A}})=m(\boldsymbol{Z})$; непрерывность $\lim\limits_{i\to\infty}m(\boldsymbol{A}_i)=m(\lim\limits_{i\to\infty}\boldsymbol{A}_i)$; эластичность $m(h\bullet\boldsymbol{A})=hm(\boldsymbol{A})$ [11]. Меру мультимножества можно задать различными способами, например, как линейную комбинацию функций кратности: $m(\boldsymbol{A})=\sum_j w_j k_A(x_j)$, $w_j>0$. Заметим, что мощность мультимножества $|\boldsymbol{A}|$ также будет мерой мультимножества

Различные метрические пространства мультимножеств (A,d) определяются следующими видами расстояний между мультимножествами [11]:

$$d_{1p}(\boldsymbol{A},\boldsymbol{B}) = [m(\boldsymbol{A}\Delta\boldsymbol{B})]^{1/p};\quad d_{2p}(\boldsymbol{A},\boldsymbol{B}) = [m(\boldsymbol{A}\Delta\boldsymbol{B})/m(\boldsymbol{Z})]^{1/p};\quad d_{3p}(\boldsymbol{A},\boldsymbol{B}) = [m(\boldsymbol{A}\Delta\boldsymbol{B})/m(\boldsymbol{A}\bigcup\boldsymbol{B})]^{1/p},\quad (1)$$

где p – целое. Функции $d_{2p}(\boldsymbol{A},\boldsymbol{B})$ и $d_{3p}(\boldsymbol{A},\boldsymbol{B})$ удовлетворяют условию нормировки $0\leq d(\boldsymbol{A},\boldsymbol{B})\leq 1$. В силу непрерывности меры мультимножества функция $d_{3p}(\boldsymbol{A},\boldsymbol{B})$ не определена для $\boldsymbol{A}=\boldsymbol{B}=\varnothing$, поэтому по определению принимается $d_{3p}(\varnothing,\varnothing)=0$. Основное расстояние $d_{1p}(\boldsymbol{A},\boldsymbol{B})$ является метрикой типа Хемминга, традиционно используемой во многих приложениях. Полностью усредненное расстояние $d_{2p}(\boldsymbol{A},\boldsymbol{B})$ характеризует различие между двумя мультимножествами $\boldsymbol{A}$ и $\boldsymbol{B}$, отнесенное к расстоянию, максимально возможному в исходном пространстве. Локально усредненное расстояние $d_{3p}(\boldsymbol{A},\boldsymbol{B})$ задает различие, отнесенное к максимально возможной «общей части» только этих двух мультимножеств в исходном пространстве.

3. Представление многопризнаковых объектов

Выбор той или иной модели для представления и исследования структуры совокупности рассматриваемых объектов определяется их свойствами, которые выражаются признаками (атрибутами) объектов. Признаки объектов могут быть непрерывными и дискретными, количественными, качественными или смешанными.

Обычно совокупность объектов представляется множеством точек в некотором многомерном (как правило, метрическом) пространстве, оси которого соотносятся с соответствующими признаками. В прикладных задачах в качестве такого пространства достаточно часто (но, заметим, не всегда обоснованно) выбирается пространство типа евклидового. Задание расстояния между объектами позволяет оценивать близость или удаленность этих объектов относительно друг друга вне зависимости от их природы, исследовать структурные особенности совокупности объектов и всего пространства в целом.

Рассмотрим совокупность $A=\{A_1,\dots,A_k\}$ объектов, которые описываются m дискретными признаками $Q_1,\dots,Q_m$, имеющими конечное число $q_s^{e_s}$, $e_s=1,\dots,h_s$, $s=1,\dots,m$ количественных (числовых) или качественных (номинальных, либо порядковых) значений. Каждый объект A_i, i=1,…,k из совокупности $\mathcal{A}$ можно представить как точку q_i в m-мерном декартовом пространстве $Q=Q_1\times Q_2\times\ldots\times Q_m$, являющемся прямым произведением шкал значений признаков Q_s, и поставить объекту A_i в соответствие m-мерный вектор $A_i=(q_{i1}^{e_1}, q_{i2}^{e_2}, \ldots, q_{im}^{e_m})$ [5, 7, 8].

Ситуация существенным образом усложняется, если одному и тому же объекту A_i может соответствовать не один, а несколько m-мерных векторов с различающимися значениями признаков. Подобная ситуация возникает, например, когда необходимо одновременно учесть m параметров объекта A_i, измеренных n различными способами, либо когда объект A_i оценивается n независимыми экспертами по m критериям. В таком случае объект A_i представляется в m-мерном пространстве Q уже не одной точкой q_i, а группой ("облаком"), состоящей из n точек $\{q_i^{(1)},\ldots,q_i^{(n)}\}$ вида $A_i=\{(q_{i1}^{e_1(1)}, q_{i2}^{e_2(1)}, \ldots, q_{im}^{e_m(1)}),\ldots,(q_{i1}^{e_1(n)}, q_{i2}^{e_2(n)}, \ldots, q_{im}^{e_m(n)})\}$, которая должна рассматриваться и анализироваться как единое целое. При этом, очевидно, измеренные разными способами значения параметров, как и индивидуальные оценки экспертов, могут быть похожими, различающимися и даже противоречивыми, что в свою очередь может приводить к несравнимости m-мерных векторов $q_i^{(j)}=(q_{i1}^{e_1(j)}, q_{i2}^{e_2(j)}, \ldots, q_{im}^{e_m(j)})$, характеризующих один и тот же объект A_i.

