

СИСТЕМНЫЙ АНАЛИЗ МОДЕЛЕЙ КЛАССИФИКАЦИИ И УПОРЯДОЧЕНИЯ ОБЪЕКТОВ

С. В. МИКОНИ

*Санкт-Петербургский федеральный исследовательский центр Российской академии наук,
Санкт-Петербург, Россия
smikoni@mail.ru*

Аннотация. Рассматривается проблема объединения моделей классификации и упорядочения объектов в рамках одной системы поддержки принятия решений. Проблема решается путем поиска сходства и различия соответствующих моделей с применением методов системного анализа. Рассмотрены связи между моделями классификации с четкими и нечеткими границами классов, моделями упорядочения объектов с жесткими и мягкими ограничениями к значениям показателей. Введено понятие реальной цели, позволяющее связать между собой методы многоцелевой и условной оптимизации, а их, в свою очередь, с методами многомерной теории ценности и полезности. Трактующаяся как граница между классами реальная цель позволяет связать методы классифицирования и упорядочения объектов. Найденные связи являются основанием для объединения рассмотренных методов в рамках общей системы поддержки принятия решений, что дает большие возможности для принятия решений разными методами и сопоставления получаемых результатов.

Ключевые слова: классификация, класс, упорядочение, выбор, цель, оценочная функция, система поддержки принятия решений

Благодарности: работа выполнена при финансовой поддержке бюджетной темы FFZF-2022-0004; автор выражает благодарность канд. техн. наук Д. П. Буракову.

Ссылка для цитирования: Микони С. В. Системный анализ моделей классификации и упорядочения объектов // Изв. вузов. Приборостроение. 2022. Т. 65, № 11. С. 796—801. DOI: 10.17586/0021-3454-2022-65-11-796-801.

SYSTEM ANALYSIS OF CLASSIFICATION AND ORDERING MODELS OF OBJECTS

S. V. Mikoni

*St. Petersburg Federal Research Center of the RAS,
St. Petersburg, Russia
smikoni@mail.ru*

Abstract. The problem of combining models of classification and ordering of objects within one decision support system is considered. The problem is solved by searching for similarities and differences of the corresponding models using the system analysis methods. Relations between classification models with clear and fuzzy class boundaries, models of ordering objects with hard and soft restrictions on the indicators' values are considered. The concept of a real goal is introduced, which makes it possible to link the methods of multi-objective and conditional optimization, and these, in turn, with the methods of the multidimensional theory of value and utility. Interpreted as a boundary between classes, the real goal allows to link the methods of classifying and emphasizing objects. The revealed connections are the basis for combining the considered methods within the framework of a common decision support system, which gives great opportunities for making decisions using different methods and comparing the results obtained.

Keywords: classification, class, ordering, choice, goal, evaluation function, decision support system

Acknowledgments: the work was carried out with the financial support of the budget theme FFZF-2022-0004; the author expresses his gratitude to D. P. Burakov, PhD.

For citation: Mikoni S. V. System analysis of classification and ordering models of objects. *Journal of Instrument Engineering*. 2022. Vol. 65, N 11. P. 796—801 (in Russian). DOI: 10.17586/0021-3454-2022-65-11-796-801.

Введение. Конечной целью принятия решения является либо выбор класса, к которому принадлежит оцениваемый объект, либо выбор предпочтительного объекта из класса на основе отношения порядка. Эти задачи обычно решаются порознь и автоматизируются в рамках различных программных систем [1—4]. Более того, даже методы одного назначения реализуются в рамках разных систем поддержки принятия решений (СППР). Однако то, что оправдано для специализированных систем, не может претендовать на универсальность. Универсальность системы, моделирующей принятие решения, востребована для оценивания объектов разными методами. Помимо этого, существует потребность объединить методы классификации и упорядочения объектов в рамках одной системы, опираясь на общность свойств объекта, подлежащих оцениванию в обеих группах методов. Для решения такой задачи необходимо найти связи между методами каждой из групп и между группами. В настоящей статье предлагается подход к решению этой задачи.

