УДК 510.22

ГРУППОВОЕ МНОГОКРИТЕРИАЛЬНОЕ ПРИНЯТИЕ РЕШЕНИЙ С НЕСОВПАДАЮЩИМИ ПРЕДПОЧТЕНИЯМИ*

А. Б. ПЕТРОВСКИЙ

Институт системного анализа Российской академии наук

e-mail: pab@isa.ru

Рассматривается новый подход к групповому многокритериальному принятию решений, где предпочтения нескольких лиц, принимающих решения, не совпадают и, возможно, противоречивы, а варианты решения оцениваются по многим разнородным критериям. Подход основан на методологии вербального анализа решений и теории мультимножеств. Приведены примеры применения подхода для решения практических задач классификации и упорядочения многопризнаковых объектов.

Ключевые слова: групповое принятие решений, теория мультимножеств, классификация, упорядочение, многопризнаковые объекты.

Введение

Типичными задачами теории принятия решений является выделение одного или нескольких лучших объектов (вариантов, альтернатив, кандидатов, и т.п.), упорядочение или ранжирование объектов по их свойствам, классификация или сортировка объектов по заданным категориям. Свойства рассматриваемых объектов характеризуются многими разнородными признаками (атрибутами) или оценками по многим критериям, имеющим количественные и/или качественные шкалы.

Решение задач индивидуального выбора осуществляется на основе предпочтений единственного лица, принимающего решение (ЛПР). Поэтому требуется, чтобы его субъективные предпочтения были согласованными. При наличии нескольких независимых ЛПР коллективный выбор зависит от различных предпочтений многих участников, которые не обязаны совпадать. Несогласованность индивидуальных суждений обусловливается разнообразием и несовпадением целей и интересов разных ЛПР, собственным пониманием решаемой задачи, специфичностью знаний самих ЛПР, и многими другими обстоятельствами. Еще одним источником несогласованности информации в задачах группового выбора служит неоднозначность представления объектов, когда один и тот же объект может существовать в нескольких версиях (экземплярах), имеющих, в частности, и противоречивые описания. При этом сами объекты должны рассматриваться и анализироваться как единое целое, а свертка значений признаков или невозможна, или математически некорректна. Множественность представления многопризнаковых объектов возникает, например, когда объекты оцениваются несколькими экспертами по многим критериям.

Известно достаточно много методов индивидуального выбора объектов, описываемых многими количественными и/или качественными признаками [3, 6, 8-11, 15-18]. Сравнительно немного методов группового упорядочения многопризнаковых объектов [6, 12] и практически отсутствуют методы групповой классификации. Главные трудности при коллективном принятии решений связаны с необходимостью учета несогласованных и противоречивых предпочтений нескольких ЛПР, а также обработкой больших объемов числовых и вербальных данных, не прибегая к дополнительным преобразованиям типа усреднения, смешивания, взвешивания, которые могут привести к необоснованным и необратимым искажениям исходной информации.

---

* Работа поддержана программами фундаментальных исследований президиума РАН «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация» и ОНИТ РАН «Информационные технологии и методы анализа сложных систем», Российским фондом фундаментальных исследований (проекты 08-01-00247, 08-07-13532, 09-07-00009, 09-07-12111).

Преодолеть указанные выше трудности оказалось возможным, воспользовавшись аппаратом теории мультимножеств для представления многопризнаковых объектов и обработки данных. В работе рассмотрены методы групповой классификации и упорядочения объектов, существующих в нескольких версиях, которые описываются многими разнородными и несовпадающими признаками. Методы основаны на теории метрических пространств мультимножеств и использованы в практических задачах группового многокритериального принятия решений.

1. Методы многокритериального принятия решений

Предпочтения ЛПР являются ключевым фактором рациональности выбора. ЛПР формализует свои предпочтения, задавая характеристики анализируемой проблемы и свойства рассматриваемых объектов, сравнивая варианты решения, оценивая качество сделанного выбора. Предпочтения могут быть заданы бинарными отношениями, функциями (ценности, полезности, целевыми), решающими правилами, которые имеют логическую, математическую и вербальную форму. В то же время, решая задачу, человек может выражать свои предпочтения непоследовательно, совершать ошибки, допускать противоречия.

В случае индивидуального выбора постулируется согласованность субъективных предпочтений отдельного ЛПР, которая обычно состоит в транзитивности его суждений. Для выявления и устранения возможных несоответствий в суждениях единственного ЛПР в методах принятия решений применяют специальные процедуры.

Коллективный выбор, осуществляемый группой независимых ЛПР, принципиально отличается от индивидуального выбора, так как связан с необходимостью агрегирования многих предпочтений. Каждое ЛПР может иметь свои персональные ценности, собственные цели, интересы и источники информации. Как следствие, индивидуальные субъективные предпочтения нескольких персон могут быть схожими, различными и несовпадающими. В известных методах группового принятия решений обычно стремятся исключить несогласованность и противоречивость индивидуальных суждений, либо заменяют совокупность многих мнений одним общим мнением, выражающим некоторую «усредненную» точку зрения или наиболее согласованную со всеми суждениями. Вместе с тем, индивидуальные предпочтения не всегда допустимо усреднять или согласовывать. Поэтому существует потребность в новых подходах, обеспечивающих возможность рационального выбора, основываясь на противоречивой информации.