Совокупность таких многомерных объектов может иметь в пространстве Q сложную структуру, достаточно трудную для анализа. Непросто ввести в этом пространстве и метрику для измерения расстояний между объектами. Указанные трудности можно преодолеть, воспользовавшись иным способом представления многопризнаковых объектов, основанным на формализме мультимножеств, которые позволяют одновременно учесть все комбинации значений количественных и качественных признаков, а также число значений каждого из этих признаков. Вместо прямого произведения m шкал значений признаков $Q=Q_1\times Q_2\times\ldots\times Q_m$ введем обобщенную шкалу признаков – множество $G=\{Q_1,Q_2,\ldots,Q_m\}$, состоящее из m групп признаков, и представим объект $A_i\in\mathcal{A}$ в виде мультимножества:

$$\boldsymbol{A}_i = \{k_{Ai}(q_1^{\,1})\bullet q_1^{\,1},\ldots,k_{Ai}(q_1^{h_1})\bullet q_1^{h_1},\ldots,k_{Ai}(q_m^{\,1})\bullet q_m^{\,1},\ldots, k_{Ai}(q_m^{h_m})\bullet q_m^{h_m}\},$$

где число $k_{Ai}(q_s^{e_s})$ указывает, сколько раз признак $q_s^{e_s}\in Q_s$ встречается в описании объекта A_i. Например, при многокритериальной оценке объекта A_i несколькими экспертами число $k_{Ai}(q_s^{e_s})$ равно числу экспертов, давших объекту A_i оценку $q_s^{e_s}$ по критерию Q_s. Объект A_i можно записать и в стандартном обозначении как $\boldsymbol{A}_i=\{k_{Ai}(x_1)\bullet x_1,\ldots,k_{Ai}(x_h)\bullet x_h\}$, определив элементы множества $G=\{x_1,\ldots,x_h\}$ следующим образом:

$$x_1=q_1^{\,1}, x_2=q_1^{\,2},\ldots,x_{h_1}=q_1^{h_1},\ x_{h_1+1}=q_2^{\,1},\ldots,x_{h_1+h_2}=q_2^{h_2},\ldots,x_{h_1+\ldots+h_{m-1}+1}=q_m^{\,1},\ldots,x_{h_1+\ldots+h_m}=q_m^{h_m},$$

где $h=h_1+\ldots+h_m$. Множество G определяет свойства совокупности объектов $\mathcal{A}=\{A_1,\ldots,A_k\}$ и играет роль домена для семейства мультимножеств $\mathcal{A}=\{\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k\}$, которые можно представлять точками метрических пространств мультимножеств $(\mathcal{A},d)$ с метриками (1).

Приведем еще несколько аналогичных примеров из других предметных областей. Пусть $\mathcal{A}=\{A_1,\ldots,A_k\}$ является массивом текстовых документов по некоторой проблематике. Этими документами могут быть, к примеру, статьи, книги, проекты, патенты, отчеты, обзоры и тому подобное. Содержание документа принято отражать с помощью отдельных слов или так называемых лексических единиц – дескрипторов, ключевых слов, терминов и прочее. Множество $G=\{x_1,\ldots,x_h\}$ таких лексических единиц образует тезаурус или проблемно-ориентированный терминологический словарь, характеризующий особенности проблематики. В этом случае каждый текстовый документ A_i можно рассматривать как совокупность повторяющихся лексических единиц x_j и представить его в виде $\boldsymbol{A}_i=\{k_{Ai}(x_1)\bullet x_1,\ldots,k_{Ai}(x_h)\bullet x_h\}$, где $k_{Ai}(x_j)$ равно числу лексических единиц x_j, присутствующих в документе $\boldsymbol{A}_i$. Заметим, что указанное представление может соответствовать и полному тексту документа, и некоторому краткому описанию его содержания, индексированному лексическими единицами.

Пусть теперь $A = \{A_1,...,A_k\}$ представляет собой коллекцию распознаваемых графических объектов: печатных или рукописных символов, линий, изображений, строк или страниц [1]. В ходе распознавания каждый распознаваемый объект A_i сравнивается с набором $G=\{x_1,...,x_h\}$ стандартных образцов целых символов или отдельных структурных элементов (фрагментов), составляющих в совокупности базу образцов, и относится к одному или нескольким образцам из стандартного набора с какой-то определенной точностью. Результат распознавания объекта $\boldsymbol{A}_i$ также можно выразить соотношением $\boldsymbol{A}_i=\{k_{Ai}(x_1)\bullet x_1,\ldots,k_{Ai}(x_h)\bullet x_h\}$, где число $k_{Ai}(x_j)$ обозначает некоторую вычисленную оценку распознавания объекта $\boldsymbol{A}_i$ по отношению к стандартному образцу x_j.

