

СИСТЕМА ГРУПУВАННЯ ОБ'ЄКТІВ КЛАСТЕРАМИ РІЗНИХ ГЕОМЕТРИЧНИХ ФОРМ

КОНДРУК Н. Е.

кандидат технічних наук, доцент,

доцент кафедри кібернетики і прикладної математики

*Державний вищий навчальний заклад «Ужгородський національний
університет»*

м. Ужгород, Україна

Кластеризація є однією із потужних додаткових засобів в сфері аналізу даних, коли відсутня апріорна інформація про групування об'єктів. В даний час багато алгоритмів кластерного аналізу з успіхом використовуються в різних прикладних областях, де є необхідність розділити схожі за певними ознаками об'єкти на підмножини. Чіткий поділ на кластери можливий тільки при сильно відмінних ознаках об'єктів кластеризації. Тому для вирішення реальних завдань все частіше застосовуються нечіткі методи, в яких розбиття об'єктів проводиться із визначенням ступеню належності об'єктів кластерам.

На сьогодні математичний апарат кластерного аналізу бурхливо розвивається і забезпечує сучасні засоби ефективного розв'язання багатьох прикладних задач. Умовно, всі існуючі методи можна класифікувати за геометричною формою тих кластерів, які вони утворюють. Крім того, геометрична форма кластеру дозволяє отримати якісно різні прикладні інтерпретації отриманих однорідних груп. Тому саме специфіка прикладних задач робить неможливим автоматичне перенесення методів в іншу прикладну область без ризику свідомо отримати неякісний розв'язок. Таким чином, доцільним є розробка інформаційної системи, яка мала би досить широкий спектр засобів для групування об'єктів на кластери різних геометричних форм. Це дає можливість ефективно розв'язувати достатньо широкі класи прикладних

задач із різних предметних областей. Математичний апарат, що дозволяє реалізувати поставлену задачу описаний в [1-3].

Ядром системи є однорівневий метод кластеризації оснований на нечітких бінарних відношеннях описаний в [2]. Гнучкість даного алгоритму забезпечує зміну форми утворених кластерів шляхом лише однієї зміни виду міри подібності об'єктів. При цьому, подібність об'єктів O_i та O_j за деяким критерієм характеризується нечітким бінарним відношенням R на множині векторних ознак $C = \{\bar{c}_i | i = \overline{1, m}\}$ із функцією належності $\mu_R(\bar{c}_i, \bar{c}_j)$, де $\mu_R : C^2 \rightarrow [0, 1]$. Чим більше значення величини $\mu_R(\bar{c}_i, \bar{c}_j)$ близьке до 1, тим в більшому ступені дані об'єкти будуть подібними за цим критерієм. Так в працях [1-3] запропоновано три види мір подібності об'єктів: R^V , R^K , R^D .

Для утворення еліптично подібних кластерів зручно та ефективно користуватись мірою подібності «відстань», що описується нечітким бінарним відношенням R^V [2]. Нечітке бінарне відношення R^K [1, 4] характеризує кут відхилення між векторами ознак. Його використання дає можливість проводити кластеризацію кінчними кластерами. «Довжинна» міра подібності R^D дозволяє розбивати вектори ознак об'єктів на кластери концентричними сферами [3].

Кластеризацію кінчними кластерами ефективно використовувати для розв'язання багатокритеріальних задач лінійного програмування із критеріальним простором великої розмірності [4, 5], що виникають, зокрема, при математичному моделюванні задач збалансованого харчування [6, 7]. Одним із етапів розв'язання таких задач є кластеризація їх критеріального простору. При цьому взаємозв'язки між критеріями визначаються їх кутовою схожістю R^K . Кластеризація еліптичними кластерами є найбільш поширена при розв'язанні багатьох прикладних задач, так як схожість об'єктів проводиться на основі «відстаневої» міри подібності. Але використання саме однорівневого методу показало дуже хороші результати на практиці, які описані в [2]. Кластеризація концентричними кластерами (кластерами у формі концентричних сфер) [3] дала можливість групувати об'єкти за довжинною

схожістю їх векторів ознак та отримувати якісно нову прикладну змістовну інтерпретацію утворених однорідних груп на практиці.

Таким чином розроблена система дозволяє отримувати якісно нові прикладні результати та може бути використана для розв'язання широкого кола задач.

Використана література

1. Кондрук, Н. Е. Деякі методи автоматичного групування об'єктів /Н. Е. Кондрук // Південно-Європейський журнал передових технологій. – 2014. – Т. 2. – № 4 (68). – С. 20–24.

2. Kondruk N. Clustering method based on fuzzy binary relation / N. Kondruk // Eastern-European Journal of Enterprise Technologies. - 2017. - № 2(4). - С. 10-16. DOI: 10.15587/1729–4061.2017.94961 2.

3. Кондрук, Н. Е. Використання довжинної міри подібності в задачах кластеризації /Н. Е. Кондрук //Радіоелектроніка, інформатика, управління. – 2018. – № 3 (46) – С. 98-105. DOI: 10.15588/1607-3274-2018-3-11.

4. Кондрук, Н. Е. Алгоритм кластеризації критеріального простору для задач вибору /Н. Е. Кондрук, М. М. Маляр //Вісник Київського університету. Серія: фіз.-мат. наук, Вип. – Т. 3. – 2006. – Вип. 3. – С. 225–229.

5. Кондрук, Н. Э. Некоторые применения кластеризации критериального пространства для задач выбора / Кондрук Н. Э., Маляр Н. Н. // Компьютерная математика. – 2009. – № 2.– С. 142-149.

6. Кондрук, Н. Е. Застосування багатокритеріальних моделей для задач збалансованого харчування / Н. Е. Кондрук, М. М. Маляр // Вісник Черкаського державного технологічного університету. Серія: технічні науки. – 2010. – Вип. 1, № 1. – С. 3–7.

7. Кондрук Н. Е. Системи підтримки прийняття рішень для автоматизованого складання дієт / Н. Е. Кондрук // Управління розвитком складних систем. – 2015. – Вип. 23(1). – С. 110–114.