Перечислим вкратце некоторые известные методы принятия решений [6]. Выделение лучших вариантов проводится с помощью разнообразных методов оптимального выбора, основанных на поиске экстремума одной или многих функций, характеризующих эффективность или качество решения. В методах многокритериальной оптимизации и многомерной полезности обобщенный критерий обычно задается сверткой многих частных числовых критериев в виде «взвешенной суммы». Однако определение весов исходных показателей является достаточно серьезной проблемой. При большом числе частных критериев построение обобщенной функции полезности сопряжено с большими трудозатратами ЛПР.

Для упорядочения объектов в целом или по многим критериям часто используются методы, основанные на парных сравнениях объектов. При наличии многих критериев и/или нескольких ЛПР результирующее упорядочение объектов строится в методе TOPSIS [6, 12] на основе парных сравнений векторов оценок, представляющих объекты. В методах аналитической иерархии [6, 8] варианты ранжируются по их приоритетности, которая последовательно вычисляется путем попарного сравнения вариантов, критериев их оценки и участников по отношению к глобальной цели решаемой задачи. Полное упорядочение объектов получается, если сравнимы все пары вариантов и транзитивны предпочтения единственного ЛПР. Если некоторые из вариантов окажутся несравнимыми, то упорядочение будет частичным. Если же все варианты несравнимы, то упорядочить их не удаётся. При увеличении числа объектов, критериев и ЛПР число возможных сопоставлений резко возрастает. В таких случаях построение итогового упорядочения объектов существенно затрудняется из-за возможных неточностей и противоречий в суждениях ЛПР.

Прямая сортировка объектов по заданным классам – один из наиболее популярных методов классификации вследствие простоты применения. Каждый объект, оцененный по одному числовому критерию, немедленно попадает в один из указанных классов. Достаточно часто для упорядочения и классификации объектов применяют функции ценности или полезности. Весьма распространены методы, основанные на взвешенной свертке критериев, например, методы многомерной полезности MAUT [6]. Порядковая сортировка многопризнаковых объектов в методе TOMASO [15] проводится с помощью интеграла Шоке, который агрегирует семейства дискриминантных функций. В интерактивной процедуре классификации [13] предпочтения единственного ЛПР описываются линейной функцией полезности, являющейся взвешенной суммой многих скалярных критериев. Однако эти и подобные им методы не позволяет объяснить полученные результаты, поскольку невозможно восстановить исходные данные по агрегированным показателям. Кроме того, назначение ЛПР весов критериев является субъективной и не имеющей строгих обоснований процедурой, а построение функции полезности при большом числе критериев сопряжено с большими трудозатратами ЛПР.

В семействе методов ELECTRE [16-18] объекты оцениваются по многим критериям, имеющим балльные шкалы и разные веса, которые задаются ЛПР. Объекты сравниваются попарно по так называемому отношению ограниченной предпочтительности (outranking relation), которое определяется специальными индексами согласия (конкорданса) и несогласия (дискорданса). Значения этих индексов вычисляются в процессе решения задачи и используются для упорядочения объектов и построения границ классов. Классификация многокритериальная альтернатив с применением огрубленных множеств предложена в [9, 11]. Предпочтения ЛПР выражаются с помощью наборов решающих правил, которые с разной степенью определенности относят альтернативы в заданные классы. Метод оперирует с достаточно большим числом решающих правил классификации, трудным для непосредственного анализа ЛПР, и требует обучения на специально выделенных массивах данных.

Иной подход к упорядочению и порядковой классификации многопризнаковых объектов предлагает методология вербального анализа решений, который ориентирован на слабо структурируемые проблемы выбора, описываемые качественными и количественными признаками. Отличительные особенности вербального анализа решений: получение информации от ЛПР и экспертов на привычном для их профессиональной деятельности языке; использование нечисловой информации без каких-либо ее преобразований в числовую на всех этапах анализа и решения задачи выбора; проверка согласованности суждений ЛПР и устранение выявленных противоречий; возможность объяснения полученного решения [3, 10]. Оценка и сравнение могут проводиться как для всех гипотетически возможных, так и для конкретных альтернатив. Для упорядочения вариантов в семействе методов ЗАПРОС используется специальная так называемая единая порядковая шкала, которая конструируется из комбинаций оценок на шкалах критериев. Методы ОРКЛАСС, СТЕПКЛАСС и ЦИКЛ позволяют строить полную и непротиворечивую классификацию многокритериальных альтернатив.

2. Групповой вербальный анализ решений

Групповой вербальный анализ решений – новый методологический подход в теории многокритериального принятия решений. Объектами группового вербального анализа служат слабо структурируемые проблемы выбора, где предпочтения нескольких ЛПР могут быть несогласованными, а варианты описываются многими разнородными количественными и качественными признаками и могут существовать в нескольких версиях (экземплярах).