4. Примеры практических задач

Рассмотрим важную практическую задачу конкурсного отбора проектов для их финансирования из некоторого фонда или для включения в состав программы, направленной на решение какой-либо важной проблемы (научно-технической, экономической, экологической, производственной). Каждая конкурсная заявка A_i $i=1,\ldots,k$ независимо оценивается несколькими экспертами по определенным критериям. Такими критериями могут быть, например: Q_1. Важность проекта для программы; Q_2. Перспективность проекта; Q_3. Новизна подхода к решению поставленных задач; Q_4. Квалификация исполнителей проекта; Q_5. Ресурсное обеспечение работ; Q_6. Возможность быстрого выхода результатов в практику.

Каждый критерий имеет порядковую или номинальную шкалу оценок с развернутыми словесными формулировками градаций качества. Так, шкала оценок по критерию Q_6.«Возможность быстрого выхода результатов в практику» выглядит следующим образом:

q_6^1 – результаты будут обладать достаточной степенью технологичности, обеспечивающей их быстрое использования в практике;

q_6^2 – для использования запланированных результатов на практике потребуются дополнительные исследования и разработки;

q_6^3 – результаты будут носить в основном теоретический характер.

Каждый эксперт, наряду с оценкой заявки по всем критериям, дает одну из следующих рекомендаций:

r_1 – включить проект в программу;

r_2 – отклонить проект;

r_3 – отложить рассмотрение заявки и отправить проект на доработку.

Указанные рекомендации экспертов являются, по существу, правилами предварительной классификации (сортировки) множества рассматриваемых заявок $A=\{A_1,...,A_k\}$. Основываясь на заключениях экспертов, конкурсная комиссия принимает решение о включении того или иного проекта в соответствующий раздел программы. В других задачах критерии оценки объектов и правила их сортировки могут быть и иными.

Если бы заявка оценивалась только одним экспертом, то найти на множестве многокритериальных оценок обобщенное решающее правило для отбора предложений не составило бы особого труда. Известно большое число разных подходов к решению подобного рода задач классификации [2, 5, 7, 13]. В случае оценки несколькими экспертами появляется несколько различных вариантов («экземпляров») одной и той же заявки, причем и многокритериальные экспертные оценки, и заключения экспертов могут быть как схожими, так и противоречивыми. Эксперты могут относить сильно различающиеся объекты в один и тот же класс, а объекты со сходными значениями признаков – в разные классы. Несогласованность индивидуальных решающих правил может быть вызвана неоднозначностью понимания экспертами решаемой задачи, ошибками или неточностями, допущенными экспертами при первоначальной классификации объектов, субъективным различием решающих правил, используемых разными экспертами, специфичностью знаний самих экспертов, нетранзитивностью отдельных экспертных суждений и многими другими причинами. В силу качественного характера экспертных данных их агрегирование тем или иным способом представляет самостоятельную, достаточно сложную проблему. Помимо этого, вырабатывая решение о включении заявки в программу, необходимо учесть все, даже и не совпадающие заключения экспертов по принятию или отклонению заявки. Желательно поэтому иметь некое единое решающее правило для отнесения заявки к какому-либо классу, которое, во-первых, базировалось бы на характеристиках заявок, выраженных их многокритериальными оценками, а во-вторых, в наибольшей степени соответствовало бы всем индивидуальным экспертным правилам сортировки.

Другой достаточно часто встречающейся на практике задачей является нахождение рейтинга компаний, основываясь на фактических показателях деятельности компаний и/или экспертных оценках по многим критериям. Перечень критериев зависит от целей анализа. Например, компании, действующие в некотором секторе рынка, можно оценивать по следующим критериям: Q_1. Уровень деловой активности; Q_2. Объем прибыли от реализации продукции; Q_3. Объем продаж; Q_4. Число выполненных проектов; Q_5.Квалификация персонала; Q_6. Численность сотрудников компании; и тому подобное.

Шкалы критериев оценки могут быть как количественными, так и качественными. Для удобства оценки и сравнения компаний количественные критерии можно трансформировать в качественные с небольшим числом упорядоченных градаций шкал. Шкалы критериев Q_4.«Число выполненных проектов» и Q_6. «Численность сотрудников компании» могут иметь, например, такой вид:

q_4^1 – очень высокое (больше ста);
q_4^2 – высокое (от пятидесяти до ста);
q_4^3 – среднее (от десяти до пятидесяти);
q_4^4 – низкое (меньше десяти).

Каждая компания из совокупности $A=\{A_1,...,A_k\}$ оценивается несколькими экспертами по всем критериям. Оценки разных экспертов могут отличаться друг от друга и даже противоречить друг другу. Каждая компания A_i представляет собой многопризнаковый объект, и определение рейтинга можно рассматривать как задачу упорядочивания многопризнаковых объектов. Основной трудностью при решении таких задач является необходимость учета всех описаний объекта – различающихся оценок, сделанных разными экспертами, при условии, что не существует «главного» эксперта, и мнения всех экспертов считаются одинаково важными.