Классификация. По мнению Ж. Б. Ламарка, „всякая наука начинается с классификации“, а по мнению Д. И. Менделеева, „наука начинается там, где начинают измерять“. Кажущееся противоречие этих высказываний о начале науки исключается, если обобщить измерения на номинальную шкалу. Именно в номинальной шкале, т.е. именами, и измеряются классифицируемые понятия. Примером применения номинальной шкалы является присваивание объекту x некоторого имени C . Моделью такой шкалы является двухместный предикат *Присвоить* (x , C). Поскольку шкала именований менее информативна, чем порядковая шкала, применяемая при упорядочении объектов, начнем с рассмотрения моделей классификации.

Простейшая классификация — это деление множества объектов на классы относительно классификационных признаков с двумя противоположными значениями. Условием правильной n -мерной классификации является независимость классификационных признаков, их ортогональность в n -мерном пространстве. Задача n -мерной классификации решается либо последовательно, либо параллельно. Примером первого варианта является дихотомическая классификация. Параллельная классификация представляет собой набор независимых фасет с противоположными значениями признаков.

В настоящей работе рассматривается обратная задача классификации, т.е. не создание классов, а отнесение объекта x к одному из известных классов $h \in H$, $|H| = m$, $m > 1$, по значениям характеризующих его показателей. Под классом h_k будем понимать отрезок $[c_{jн}^k, c_{jв}^k]$, принадлежащий шкале $[y_{j, \min}, y_{j, \max}]$ j -го показателя, $j = \overline{1, n}$, $k = \overline{1, m}$.

Классификация с четкими границами классов. Условие четких границ между смежными классами h_k и h_{k+1} на шкале j -го показателя определяется как

$$[c_{jн}^k, c_{jв}^k] \cap [c_{jн}^{k+1}, c_{jв}^{k+1}] = \emptyset, \quad k = \overline{1, m-1}. \quad (1)$$

Объект x принадлежит к классу h , представленному отрезком $[c_{jн}, c_{jв}]$ с нижней и верхней границами $c_{jн}$ и $c_{jв}$ на шкале $[y_{j, \min}, y_{j, \max}]$ j -го показателя, $y_{j, \min} \leq c_{jн} \leq c_{jв} \leq y_{j, \max}$, если $y_j(x) \in [c_{jн}, c_{jв}]$. В префиксной (функциональной) записи это условие имеет вид: $\in(y_j(x), [c_{jн}, c_{jв}])$. Как логическая функция оно представляет собой двухместный предикат в значении „истина“, задающий логическое правило формирования класса h . Значению „ложь“ этого предиката соответствует выражение $\notin(y_j(x), [c_{jн}, c_{jв}])$.

Отсюда следует, что условиями отнесения объекта x по j -му показателю к одному из классов $h \in H$ являются:

- 1) наличие не менее двух классов $|H| = m > 1$;
- 2) задание интервала значений $[c_{jн}^k, c_{jв}^k]$ j -го показателя для каждого класса, $k = \overline{1, m}$;
- 3) применение отношения *соответствия*, реализуемого предикатом принадлежности значения к k -му отрезку шкалы:

$$\in \left(y_j(x), \left[c_{jн}^k, c_{jв}^k \right] \right), \quad k = \overline{1, m}. \quad (2)$$

Частный случай отнесения объекта x к одному из двух взаимоисключающих классов („норма/ не норма“) представляет собой задачу отбора.

Если классы $h_k \in H$, $k = \overline{1, m}$, характеризуются n показателями, то принадлежность к k -му классу представляется конъюнкцией двухместных предикатов по j :

$$\bigwedge_{j=1}^n \in \left(y_j(x), \left[c_{jн}^k, c_{jв}^k \right] \right), \quad k = \overline{1, m}. \quad (3)$$

Значение „истина“ логической функции (3) означает принадлежность объекта x к k -му классу по всем показателям. Значение „ложь“ означает либо отсутствие принадлежности, либо частичную принадлежность объекта x к k -му классу, если хотя бы один из предикатов принадлежности к k -му классу в формуле (3) является истинным. При r истинных предикатах, $0 < r < n$, оценкой частичной принадлежности к k -му классу является относительная величина r_k / n . В этом случае выбирается тот класс h^* , частичная принадлежность к которому принимает максимальное значение:

$$h^* = h_t, \quad t = \arg \max_k \frac{r_k}{n}. \quad (4)$$

При классифицировании нескольких объектов величина r_k/n может быть использована для их упорядочения по частичной принадлежности к k -му классу, $k = \overline{1, m}$.