Сформулируем основные задачи группового многокритериального выбора. Задана совокупность объектов (вариантов, альтернатив) $A=\{A_1,...,A_n\}$, каждый из которых оценивается k экспертами по m критериям $Q_1,...,Q_m$, характеризующим особенности объекта. Каждый критерий (содержательный признак) Q_s, $s=1,...,m$ имеет свою собственную порядковую или номинальную шкалу $X_s=\{x_s^1,...,x_s^{h_s}\}$, числовую, символьную или вербальную. Предполагается, что разные критерии могут иметь различную относительную важность $w_s>0$. Будем также считать, что каждый объект оценивается всеми k экспертами по всем m критериям, и что оценки экспертов независимы. Таким образом, многопризнаковый объект A_i, $i=1,...,n$ присутствует в k версиях (экземплярах) $A_i^{(f)}$, $f=1,...,k$. Основываясь на предпочтениях ЛПР, требуется упорядочить объекты, например, от лучших к худшим.

В задаче классификации с несколькими учителями каждый эксперт наряду с оценкой относит каждый объект A_i, а точнее его версию $A_i^{(f)}$, к одному из классов $C_1,...C_g$, которые могут быть упорядочены по предпочтительности, либо нет. Принадлежность объекта A_i к классу C_l, $l=1,...,g$ описывается сортирующим атрибутом R со шкалой значений $R=\{r_1,...,r_g\}$, который можно считать еще одним качественным признаком. Таким образом, имеется k версий (экземпляров) каждого объекта A_i, различающихся оценками по критериям $Q_1,...,Q_m$, и k индивидуальных экспертных правил сортировки этого объекта

$$\text{IF } \langle\text{условия}\rangle, \text{ THEN } \langle\text{решение}\rangle, \tag{1}$$

обычно не согласованных между собой. Здесь терм ⟨условия⟩ включает совокупность значений признаков $x_{i1}^{e_1(f)},...,x_{im}^{e_m(f)}$; $r_{il}^{(f)}$, $x_{is}^{e_s(f)}\in X_s$, $r_{il}^{(f)}\in R$; терм ⟨решение⟩ указывает на принадлежность версии $A_i^{(f)}$ объекта A_i к классу C_l. Требуется найти одно или несколько достаточно простых обобщенных правил групповой классификации вида (1), которые в наибольшей степени совпадают с несогласованными индивидуальными правилами экспертной классификации объектов, позволяют отнести объекты к заданным классам, не отвергая возможную противоречивость оценок, и выявить противоречиво классифицированные объекты.

Многопризнаковый объект A_i обычно принято представлять в пространстве $X=X_1\times...\times X_m$ как вектор или кортеж $x_i=(x_{i1}^{e_1},...,x_{im}^{e_m})$, составленный из количественных (числовых) или качественных (символьных, вербальных) значений признаков [3, 6, 9-11]. Качественные переменные нередко и не всегда обоснованно трансформируются в числовые, например, используя лексикографические шкалы или функции нечеткой принадлежности [12, 19]. Ситуация существенным образом усложняется, если одному и тому же объекту A_i будет соответствовать не один, а несколько m-мерных кортежей. В таких случаях объект A_i представляется в пространстве X группой, состоящей из k кортежей $\{x_i^{(1)},...,x_i^{(k)}\}$ вида $x_i^{(f)}=(x_{i1}^{e_1(f)},...,x_{im}^{e_m(f)})$, $f=1,...,k$, которая должна рассматриваться как единое целое. Значения признаков могут быть похожими, различающимися и даже противоречивыми, что может приводить к несравнимости m-мерных кортежей $x_i^{(f)}=(x_{i1}^{e_1(f)},...,x_{im}^{e_m(f)})$, характеризующих один и тот же объект A_i. Более того, если разные объекты или разные версии одного и того же объекта имеют одинаковые значения признаков, но при их классификации они отнесены экспертами к различным классам, то в пространстве X невозможно даже построить разделяющие классы гиперповерхности. Поэтому анализ совокупностей таких многомерных объектов в пространстве X достаточно труден. Эти трудности можно преодолеть, если воспользоваться иным представлением многопризнаковых объектов.

Введем вместо прямого произведения m шкал признаков $X=X_1\times...\times X_m$ множество $X=X_1\cup...\cup X_m$ – обобщенную шкалу признаков, состоящую из m групп признаков, и представим объект в таком символическом виде:

$$A_i = \{k_{Ai}(x_1^1)\circ x_1^1,...,k_{Ai}(x_1^{h_1})\circ x_1^{h_1},..., k_{Ai}(x_m^1)\circ x_m^1,..., k_{Ai}(x_m^{h_m})\circ x_m^{h_m}\}, \tag{2}$$

где число $k_{Ai}(x_s^{e_s})$ указывает, сколько раз признак $x_s^{e_s}\in X_s$ встречается в описании объекта A_i, а знак $\circ$ обозначает кратность признака $x_s^{e_s}$. Множество X характеризует свойства совокупности объектов $A=\{A_1,...,A_n\}$. Такая запись объекта A_i представляет его как мультимножество A_i, порожденное множеством X, которое задано функцией кратности k_A:$X\to Z_+=\{0,1,2,3,...\}$. Мультимножество A_i конечно, когда конечны все $k_{Ai}(x_s^{e_s})$. Мультимножества A и B равны ($A=B$), если $k_A(x_s^{e_s})=k_B(x_s^{e_s})$. Мультимножество B включено в мультимножество A ($B\subseteq A$), если $k_B(x_s^{e_s})\leq k_A(x_s^{e_s})$, $\forall x_s^{e_s}\in X$ [4].