В существующих методах классификации и упорядочивания объектов построение итогового решения производится либо на основе информации, полученной от одного источника, либо путем согласования или усреднения различных оценок [5, 7, 13]. Однако бывает крайне сложно, а иногда и невозможно найти согласованное мнение экспертов, например, если эксперты работают независимо друг от друга и не могут знать оценок, данных другими экспертами. Поэтому необходимы методы классификации и упорядочения многопризнаковых объектов, которые позволяли бы одновременно учитывать оценки, в том числе и противоречивые, всех экспертов без поиска компромисса между мнениями отдельных экспертов.

5. Классификация многопризнаковых объектов

Перейдем к формальной постановке задачи классификации совокупности объектов $A=\{A_1,...,A_k\}$, которые описываются m дискретными признаками $Q_1,\dots,Q_m$, имеющими качественные (номинальные, либо порядковые) значения $q_s^{e_s}$, $e_s=1,\dots,h_s$, $s=1,\dots,m$. Порядковые значения качественного признака обычно считаются упорядоченными от лучшего значения к худшему $q_s^1 \succ q_s^2 \succ \ldots \succ q_s^{h_s}$. Пусть каждый из объектов A_i, i=1,…,k может существовать в n экземплярах, которые отличаются наборами признаков, его характеризующих. Различными экземплярами многопризнакового объекта являются, например, наборы экспертных оценок проекта, которые даны n экспертами. Каждый из экспертов предварительно относит объекты A_i к одному из нескольких классов X_t, t=1,…,f, тем самым появляется n в общем случае несовпадающих индивидуальных сортировок совокупности объектов A. Принадлежность экземпляра объекта A_i к некоторому классу X_t выражается индивидуальным правилом сортировки R, которое может считаться еще одним качественным признаком объекта со шкалой значений $R=\{r_t\}$. Таким образом, каждый экземпляр объекта описывается только одним каким-то значением признака из каждой группы $Q_1,\ldots,Q_m,R$. Других дополнительных предположений об особенностях классов, признаков объектов и их значений (важности, предпочтительности, характерности, упорядоченности и прочее) не делается. Требуется построить одно или несколько обобщенных решающих правил, составленных из небольшого числа значений признаков, которые относили бы объекты к заданным классам наилучшим (в смысле близости к предварительным индивидуальным сортировкам) образом.

Сопоставим многопризнаковому объекту A_i мультимножество вида

$$\boldsymbol{A}_i = \{k_{Ai}(q_1^{\,1})\bullet q_1^{\,1},\ldots,k_{Ai}(q_1^{h_1})\bullet q_1^{h_1},\ldots,k_{Ai}(q_m^{\,1})\bullet q_m^{\,1},\ldots,k_{Ai}(q_m^{h_m})\bullet q_m^{h_m},k_{Ai}(r_1)\bullet r_1,\ldots,k_{Ai}(r_f)\bullet r_f\}$$

над доменом $G=\{Q_1,\ldots,Q_m,R\}$, где, напомним, $k_{Ai}(q_s^{e_s})$ и $k_{Ai}(r_t)$ равны числу экспертов, давших объекту A_i оценки $q_s^{e_s}$ и r_t. Представление объекта A_i в таком виде может трактоваться также как способ выражения экспертных правил сортировки:

$$\text{ЕСЛИ } \langle\text{условия}\rangle\text{, ТО } \langle\text{решение}\rangle. \qquad (2)$$

Терм ⟨условия⟩ ассоциируется с различными комбинациями значений признаков $q_s^{e_s}$, описывающими свойства объекта A_i, а терм ⟨решение⟩ – с принадлежностью объекта A_i к классу X_t. В терм ⟨решение⟩ входят совокупность индивидуальных заключений экспертов по предварительной сортировке объекта A_i, а также некоторое правило, позволяющее относить объект A_i к определенному классу X_t. Это может быть, например, правило простого большинства голосов, в соответствии с которым объект A_i будет считаться принадлежащим к классу X_t, если $k_{Ai}(r_t)>k_{Ai}(r_p)$ для всех $p\neq t$, или правило квалифицированного большинства голосов, по которому должно выполняться условие $k_{Ai}(r_t)>\sum_{p\neq t} k_{Ai}(r_p)$, или любое другое правило.

Для простоты будем считать, что результатом классификации должно быть разложение совокупности объектов A только на два класса X_a и X_b. Возможны различные подходы к агрегированию многопризнаковых объектов [11,12]. Рассмотрим наиболее простой и типичный случай, когда каждая группа объектов формируются путем сложения соответствующих им мультимножеств. В этом случае учитываются значения всех признаков, характеризующих все объекты, входящие в группу. Тогда мультимножество

$$\boldsymbol{X}_t = \{k_{Xt}(q_1^{\,1})\bullet q_1^{\,1},\ldots,k_{Xt}(q_1^{h_1})\bullet q_1^{h_1},\ldots,k_{Xt}(q_m^{\,1})\bullet q_m^{\,1},\ldots,k_{Xt}(q_m^{h_m})\bullet q_m^{h_m},k_{Xt}(r_a)\bullet r_a,k_{Xt}(r_b)\bullet r_b\},$$

