

БЕЗМЕНОВА О. М., КУЦЕНКО О. С., докт. техн. наук

РОЗВ'ЯЗАННЯ ЗАДАЧІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ, ОПИСУВАНИХ КІЛЬКІСНИМИ ТА ЯКІСНИМИ ОЗНАКАМИ

При аналізі технічних, економічних і багатьох інших систем часто виникають задачі, пов'язані з розбиттям систем на сукупність підсистем, які характеризуються великою схожістю між об'єктами, що увійшли до них.

Розв'язувана задача може бути сформульована так. Нехай є деяка сукупність об'єктів, що описується множиною ознак. Необхідно сформувати такі класи об'єктів з сукупності, щоб схожість між об'єктами, що належать різним класам, була менша схожості між об'єктами, які входять в один клас [1, 2].

Міра схожості будь-яких двох об'єктів, що описуються кількісними ознаками, може бути побудована на базі відстані Евкліда у багатовимірному просторі нормованих ознак. У разі номінальних ознак використовуються інші показники міри схожості. Якщо об'єкти характеризуються кількісними і якісними (номінальними) ознаками, як агреговану міру схожості між ними пропонується використовувати величину, визначувану формулою:

$$\mu(X_i, X_l) = \frac{m_k \mu_k(X_i, X_l) + m_n \mu_n(X_i, X_l)}{m_k + m_n}, \quad (1)$$

де $\mu_k(X_i, X_l)$ – міра схожості між об'єктами X_i і X_l , m_k та m_n – відповідно число кількісних і номінальних ознак ($m_k + m_n = m$, де m – загальна кількість ознак). Величини $\mu_k(X_i, X_l)$ і $\mu_n(X_i, X_l)$ повинні збільшуватися при збільшенні міри схожості.

Кожній підмножині об'єктів пропонується співставляти деякий еталон, за якого брати елемент підмножини, максимально схожий з усіма іншими її елементами. Для розв'язання задачі був розроблений алгоритм, який базується на ітеративному уточненні розбиття множини об'єктів на підмножини за їх мірою схожості з еталонами множин і переобчисленню самих еталонів.

Для задачі великої розмірності, для її зменшення, розроблені спеціальні алгоритми, що базуються на розбитті початкової задачі на підзадачі, розв'язанні підзадач і подальшому отриманні наближеного рішення початкової задачі.

Розроблений на базі запропонованих алгоритмів комплекс програм був протестований на статистичних даних, що відображують стан шоккових хворих при початковому і кінцевому обстеженнях. Результати тестування підтвердили достатню ефективність алгоритмів.

Список літератури: **1.** Айвазян С. А. Теория вероятностей и математическая статистика / С. А. Айвазян, В. С. Мхитарян. – М. : ЮНИТИ-ДАНА. – 2001. – 656 с. **2.** Классификация и кластер / ред. Дж. Вэн Райзин. – М. : Мир, 1980. – 391 с.