Введены следующие операции над мультимножествами: объединение $A\bigcup B$, пересечение $A\bigcap B$, сложение $A+B$, вычитание $A-B$, симметрическая разность $A\Delta B$, умножение на число $b\bullet A$, $b_i>0$, умножение $A\bullet B$, n-ая степень A^n, прямое произведение $A\times B$, прямая n-ая степень $(\times A)^n$. Новые типы операций над мультимножествами открывают новые возможности для агрегирования многопризнаковых объектов. Например, группа X_a объектов может быть получена как сумма $X_a=\sum_i A_i$, объединение $X_a=\bigcup_i A_i$ или пересечение $X_a=\bigcap_i A_i$ мультимножеств A_i, описывающих объекты A_i, либо как линейная комбинация мультимножеств вида $X_a=\sum_i b_i\bullet A_i$, $X_a=\bigcup_i b_i\bullet A_i$ или $X_a=\bigcap_i b_i\bullet A_i$.

На семействе мультимножеств $A=\{A_1,...,A_n\}$ можно определить новые классы метрических пространств мультимножеств (A,d), которые задаются метриками:

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

где p – целое число, m – мера мультимножества, заданная на алгебре мультимножеств $L(Z)$, Z – максимальное мультимножество с $k_Z(x)=\max_{A\in A}k_A(x)$. Метрика $d_{1p}(A,B)$ называется основной и является метрикой типа Хемминга, традиционно используемой во многих приложениях. Метрика $d_{2p}(A,B)$ называется полностью усредненной и характеризует различие между двумя мультимножествами A и B, отнесенное к расстоянию, максимально возможному в исходном пространстве. Метрика $d_{3p}(A,B)$ называется локально усредненной и задает различие, отнесенное к максимально возможной «общей части» $A\bigcup B$ только этих двух мультимножеств в исходном пространстве. При $p=1$ в случае обычных множеств $d_{11}(A,B)=m(A\Delta B)$ называют расстоянием Фреше, а $d_{31}(A,B)=m(A\Delta B)/m(A\bigcup B)$ – расстоянием Штейнхауса [1].

Меру мультимножества можно определить различными способами, например, как линейную комбинацию функций кратности: $m(A)=\sum_s w_s k_A(x_s^{e_s})$, $w_i>0$. В этом случае, например, основная метрика имеет при $p=1$ следующий вид

$$d_{11}(A,B)=m(A\Delta B)=\sum_{s=1}^{m} w_s \sum_{e_s=1}^{h_s}|k_A(x_s^{e_s})-k_B(x_s^{e_s})|.$$

Различные свойства мультимножеств и метрических пространств мультимножеств рассматриваются и обсуждаются в [4].

3. Групповое упорядочение многопризнаковых объектов

Опишем метод группового упорядочения многопризнаковых объектов АРАМИС (Агрегирование и Ранжирование Альтернатив относительно Многопризнаковых Идеальных Ситуаций). Метод предназначен для ранжирования объектов, оценённых несколькими экспертами по многим качественным критериям, без построения индивидуальных ранжировок объектов [6].

Будем считать многопризнаковые объекты точками некоторого метрического пространства мультимножеств (A,d), например, с основной метрикой d_{11}. В пространстве (A,d) можно выделить два объекта (возможно, гипотетических) – абсолютно лучший и абсолютно худший, которые имеют соответственно самые лучшие и самые худшие оценки по всем критериям. Этим объектам соответствуют мультимножества

$$A_{\max}=\{k\circ x_1^1,0,...,0,\ k\circ x_2^1,0,...,0,...,\ k\circ x_m^1,0,...,0\},$$
$$A_{\min}=\{0,...,0,k\circ x_1^{h_1},\ 0,...,0,k\circ x_2^{h_2},...,\ 0,...,0,k\circ x_m^{h_m}\},$$

которые принято называть идеальным и антиидеальным решениями. Таким образом, задача упорядочения многопризнаковых объектов сводится к задаче упорядочения мультимножеств.

Упорядочим все объекты по величине их расстояния от идеального решения $A_{\max}$. Если $d_{11}(A_{\max},A_i)<d_{11}(A_{\max}, A_j)$, то объект A_i предпочтительнее объекта A_j ($A_i\succ A_j$). Если для некоторых объектов $d_{11}(A_{\max},A_i)=d_{11}(A_{\max},A_j)$, то объекты A_i и A_j будут или эк-

вивалентными, или несравнимыми. Тем самым получаемое ранжирование объектов окажется нестрогим.

Так как каждый объект A_i оценивается k экспертами по всем m критериям, то выражение для расстояния от идеального решения $A_{\max}$ до мультимножества A_i можно записать как

$$d_{11}(A_{\max}, A_i) = 2\sum_{s=1}^{m} w_s\, [k - k_{Ai}(x_s^1)].$$

Условие сравнения многопризнаковых объектов приобретает тогда следующую форму: объект A_i предпочтительнее объекта A_j ($A_i \succ A_j$), если