которое представляет свой класс объектов, можно записать как сумму мультимножеств

$$\boldsymbol{X}_t = \sum\nolimits_{i\in I_t} \boldsymbol{A}_i\ .$$

Здесь I_t – подмножество индексов i для объектов класса X_t, $t=a,b$, $I_a\bigcup I_b=\{1,\ldots,k,\}$, $I_a\bigcap I_b=\varnothing$. Перепишем мультимножество $\boldsymbol{X}_t$ в виде следующего разложения на новые мультимножества:

$$\boldsymbol{X}_t = \sum\nolimits_{s=1}^{m} \boldsymbol{Q}_{st} + \boldsymbol{R}_t,$$

где слагаемые суть «однопризнаковые» мультимножества

$$\boldsymbol{Q}_{st} = \sum\nolimits_{i\in I_t} \boldsymbol{A}_{iqs}\ ,\quad \boldsymbol{A}_{iqs} = \{k_{Ai}(q_s^{\,1})\bullet q_s^{\,1},\ldots,k_{Ai}(q_s^{h_s})\bullet q_s^{h_s}\};$$

$$\boldsymbol{R}_t = \sum\nolimits_{i\in I_t} \boldsymbol{A}_{ir}\ ,\quad \boldsymbol{A}_{ir} = \{k_{Ai}(r_a)\bullet r_a,\ k_{Ai}(r_b)\bullet r_b\}.$$

Так как каждый экземпляр объекта $\boldsymbol{A}_i$ может обладать только единственными значениями признаков $k_{Ai}(q_s^{e_s})$ и $k_{Ai}(r_t)$ из каждой группы признаков Q_s и R, то выполняются следующие условия для мощностей мультимножеств:

$$|\boldsymbol{Q}_{sa}| + |\boldsymbol{Q}_{sb}| = kn, \quad |\boldsymbol{R}_a| + |\boldsymbol{R}_b| = kn, \quad |\boldsymbol{X}_a| + |\boldsymbol{X}_b| = kn(m+1),$$

где, напомним, k равно числу объектов, m – числу групп признаков, n – числу экспертов.

Требование бинарной декомпозиции совокупности объектов только на два класса не является принципиальным ограничением. Если необходимо рассортировать объекты на большее число классов, можно сначала разбить совокупность объектов на две группы, затем одну из них или обе группы – на подгруппы, и так далее. Например, конкурсные заявки можно разделить на принятые и отклоненные, отклоненные заявки – на отложенные для дальнейшей доработки и окончательно не принятые, и так далее.

Рассмотрим теперь метрическое пространство мультимножеств (A, d) с метрикой d, определяемой одним из выражений (1). Очевидно, что объекты $\boldsymbol{A}_i$, которые попали в разложение $\{\boldsymbol{R}_a, \boldsymbol{R}_b\}$, сделанное по результатам предварительных индивидуальных сортировок, образуют наилучшую из всех возможных декомпозиций рассматриваемой совокупности объектов $A=\{\boldsymbol{A}_1,...,\boldsymbol{A}_k\}$ на два класса. Расстояние $d^*=d(\boldsymbol{R}_a, \boldsymbol{R}_b)$ между мультимножествами $\boldsymbol{R}_a$ и $\boldsymbol{R}_b$ будет предельно возможным расстоянием в метрическом пространстве мультимножеств (A, d) между объектами, входящими в разные классы. При идеальных предварительных сортировках объектов, в которых отсутствуют противоречия между индивидуальными правилами, расстояние d^* будет равно соответственно $d_{1p}{}^*=[kn]^{1/p}$, $d_{2p}{}^*=[1/h]^{1/p}$, $d_{3p}{}^*=1$. Здесь n есть число индивидуальных сортировок, совпадающее с числом экспертов, h – общее число значений всех признаков, описывающих объекты, равное в нашем случае $h=h_1+...+h_m+f$.

Задача поиска обобщенного решающего правила классификации многопризнаковых объектов сводится к m оптимизационным задачам для каждой группы признаков Q_s

$$d(\boldsymbol{Q}_{sa}, \boldsymbol{Q}_{sb}) \to \max d(\boldsymbol{Q}_{sa}, \boldsymbol{Q}_{sb}) = d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*). \tag{3}$$

Таким образом, требуется найти в каждой группе признаков Q_s новые мультимножества $\boldsymbol{Q}_{sa}^*$ и $\boldsymbol{Q}_{sb}^*$, которые будут расположены на максимально возможном расстоянии в метрическом пространстве мультимножеств (A, d) и принадлежат к разным классам. Мультимножество $\boldsymbol{Q}_{st}^*$ ($t=a,b$), относящееся к одному и тому же классу, представляется в виде суммы двух подмультимножеств $\boldsymbol{Q}_{st}^*=\boldsymbol{Q}_{st}^{*1}+\boldsymbol{Q}_{st}^{*2}$. Решение каждой из задач (3) выражается через подмультимножества $\boldsymbol{Q}_{st}^{*1}$, $\boldsymbol{Q}_{st}^{*2}$ и определяет наилучшую бинарную декомпозицию $\{\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*\}$ имеющейся совокупности многопризнаковых объектов $A=\{\boldsymbol{A}_1,...,\boldsymbol{A}_k\}$ по s-ой группе признаков. Когда число h_s значений $q_s^{e_s}$ каждого из признаков невелико, решение задачи (3) не вызывает существенных трудностей и может быть получено даже путем перебора.