Классификация с нечеткими границами классов. Нечеткая классификация следует из нарушения условия (1). Она характеризуется исходной частичной принадлежностью объекта более чем к одному классу. Этот случай моделируется бесконечнозначной логикой Л. Заде. С применением функции принадлежности $\mu_{jk}(x)$, $j = \overline{1, n}$, $k = \overline{1, m}$, с областью значений $[0, 1]$ появляется возможность установления отношения частичного порядка на множестве классов H по значениям функции $\mu_{jk}(x)$ в интервале $0 < \mu_{jk} < 1$. Выбирается тот класс h_j^* , для которого эта функция по j -му показателю принимает максимальное значение:

$$h_j^* = h_t, \quad t = \arg \max_k \mu_{jk}(x). \quad (5)$$

В выражении (5) функция $\mu_{jk}(x)$ играет роль целевой функции, применяемой в задачах оптимизации: $\mu_{jk}(x) \rightarrow \max$. Иными словами, задача классификации решается методом оптимизации. Для классифицирования объекта x по n показателям вычисляется средневзвешенная аддитивная свертка частных принадлежностей к каждому классу, где w_j — значимость j -го показателя:

$$\mu_k(x) = \sum_{j=1}^n w_j \mu_{jk}(x), \quad k = \overline{1, m}. \quad (6)$$

По n показателям выбирается класс h^* с максимальным значением μ_k :

$$h^* = h_t, \quad t = \arg \max_k \mu_k(x). \quad (7)$$

При классифицировании объектов из множества X функция $\mu_k(x)$ может использоваться для установления отношения частичного порядка на этом множестве по степени принадлежности к каждому классу [5—7]. Если классы имеют одинаковый порядок на шкале каждого показателя, то существует возможность сквозного упорядочения объектов из множества X на основе значений функции принадлежности к классам [8].

Для случая двух нечетких классов („норма/не норма“) может решаться задача мониторинга и ранней диагностики объекта по степени отклонения от нормы как каждого параметра, так и их совокупности [9].

Упорядочение объектов. Установление отношения порядка на множестве объектов X выполняется при следующих условиях:

- 1) наличие не менее двух объектов $|X| = N > 1$;
- 2) точечное задание целевого значения c_j на шкале каждого j -го показателя;
- 3) применение отношения *превосходства*, реализуемого предикатом:

$$\succ(y_j(x), c_j). \quad (8)$$

Отношение строгого превосходства реализуется арифметическими предикатами „больше“ ($>$) или „меньше“ ($<$). Им соответствуют *целевые* значения j -го показателя, совпадающие с одной из границ шкалы $[y_{j,\min}, y_{j,\max}]$: $c_j = y_{j,\min}$ или $c_j = y_{j,\max}$. В [10] такое целевое значение названо *идеальной* целью, поскольку оно определяет размах выборки объектов либо теоретическую границу показателя. Если все компоненты вектора целей $\mathbf{c} = (c_1, \dots, c_j, \dots, c_n)$ являются идеальными целями, осуществляется упорядочение объектов относительно идеальной цели. Степень ее достижения определяется средневзвешенной оценкой объекта по всем показателям. Этот подход к упорядочению объектов реализуется методом многоцелевой оптимизации (Multiobjective Optimization) [11]. В случае когда объекты несравнимы по ресурсам, для каждого из них формируется свой вектор целей $\mathbf{c}_i = (c_{i1}, \dots, c_{ij}, \dots, c_{in})$, $i = \overline{1, N}$, и они упорядочиваются по процентам достижения своих *индивидуальных* целей [12].

В более общем случае сочетаются отношения строгого и нестрогого превосходства. Второе из них реализуется арифметическими предикатами „не менее“ ($\geq$) или „не более“ ($\leq$). Если они используются в роли ограничений, не удовлетворяющие им объекты исключаются из процесса оценивания, а остальные упорядочиваются относительно идеальной цели. Эта задача решается методом условной оптимизации (Multiobjective Optimization with Constraints) [13]. В частном случае при одной целевой функции имеет место классическая задача оптимизации.

Упорядочение объектов согласно мягким ограничениям. Это название такое упорядочение получило в силу того, что объект, не удовлетворяющий ограничению, не исключается, а подвергается упорядочению наравне с другими [14]. В дальнейшем этот подход был назван упорядочением по реальным целям. *Реальной* целью называется целевое значение c_j , не совпадающее с границей шкалы показателя: $y_{j,\min} < c_j < y_{j,\max}$. Такое название обосновывается согласованием целевого значения показателя с ресурсами, имеющимися для его достижения. Ценность этого показателя моделируется кусочно-линейной монотонной нормирующей функцией, обобщением которой является нелинейная функция ценности [15].