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

Тем самым правило упорядочения многопризнаковых объектов сводится к сравнению взвешенных сумм $S_{Ai}^1=\sum_s w_s k_{Ai}(x_s^1)$ первых (наилучших) оценок объектов по всем критериям Q_s. Лучшим будет тот объект A_i, у которого эта сумма S_{Ai}^1 будет больше. Если найдутся группы эквивалентных или несравнимых объектов $A_{i1},\ldots,A_{it}$, имеющих одинаковые суммы S_{Ai}^1, то нужно вычислить для каждого объекта A_{ir}, $r=1,\ldots t$ в соответствующей группе взвешенную сумму $S_{Air}^2=\sum_s w_s k_{Air}(x_s^2)$ 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 $A_{iru},\ldots,A_{irv}$, имеющих одинаковые суммы S_{Air}^2, то вычислить для каждого объекта A_{irp} в соответствующей подгруппе взвешенную сумму $S_{Airp}^3=\sum_s w_s k_{Airp}(x_s^3)$ 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 $A=\{A_1,\ldots,A_n\}$. Если число h_s значений оценок $x_s^{e_s}$ у некоторых критериев Q_s окажется меньше требуемого на данном b-ом шаге алгоритма, то следует считать $k_{Air\ldots p}(x_s^b)=0$. В случае, когда все критерии считаются одинаково важными, все коэффициенты w_s полагаются равными 1.

Аналогично строится процедура упорядочения многопризнаковых объектов A_i по отношению к антиидеальному решению $A_{\min}$, считая, что объект A_i лучше объекта A_j ($A_i \succ A_j$), если мультимножество A_i находится дальше от антиидеального решения $A_{\min}$, то есть если выполняется условие $d_{11}(A_{\min},A_i)>d_{11}(A_{\min},A_j)$. Подчеркнем, что упорядочение совокупности многопризнаковых объектов $A=\{A_1,\ldots,A_n\}$ по отношению к антиидеальному решению может не совпадать с упорядочением по отношению к идеальному решению.

4. Согласованная групповая классификация многопризнаковых объектов

Изложим теперь метод групповой классификации многопризнаковых объектов МАСКА (МногоАспектная Согласованная Классификация Альтернатив). Метод предназначен для построения согласованных правил классификации объектов, оценённых несколькими экспертами по многим качественным критериям, которые агрегируют противоречивые индивидуальные правила сортировки объектов и выявляют противоречиво классифицированные объекты [14].

Введем иную гипершкалу признаков – множество $X'=X_1\cup\ldots\cup X_m\cup R$, которое объединяет подмножества критериальных оценок X_s и подмножество сортирующих признаков R. Сопоставим многопризнаковому объекту A_i мультимножество:

$$A_i=\{k_{Ai}(x_1^1)\circ x_1^1,\ldots,k_{Ai}(x_1^{h_1})\circ x_1^{h_1},\ldots, k_{Ai}(x_m^1)\circ x_m^1,\ldots,k_{Ai}(x_m^{h_m})\circ x_m^{h_m}, k_{Ai}(r_1)\circ r_1,\ldots,k_{Ai}(r_g)\circ r_g\}, \qquad (4)$$

порожденное множеством X'. Значения функции кратности $k_{Ai}(x_s^{e_s})$ и $k_{Ai}(r_l)$ мультимножества A_i равны соответственно числам экспертной оценки $x_s^{e_s}$ объекта A_i по критерию Q_s и заключения r_l.

Мультимножество (4) можно интерпретировать как коллективное правило классификации вида (1), объединяющее k индивидуальных правил сортировки объекта A_i. В этом правиле терм ⟨условия⟩ включает различные комбинации оценок по критериям $x_s^{e_s}$ и признаков r_l, а терм ⟨решение⟩ обозначает принадлежность объекта A_i к некоторому классу C_l. Например, объект A_i относится к классу C_l согласно одному из следующих правил «большинства голосов»: если $k_{Ai}(r_l)>k_{Ai}(r_p)$ для всех $p\neq l$, или если $k_{Ai}(r_l)>\sum_{p\neq l} k_{Ai}(r_p)$. Очевидно, что оценки и заключения многих экспертов могут быть схожими, различающимися и противоречивыми. Эти расхождения не являются случайными ошибками, а выражают субъективность предпочтений экспертов.

Для упрощения задачи предположим, что совокупность объектов $A=\{A_1,\ldots,A_n\}$ нужно разделить только на два класса C_a (более предпочтительные) и C_b (менее предпочтительные). Разбиение именно 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, если требуется выделить несколько групп конкурсных проектов, то сначала заявки можно разделить на принятые и отложенные, затем принятые заявки – на безусловно и условно поддержанные, а отложенные заявки – на возможно поддержанные и неподдержанные, и так далее.

Сопоставим каждому классу объектов мультимножество, которое образуем, суммируя мультимножества (4), которые представляют многопризнаковые объекты, принадлежащие к этому классу по принятому правилу большинства голосов. Тогда классу C_t, $t=a,b$ соответствует мультимножество:

$$Y_t=\{k_{Yt}(x_1^1)\circ x_1^1,\ldots,k_{Yt}(x_1^{h_1})\circ x_1^{h_1},\ldots, k_{Yt}(x_m^1)\circ x_m^1,\ldots,k_{Yt}(x_m^{h_m})\circ x_m^{h_m}, k_{Yt}(r_1)\circ r_1,\ldots,k_{Yt}(r_g)\circ r_g\}, \quad (5)$$

где $k_{Yt}(x_s^{e_s})=\sum_{i\in It}k_{Ai}(x_s^{e_s})$, $k_{Yt}(r_l)=\sum_{i\in It}k_{Ai}(r_l)$, $I_a\bigcup I_b=\{1,\ldots,n\}$, $I_a\bigcap I_b=\varnothing$. Мультимножество (5) можно считать групповым правилом классификации объектов вида (1), объединяющим индивидуальные правила всех экспертов по сортировке совокупности объектов $A=\{A_1,\ldots,A_n\}$ в класс C_t.