Значение признака q_s^*, которое определяет границу между сгенерированными слагаемыми $\boldsymbol{Q}_{st}^{*1}$ и $\boldsymbol{Q}_{st}^{*2}$ внутри каждой пары, назовем граничным. Различные комбинации граничных значений признаков $\{q_s^*\}$ из разных групп признаков Q_s, задают условия отнесения объекта $\boldsymbol{A}_i$ к соответствующему классу $\boldsymbol{X}_t$ и образуют возможные обобщенные правила классификации совокупности многопризнаковых объектов вида (2).

Граничные признаки q_s^* можно упорядочить по величине расстояния $d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*)$. Для построения обобщенных правил классификации следует использовать граничные признаки q_s^*, занимающие первые места в такой ранжировке. Чем ближе значение $d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*)$ к расстоянию $d^*=d(\boldsymbol{R}_a, \boldsymbol{R}_b)$, тем более точной будет аппроксимация первоначальной индивидуальной экспертной сортировки объектов. Оценить точность аппроксимации по s-ой группе признаков можно, например, выражением

$$\rho_s = d(\boldsymbol{Q}_{sa}^*, \boldsymbol{Q}_{sb}^*)/d(\boldsymbol{R}_a, \boldsymbol{R}_b).$$

В итоговое обобщенное решающее правило следует включать граничные значения признаков q_s^*, имеющие показатель точности ρ_s, превышающей желаемый пороговый уровень ρ_0. Заметим, что величина ρ_s показателя точности аппроксимации характеризует своего рода относительную важность s-ой группы признаков Q_s в обобщенном решающем правиле классификации.

6. Упорядочение многопризнаковых объектов

Дадим формальную постановку задачи упорядочения совокупности многопризнаковых объектов $\mathsf{A}=\{A_1,...,A_k\}$, которые оцениваются n экспертами по m критериям $Q_1,\dots,Q_m$. Каждый критерий Q_s имеет порядковую шкалу количественных или качественных оценок $\{q_s^{e_s}\}$, $e_s=1,\dots,h_s$, $s=1,\dots,m$, которые упорядочены от лучшего значения к худшему $q_s^1\succ q_s^2\succ\dots\succ q_s^{h_s}$. Предполагается, что разные критерии могут иметь различную относительную важность, но значения оценок, относящихся к одному и тому же критерию, равноценны. Будем также считать, что каждый объект оценивается всеми n экспертами по всем m критериям, и что экспертные оценки независимы. В таком случае можно выделить два объекта (возможно, гипотетических) – абсолютно лучший и абсолютно худший, которым все эксперты дали соответственно самые лучшие и самые худшие оценки по всем критериям. Требуется, исходя из многокритериальных оценок объектов, упорядочить объекты от лучшего к худшему.

Представим многопризнаковый объект как мультимножество

$$\boldsymbol{A}_i = \{k_{Ai}(q_1^{\,1})\bullet q_1^{\,1},\dots,k_{Ai}(q_1^{h_1})\bullet q_1^{h_1},\dots,k_{Ai}(q_m^{\,1})\bullet q_m^{\,1},\dots,k_{Ai}(q_m^{h_m})\bullet q_m^{h_m}\}$$

над доменом $G=\{Q_1,\dots,Q_m\}$, являющимся множеством критериальных оценок. Функция кратности $k_{Ai}(q_s^{e_s})$ мультимножества характеризует здесь количество экспертов, давших объекту A_i оценку $q_s^{e_s}$. Наилучшему и наихудшему объектам соответствуют мультимножества

$$\boldsymbol{A}_{\max} = \{n\bullet q_1^{\,1},0,\dots,0,\ n\bullet q_2^{\,1},0,\dots,0,\dots,\ n\bullet q_m^{\,1},0,\dots,0\}, \tag{4}$$

$$\boldsymbol{A}_{\min} = \{0,\dots,0,n\bullet q_1^{h_1},\ 0,\dots,0,n\bullet q_2^{h_2},\dots,\ 0,\dots,0,n\bullet q_m^{h_m}\}, \tag{5}$$

и их принято называть идеальным и антиидеальным решениями. Задача упорядочения многопризнаковых объектов сводится, таким образом, к упорядочению мультимножеств.

Будем считать многопризнаковые объекты точками некоторого метрического пространства мультимножеств (A, d), например, с основной метрикой (типа Хемминга), которая задается формулой (1), принимающей в данном случае вид

$$d_1(\boldsymbol{A},\boldsymbol{B}) = m(\boldsymbol{A}\Delta\boldsymbol{B}) = \sum_{s=1}^{m} w_s \sum_{e_s=1}^{h_s} |k_A(q_s^{e_s}) - k_B(q_s^{e_s})|,$$

где $w_s>0$ – коэффициенты относительной важности критериев Q_s. Считаем, что объект $\boldsymbol{A}_i$ лучше объекта $\boldsymbol{A}_j$ ($\boldsymbol{A}_i\succ\boldsymbol{A}_j$), если

$$d_1(\boldsymbol{A}_{\max}, \boldsymbol{A}_i) < d_1(\boldsymbol{A}_{\max}, \boldsymbol{A}_j). \tag{6}$$

Упорядочим все объекты по величине их расстояния от идеального решения. Если для некоторых объектов $d_1(\boldsymbol{A}_{\max},\boldsymbol{A}_i)=d_1(\boldsymbol{A}_{\max},\boldsymbol{A}_j)$, то объекты $\boldsymbol{A}_i$ и $\boldsymbol{A}_j$ будут или эквивалентными, или несравнимыми. Тем самым полученное ранжирование объектов окажется нестрогим.