Принципиальной особенностью реальной цели является разбиение шкалы показателя на два участка — не удовлетворяющего и удовлетворяющего нестрогому ограничению. При использовании функции ценности с областью значений $[0, 1]$ значение показателя на первом участке может рассматриваться как невыполнение цели, а на втором — как перевыполнение. При использовании функции полезности с областью значений $[-1, 1]$ на ее отрицательной полуоси $[-1, 0)$ моделируется ущерб, обусловленный невыполнением цели, а на положительной полуоси $(0, 1]$ — доход от ее перевыполнения. В модели лотереи точка c_j на шкале показателя означает отказ от участия в ней с функцией полезности $u(c_j) = 0$ [16].

Таким образом, реальная цель является связующим звеном между методами многоцелевой (многокритериальной) оптимизации и методами теории многомерной ценности (Multi-Attribute Value Theory) и полезности (Multi-Attribute Utility Theory). Она позволяет также упорядочивать объекты по штрафам (по степени невыполнения цели) и поощрениям (по степени перевыполнения цели). С учетом изложенного введение понятия реальной цели расширяет разнообразие задач упорядочения объектов.

Заключение. Общность свойств, оцениваемых в задачах классификации и упорядочения объектов, является обоснованием инвариантности модели объекта по отношению к реализуемым методам классификации и упорядочения. Между собой модели классификации

и упорядочения различаются видом используемого отношения (соответствие/превосходство) и способом задания значений на шкале показателя (интервальное и точечное). Введение понятия реальной цели, делящей шкалу показателя на две части, связывает эти способы через полуинтервалы значений. Реальная цель также связывает через оценочные функции методы многокритериальной оптимизации с методами теории многомерной ценности и полезности. При этом имеет место определенная аналогия между частичной принадлежностью к классу в методах классификации и степенью достижимости цели в методах многомерной оптимизации. Установленные связи между методами классификации и упорядочения объектов как внутри групп методов, так и между группами позволяют объединять различные методы в рамках общей СППР для решения разнообразных практических задач.

СПИСОК ЛИТЕРАТУРЫ

1. Программное обеспечение для принятия решений [Электронный ресурс]: <<https://ru.wikipedia.org/wiki/>>.
2. *Figueira G., Amorim P., Guimar L., Amorim-Lopes M., Neves-Moreira F., Almada-Lobo B.* A decision support system for the operational production planning and scheduling of an integrated pulp and paper mill // *Computers & Chemical Engineering*. 2015. N 77. P. 85—104.
3. *Kar A. K.* A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network // *Journal of Computational Science*. 2015. N 6. P. 23—33.
4. *Kozhombardieva G., Burakov D. P., Khamchichev G. A.* Decision-Making Support Software Tools Based on Original Authoring Bayesian Probabilistic Models // *Journal of Physics Conference*. 2022. Ser. 2224(1). P. 1—8.
5. *Sun C. C.* A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS method // *Expert Syst. Appl.* 2010. N 37. P. 7745—7754.
6. *Prokopchina S. V.* Methodological foundations of scaling in modern Measurement Theory. Classification of measurement scales and their application under uncertainty based on Bayesian Intelligent Technologies // *Journal of Physics Conference*. 2020. Ser. 1703(1). P. 012003.
7. *Qiansheng Zhang, Dongfang Sun.* An Improved Decision-Making Approach Based on Interval-valued Fuzzy Soft Set // *Journal of Physics Conference*. 2021. Ser. 1828. P. 012041.
8. Микони С. В., Гарина М. И. Условие одинакового упорядочения объектов по функциям полезности и принадлежности // Тр. Конгресса IS&IT'11, Дивноморское, 3-10.09.2011. М.: Физматлит, 2011. Т. 1. С. 33—37.
9. Охтилев М. Ю., Соколов Б. В. Теоретические и прикладные проблемы разработки и применения автоматизированных систем мониторинга состояния сложных технических объектов // Тр. СПИИРАН. 2002. № 1(1). С. 167—180.
10. Микони С. В. Теория принятия управленческих решений: Учеб. пособие. СПб: Лань, 2015. 448 с.
11. *Velasquez M., Hester P. T.* An Analysis of Multi-Criteria Decision Making Methods // *Intern. Journal of Operations Research*. 2013. Vol. 10, N 2. P. 56—66.
12. Микони С. В., Тихомиров В. О., Тришанков В. В., Сорокина М. И. Определение рейтинга подразделений железной дороги по итогам выполнения планов // Тр. X конф. „Региональная информатика-2004“, 24-26.10.2006. СПб: СПОИСУ, 2006. С. 266—273.
13. *Пережудов Ф. И., Тарасенко Ф. П.* Введение в системный анализ. М.: Высш. школа, 1989. 367 с.
14. *Mikoni S. V.* Method of choice by approximation to a pattern // *Proc. Conf. NITE'2000*. Minsk: Belarus State Economic University, 2000. P.156—159.
15. *Keeney R. L., Raiffa H.* Decisions with Multiple Objectives: Preferences and Value Tradeoffs. N. Y.: Wiley, 1976.
16. *Neumann J. V., Morgenstern O.* Theory of Games and Economic Behavior. Princeton, NJ: Princeton Univ. Press, 1953.