Сконструируем новые объекты, представленные мультимножествами

$$R_l=\{k_{Rl}(y_a)\circ y_a,\ k_{Rl}(y_b)\circ y_b\},\quad Q_j=\{k_{Qj}(y_a)\circ y_a,\ k_{Qj}(y_b)\circ y_b\}$$

над множеством новых признаков $Y=\{y_a,y_b\}$, элементы которого характеризуют принадлежность объектов к классам C_a и C_b. Здесь $k_{Rl}(y_t)=k_{Yt}(r_l)$, $l=1,\ldots,g$, $k_{Qj}(y_t)=k_{Yt}(x_j)$, $j=1,\ldots,h$, $h=h_1+\ldots+h_m$. Назовем мультимножества R_l категориальными, а мультимножества Q_j – содержательными.

Нетрудно убедиться, что согласно правилу большинства голосов категориальные мультимножества формируют два мультимножества $R_a=\sum_{l\in La}R_l$ и $R_b=\sum_{l\in Lb}R_l$, где $L_a\bigcup L_b=\{1,\ldots,g\}$, $L_a\bigcap L_b=\varnothing$, расстояние $d^*=d(R_a,R_b)$ между которыми в метрическом пространстве (B,d) новых мультимножеств будет наибольшим расстоянием между объектами, входящими в разные классы. $B=\{R_1,..,R_g, Q_1,..,Q_h\}$, d – одна из метрик (3).

Категориальные мультимножества R_a и R_b определяют декомпозицию рассматриваемой совокупности объектов A на классы C_a и C_b, лучшую из всех возможных для заданных индивидуальных экспертных правил сортировки:

IF ⟨$(\sum_{l\in La}k_{Ai}(r_l)>\sum_{l\in Lb}k_{Ai}(r_l))$⟩, THEN ⟨Объект $A_i\in C_a$⟩,

IF ⟨$(\sum_{l\in La}k_{Ai}(r_l)<\sum_{l\in Lb}k_{Ai}(r_l))$⟩, THEN ⟨Объект $A_i\in C_b$⟩. (6)

При отсутствии противоречий между индивидуальными правилами экспертов по предварительной сортировке объектов, расстояние d^* будет равно соответственно $d_{11}^*=kn$, $d_{12}^*=1/(h+g)$, $d_{13}^*=1$.

Основная идея нахождения простого обобщенного правила групповой классификации многопризнаковых объектов, наилучшим образом агрегирующего несогласованные индивидуальные правила сортировки многих экспертов, формулируется так. Для каждой s-ой группы признаков Q_s нужно построить содержательные мультимножества Q_{sa}^* и Q_{sb}^*, которые будут определять s-ую лучшую декомпозицию совокупности объектов A на классы C_a и C_b.

По аналогии с категориальными мультимножествами R_a и R_b можно сформировать содержательные мультимножества Q_{sa} и Q_{sb}, которые образуются как суммы мультимножеств: $Q_{sa}=\sum_{j\in Jsa}Q_j$ и $Q_{sb}=\sum_{j\in Jsb}Q_j$, где $J_{sa}\bigcup J_{sb}=\{1,\ldots,h_s\}$, $J_{sa}\bigcap J_{sb}=\varnothing$, и задают некоторую декомпозицию объектов на классы C_a и C_b. Каждое мультимножество Q_{st} соответствует некоторой подгруппе значений признаков $Q_{st}=\bigcup_{j\in J_{st}}Q_j$, $t=a,b$.

Рассматривая содержательные мультимножества как точки в пространстве мультимножеств (B,d), требуется построить для всех признаков такие пары мультимножеств Q_{sa}^* и Q_{sb}^*, которые находились бы на максимально возможном расстоянии. Мультимножества Q_{sa}^* и Q_{sb}^* являются решением следующей оптимизационной задачи:

$$d(Q_{sa},Q_{sb}) \to \max d(Q_{sa},Q_{sb}) = d(Q_{sa}^*,Q_{sb}^*). \quad (7)$$

Значения признаков $x_j\in Q_{st}^*$, получаемые из решения m оптимизационных задач (7) и характеризующие класс C_t, будем называть классифицирующими признаками для этого класса. Классифицирующие признаки можно упорядочить по величине расстояния $d(Q_{sa}^*,Q_{sb}^*)$ или точности аппроксимации $V_s=d(Q_{sa}^*,Q_{sb}^*)/d(R_a,R_b)$. Классифицирующий признак $x_j\in Q_{st}^*$, обеспечивающий приемлемую точность аппроксимации $V_s\geq V_0$, включается в обобщенное решающее правило для групповой сортировки многопризнаковых объектов. Значение точности аппроксимации V_s характеризует относительную значимость s-ой группы признаков Q_s в обобщенном правиле классификации.