Так как каждый объект $\boldsymbol{A}_i$ оценивается n экспертами по всем m критериям, то выражение для расстояния от идеального решения $\boldsymbol{A}_{\max}$ до объекта $\boldsymbol{A}_i$ принимает вид:

$$d_1(\boldsymbol{A}_{\max}, \boldsymbol{A}_i) = \sum_{s=1}^{m} w_s \sum_{e_s=1}^{h_s} |k_{A\max}(q_s^{e_s}) - k_{Ai}(q_s^{e_s})| = 2\sum_{s=1}^{m} w_s\,[n - k_{Ai}(q_s^{\,1})].$$

Условие (6) сравнения многопризнаковых объектов приобретает тогда следующую форму: объект $\boldsymbol{A}_i$ лучше объекта $\boldsymbol{A}_j$ ($\boldsymbol{A}_i\succ\boldsymbol{A}_j$), если

$$\sum_{s=1}^{m} w_s\, k_{Ai}(q_s^{\,1}) > \sum_{s=1}^{m} w_s\, k_{Aj}(q_s^{\,1}).$$

Таким образом, правило упорядочения многопризнаковых объектов сводится к сравнению взвешенных сумм $S_{Ai}^{\,1}=\sum_s w_s\, k_{Ai}(q_s^{\,1})$ первых (наилучших) оценок объектов по всем критериям Q_s. Лучшим будет тот объект $\boldsymbol{A}_i$, у которого эта сумма S_{Ai}^{1} будет больше. Если найдутся группы эквивалентных или несравнимых объектов $\boldsymbol{A}_{i1},\ldots,\boldsymbol{A}_{it}$, имеющих одинаковые суммы S_{Ai}^{1}, то нужно вычислить для каждого объекта $\boldsymbol{A}_{ir}$, r=1,…t в соответствующей группе взвешенную сумму $S_{Air}^{2} = \sum_s w_s\, k_{Air}(q_s^{2})$ всех вторых оценок по всем критериям Q_s и упорядочить объекты внутри каждой группы от лучшего к худшему по величинам S_{Air}^{2} сумм вторых оценок. Если останутся подгруппы эквивалентных или несравнимых объектов $\boldsymbol{A}_{iru},\ldots,\boldsymbol{A}_{irv}$, имеющих одинаковые суммы S_{Air}^{2}, то вычислить для каждого объекта $\boldsymbol{A}_{irp}$ в соответствующей подгруппе взвешенную сумму $S_{Airp}^{3} = \sum_s w_s\, k_{Airp}(q_s^{3})$ всех третьих оценок по всем критериям Q_s и упорядочить объекты внутри каждой подгруппы от лучшего к худшему по величинам сумм S_{Airp}^{3} третьих оценок. Процедура продолжается до полного упорядочения всех объектов из совокупности $A = \{\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k\}$. Если число h_s значений оценок $q_s^{e_s}$ у некоторых критериев Q_s окажется меньше требуемого на данном b-ом шаге алгоритма, то следует считать $k_{Air\ldots p}(q_s^{b})=0$.

В приведенном алгоритме предполагалась различная относительная важность критериев Q_s, выражаемая коэффициентами $w_s > 0$, $\sum_s w_s = 1$. Проблема определения (вычисления) важности критериев имеет самостоятельное значение [8]. В случае, когда все критерии считаются одинаково важными, все коэффициенты w_s полагаются равными 1.

Аналогично можно построить процедуру упорядочения многопризнаковых объектов $\boldsymbol{A}_i$ по отношению к антиидеальному решению $\boldsymbol{A}_{\min}$, заданному выражением (5), считая, что объект $\boldsymbol{A}_i$ лучше объекта $\boldsymbol{A}_j$ ($\boldsymbol{A}_i \succ \boldsymbol{A}_j$), если он находится дальше от антиидеального решения $\boldsymbol{A}_{\min}$, то есть если выполняется условие $d_1(\boldsymbol{A}_{\min}, \boldsymbol{A}_i) > d_1(\boldsymbol{A}_{\min}, \boldsymbol{A}_j)$. Подчеркнем, однако, что упорядочение совокупности многопризнаковых объектов $A = \{\boldsymbol{A}_1,\ldots,\boldsymbol{A}_k\}$ по отношению к антиидеальному решению может не совпадать с упорядочением по отношению к идеальному решению.