Сведения об авторе

Станислав Витальевич Микони

— д-р техн. наук, профессор; СПбФИЦ РАН, СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; ведущий научный сотрудник; E-mail: smikoni@mail.ru

Поступила в редакцию 18.07.2022; одобрена после рецензирования 26.07.2022; принята к публикации 30.09.2022.

REFERENCES

1. <https://ru.wikipedia.org/wiki/>. (in Russ.)
2. Figueira G., Amorim P., Guimar L., Amorim-Lopes M., Neves-Moreira F., Almada-Lobo B. *Computers & Chemical Engineering*, 2015, vol. 77, pp. 85–104.
3. Kar A.K. *Journal of Computational Science*, 2015, no. 6, pp. 23–33.
4. Kozhombardieva G., Burakov D.P., Khamchichev G.A. *Journal of Physics Conference*, 2022, no. 1, series 2224, pp. 1–8.
5. Sun C.C. *Expert Syst. Appl.*, 2010, vol. 37, pp. 7745–7754.
6. Prokopchina S.V. *Journal of Physics Conference*, 2020, no. 1, series 1703, pp. 012003.
7. Qiansheng Zhang and Dongfang Sun. *Journal of Physics Conference*, 2021, series 1828, pp. 012041.
8. Mikoni S.V., Garina M.I. *Proceedings of the Congress IS&IT'11*, Divnomorskoye, 3–10.09 2011, Moscow, 2011. vol. 1, pp. 33–37. (in Russ.)
9. Okhtilev M.Yu., Sokolov B.V. *Trudy SPIIRAN* (SPIIRAS Proceedings), 2002, no. 1(1), pp. 167–180. (in Russ.)
10. Mikoni S.V. *Teoriya prinyatiya upravlencheskikh resheniy* (Theory of Managerial Decision Making), St. Petersburg, 2015, 448 p. (in Russ.)
11. Velasquez M., Hester P.T. *International Journal of Operations Research*, 2013, no. 2(10), pp. 56–66.
12. Mikoni S.V., Tikhomirov V.O., Trishankov V.V., Sorokina M.I. *Regional'naya informatika-2004* (Regional Informatics-2004), Proceedings of the X Conference, 24–26.10 2006, St. Petersburg, 2006, pp. 266–273. (in Russ.)
13. Peregudov F.I., Tarasenko F.P. *Vvedeniye v sistemnyy analiz* (Introduction to Systems Analysis), Moscow, 1989, 367 p. (in Russ.)
14. Mikoni S.V. *Proceedings of Conf. NITE'2000*, Minsk, 2000, pp.156–159.
15. Keeney R.L., Raiffa H. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, NY, Wiley, 1976.
16. Neumann J.V., Morgenstern O. *Theory of Games and Economic Behavior*, Princeton, NJ. Princeton University Press, 1953.

Stanislav V. Mikoni

Data on author

— Dr. Sci., Professor; St. Petersburg Federal Research Center of the RAS, St. Petersburg Institute for Informatics and Automation of the RAS, Laboratory of Information Technologies in System Analysis and Modeling; Leading Researcher; E-mail: smikoni@mail.ru

Received 18.07.2022; approved after reviewing 26.07.2022; accepted for publication 30.09.2022.