Разные комбинации классифицирующих признаков дают следующие обобщенные решающие правила для групповой сортировки объектов на классы C_a и C_b:

$$\text{IF } \langle(x_j\in Q_{ua}^*) \text{ AND } (x_j\in Q_{va}^*) \text{ AND } (x_j\in Q_{wa}^*) \text{ AND}\ldots\rangle, \text{ THEN } \langle\text{Объект } A_i\in C_a\rangle, \quad (8)$$

$$\text{IF } \langle(x_j\in Q_{ub}^*) \text{ AND } (x_j\in Q_{vb}^*) \text{ AND } (x_j\in Q_{wb}^*) \text{ AND}\ldots\rangle, \text{ THEN } \langle\text{Объект } A_i\in C_b\rangle. \quad (9)$$

При одновременном учете экспертных правил сортировки (6) и обобщенного решающего правила (8) или (9) среди отобранных окажутся как правильно, так и неправильно классифицированные объекты. Необходимо найти такие классифицирующие признаки, которые обеспечат максимальные числа N_a и N_b правильно классифицированных объектов и минимальные числа N_{ac} и N_{bc} неправильно классифицированных объектов.

Начнем с обобщенного решающего правила (8). Найдем сначала критерий Q_u^*, обеспечивающий максимальную разность N_a-N_{ac} между числами правильно и неправильно классифицированных объектов. Затем рассмотрим все комбинации критерия Q_u^* с остальными критериями и выберем пару критериев Q_u^*, Q_v^*, обеспечивающих максимальную разность N_a-N_{ac}. Далее, рассмотрев все комбинации пары критериев Q_u^*, Q_v^* с остальными критериями, выберем тройку Q_u^*,Q_v^*,Q_w^*, затем четверку и т.д. критериев, обеспечивающих максимальную разность N_a-N_{ac}. Последовательно шаг за шагом уменьшая разность N_a-N_{ac} между числами правильно и неправильно классифицированных объектов, найдем критерии, которые включаются в согласованное решающее правило для построения уточненного класса $C_a\backslash C_{ac}$ безусловно предпочтительных объектов:

$$\text{IF } \langle(\textstyle\sum_{x\in Qua^*}k_{Ai}(x_j)>\sum_{x\in Qub^*}k_{Ai}(x_j)) \text{ AND } (\sum_{x\in Qva^*}k_{Ai}(x_j)>\sum_{x\in Qvb^*}k_{Ai}(x_j)) \text{ AND}\ldots$$
$$\text{AND } (\textstyle\sum_{r\in Ra}k_{Ai}(r_l)>\sum_{r\in Rb}k_{Ai}(r_l))\rangle, \text{ THEN } \langle\text{Объект } A_i\in C_a\backslash C_{ac}\rangle, \quad (10)$$

Аналогичным образом получается согласованное решающее правило для построения уточненного класса $C_b\backslash C_{bc}$ безусловно непредпочтительных объектов:

$$\text{IF } \langle(\textstyle\sum_{x\in Qua^*}k_{Ai}(x_j)<\sum_{x\in Qub^*}k_{Ai}(x_j)) \text{ AND } (\sum_{x\in Qva^*}k_{Ai}(x_j)<\sum_{x\in Qvb^*}k_{Ai}(x_j)) \text{ AND}\ldots$$
$$\text{AND } (\textstyle\sum_{r\in Ra}k_{Ai}(r_l)<\sum_{r\in Rb}k_{Ai}(r_l))\rangle, \text{ THEN } \langle\text{Объект } A_i\in C_b\backslash C_{bc}\rangle, \quad (11)$$

с помощью которого минимизируется разность N_b-N_{bc} между числами правильно и неправильно классифицированных объектов.

Согласованные групповые правила (10) и (11) наилучшим образом аппроксимируют семейство индивидуальных правил сортировки многих экспертов и, вообще говоря, различны.

Одновременно формируется класс $C_c=C_{ac}\bigcup C_{bc}$ противоречиво классифицированных объектов, которые удовлетворяют следующему решающему правилу:

$$\text{IF } \langle[(\textstyle\sum_{x\in Qua^*}k_{Ai}(x_j)>\sum_{x\in Qub^*}k_{Ai}(x_j)) \text{ AND } (\sum_{x\in Qva^*}k_{Ai}(x_j)>\sum_{x\in Qvb^*}k_{Ai}(x_j)) \text{ AND}\ldots$$

AND $(\sum_{r\in Ra} k_{Ai}(r_l) < \sum_{r\in Rb} k_{Ai}(r_l))]$ OR $[(\sum_{x\in Qua^*} k_{Ai}(x_j) < \sum_{x\in Qub^*} k_{Ai}(x_j))$ AND $(\sum_{x\in Qva^*} k_{Ai}(x_j) < \sum_{x\in Qvb^*} k_{Ai}(x_j))$ AND…AND $(\sum_{r\in Ra} k_{Ai}(r_l) > \sum_{r\in Rb} k_{Ai}(r_l))]\rangle$,

THEN ⟨Объект $A_i \in C_c$⟩. (12)

Эти объекты нуждаются в дополнительном анализе. Правило (12) помогает обнаружить возможные противоречия в индивидуальных экспертных правилах сортировки.