7. Заключение

Проблемы классификации и упорядочения объектов, которые описываются многими количественными и качественными признаками, причем каждый из объектов может существовать в нескольких различающихся «экземплярах», являются достаточно трудными. Эти трудности имеют и содержательные основания (например, некорректность применения процедур «усреднения» качественных признаков), и формальные причины (например, большая размерность задачи). Преодолеть главные из перечисленных трудностей оказалось возможным благодаря использованию нового теоретического инструментария, основанного на понятии мультимножества. Теория мультимножеств позволяет разрабатывать новые методы анализа данных и решения новых классов задач, которые содержат новые способы преобразований исходной информации, не приводящие к потере или искажению данных. В работе изложены методы классификации и упорядочения многопризнаковых объектов, основанные на их представлении с помощью мультимножеств. В качестве признаков могут выступать количественные и качественные, в том числе и противоречивые данные.

Предложенный подход к построению обобщенного решающего правила для классификации многопризнаковых объектов, которое аппроксимирует большое число предварительных противоречивых сортировок, был проверен на результатах экспертной оценки и конкурсного отбора проектов при формирования государственной научно-технической программы по высокотемпературной сверхпроводимости [6]. Каждая представленная на конкурс заявка независимо оценивалась 3 экспертами по 6 качественным критериям, которые давали также свое заключение по принятию или отклонению заявки. Всего было подано более 250 заявок, около 170 из них было отобрано для включения в программу. Разработанным методом были сформулированы несколько решающих правил, одно из которых полностью совпало с примененным на практике [6]. Обобщенное решающее правило классификации объектов позволило также выделить наиболее важные для отбора проектов критерии и выявить расхождения в индивидуальных правилах сортировки проектов, применявшихся экспертами.

Упорядочивание объектов по их близости к наилучшему объекту в метрическом пространстве мультимножеств дает возможность получать как строгое, так и нестрогое ранжирование объектов при равнозначных или различных по важности критериях. Рассмотренный метод упорядочения многопризнаковых объектов был применен для построения рейтинга российских компаний, работающих в секторе информационно-коммуникационных технологий [4]. Результаты экспертизы обрабатывались по описанной процедуре в предположении равнозначности критериев. В итоге из 50 оцененных компаний были выделены 30 ведущих высокотехнологичных компаний, а также составлены рейтинги 10 наиболее динамично развивающихся компаний и 10 ведущих разработчиков программного обеспечения.

Литература

1. Арлазаров В.Л., Логинов А.С., Славин О.А. Характеристики программ оптического распознавания текста.//Программирование, 2002, №3, 45-63.
2. Дорофеюк А.А. Алгоритмы автоматической классификации.//Автоматика и телемеханика. 1971, №12, 78-113.
3. Журавлев Ю.И. Корректные алгебры над множествами некорректных (эвристических) алгоритмов.I,II,III.//Кибернетика, 1977, №4, 14-21; 1977, №6, 21-27; 1978, №2, 35-43.
4. Кто в России самый интеллектуальный? Рейтинг ведущих российских разработчиков высоких технологий.//Компания, 2000, №47(143), 38-39.
5. Ларичев О.И., Мошкович Е.М. Качественные методы принятия решений. Вербальный анализ решений. – М.: Наука, Физматлит, 1996.
6. Ларичев О.И., Прохоров А.С., Петровский А.Б., Стернин М.Ю., Шепелев Г.И. Опыт планирования фундаментальных исследований на конкурсной основе.//Вестник АН СССР, 1989, №7, 51-61.
7. Миркин Б.Г. Анализ качественных признаков и структур. – М.: Статистика, 1980.
8. Ногин В.Д. Принятие решений в многокритериальной среде: количественный подход. – М.: Физматлит, 2002.
9. Петровский А.Б. Метрические пространства мультимножеств.//Доклады Академии наук, 1995, Т.344, №2, 175-177.
10. Петровский А.Б. Основные понятия теории мультимножеств. – М.: Едиториал УРСС, 2002.
11. Петровский А.Б. Пространства множеств и мультимножеств. – М.: Едиториал УРСС, 2003.
12. Petrovsky A. Method for approximation of diverse individual sorting rules.//Informatica, 2001, V.12, №1, 109-118.
13. Roy B. Multicriteria methodology for decision aiding. – Kluwer Academic Publishers, Dodrecht, 1996.

MULTIPLE CRITERIA DECISION MAKING WITH CONTRADICTORY DATA: MULTISET THEORY APPROACH

Alexey B.Petrovsky

New methods for classifying and ordering objects, that are characterized by many diverse (quantitative and qualitative) attributes, and may exist in several copies with different, in particular, contradictory values of attributes, are considered in the paper. Methods are based on the theory of multiset metric spaces, and are applied for solving case studies: competitive selection of projects and ranking companies, which are estimated by several experts by many qualitative criteria.

Петровский А. Б. Многокритериальное принятие решений по противоречивыми данным: подход теории мультимножеств // *Информационные технологии и вычислительные системы*.— 2004.— № 2.— С. 56–66.

```
@Article{Petrovsky_2004,
  author =        "Петровский, А. Б.",
  title =         "Многокритериальное принятие решений по противоречивыми
                   данным: подход теории мультимножеств",
  journal =       "Информационные технологии и вычислительные системы",
  number =        "2",
  pages =         "56--66",
  year =          "2004",
  language =      "russian",
}
```