Заключение

Задачи групповой классификации и упорядочения объектов, которые характеризуются многими разнородными признаками, причем каждый из объектов может существовать в нескольких различающихся версиях, являются достаточно трудными для решения и до настоящего момента слабо изучены. Трудности имеют и содержательные основания (например, необходимость рассмотрения нескольких версий объектов), и формальные причины (например, некорректность процедуры «усреднения» качественных признаков). Использование аппарата теории мультимножеств при решении практических задач принятия решений, анализа данных, распознавания образов и в других областях позволяет ввести новые виды операций, создать новые методы обработки разнородной информации, не приводящие к потере или искажению данных.

Разработанные подходы были использованы для построения рейтинга российских компаний, работающих в секторе информационно-коммуникационных технологий [2], для оценки кредитоспособности держателей кредитных карт [5], при анализе конкурсов целевых фундаментальных исследований, выполняемых по грантам Российского фонда фундаментальных исследований [7]. Решающие правила (8)-(12) для групповой сортировки объектов можно легко записать на естественном языке, используя формулировки вербальных значений классифицирующих признаков. Например, согласованное групповое правило для конкурсного отбора проектов целевых фундаментальных исследований выглядит следующим образом [7]:

«Проект безусловно поддерживается, если он имеет исключительно высокую или значительную научную значимость; массовый или междисциплинарный масштабы применимости результатов исследований; предлагаемые в проекте исследования завершаются лабораторным образцом или ключевыми элементами разработки».

В работе изложены методы классификации и упорядочения многопризнаковых объектов, основанные на их представлении с помощью мультимножеств, которые применимы и в случае несогласованности и противоречивости описаний объектов. Использование мультимножеств позволяет значительно расширить круг рассматриваемых проблем и решать новые разнообразные задачи классификации, сортировки, ранжирования многопризнаковых объектов.

Литература

1. Деза М.М., Лоран М. Геометрия разрезов и метрик. – М.: МЦНМО, 2001.
2. Кто в России самый интеллектуальный? Рейтинг ведущих российских разработчиков высоких технологий.//Компания. – 2000. – №47(143) – С.38-39.
3. Ларичев О.И. Вербальный анализ решений. – М.: Наука, 2006.
4. Петровский А.Б. Пространства множеств и мультимножеств. – М: Едиториал УРСС, 2003.
5. Петровский А.Б. Модель оценки кредитоспособности владельцев кредитных карт по противоречивыми данным. // Искусственный интеллект. – 2004. – №2. – С.155-161.
6. Петровский А.Б. Теория принятия решений. – М.: Издательский центр «Академия», 2009.
7. Петровский А.Б., Тихонов И.П., Балышев А.В., Комарова Н.А. Построение согласованных групповых правил для конкурсного отбора научных проектов // Труды третьей межд. конф. "Системный анализ и информационные технологии". – М.: ПолиПринтСервис, 2009. – С.337-348.
8. Т.Саати Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993.

9. Doumpos M., Zopounidis C. Multicriteria Decision Aid Classification Methods. – Dordrecht: Kluwer Academic Publishers, 2002.
10. Furems E. Knowledge-based multi-attribute classification problems structuring. // Ruan D., Montero J., Lu J., Martinez L., D'hondt P., Kerre E. (eds.) Computational Intelligence in Decision And Control. – Singapore: World Scientific Publisher, 2008. – P.465-470.
11. Greco S., Matarazzo B., Slowinski R. Rough sets methodology for sorting problems in presence of multiple attributes and criteria. // European Journal of Operational Research. – 2002. – Vol.138. – N2. – P.247-259.
12. Hwang C.L., Lin M.J. Group Decision Making under Multiple Criteria. – Berlin: Springer-Verlag, 1987.
13. Koksalan M., Ulu C. An interactive approach for placing alternatives in preference classes. // European Journal of Operational Research. – 2003. – Vol.144. – N2. – P.429-439.
14. Petrovsky A.B. Multi-attribute sorting of qualitative objects in multiset spaces. // Multiple Criteria Decision Making in the New Millenium. – Berlin: Springer-Verlag, 2001. — P.124-131.
15. Roubens M. Ordinal multiattribute sorting and ordering in the presence of interacting points of view. // Bouyssou D., Jacquet-Lagreze E., Perny P., Slowinski R., Vanderpooten D., Vincke P. (eds.) Aiding Decisions with Multiple Criteria: Essays in Honor of Bernard Roy. – Dordrecht: Kluwer Academic Publishers, 2001. – P.229-246.
16. Roy B. Multicriteria Methodology for Decision Aiding. – Dordrecht: Kluwer Academic Publishers, 1996.
17. Roy B., Bouyssou D. Aide Multicritere a la Decision: Methodes et Cas. – Paris: Economica, 1993.
18. Vincke P. Multicriteria Decision Aid. – Chichester: Wiley, 1992.
19. Zadeh L.A. From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions. // IEEE Transactions on Circuits and Systems. – 1999. – Vol.45. – N1. – P.105-119.

GROUP MULTICRITERIAL DECISION MAKING WITH INCONSISTENT PREFERENSES

A. B. PETROVSKY

Institute for System Analysis, Russian Academy of Sciences

e-mail: pab@isa.ru

The paper considers a new approach to group multicriterial decision making, where preferences of several decision makers are inconsistent and possibly contradictory, and alternatives are estimated by manifold criteria. The approach is based on the methodology of verbal decision analysis and the theory of multisets. There are examples of the approach application to solve practical problems of classifying and ordering multi-attribute objects.

Key words: group decision making, theory of multisets, classifying, ordering, multi-attribute objects