

УДК 681.513

Е.В. Бодянский, Е.В. Горшков, И.В. Кокшенев, В.В. Колодяжный

## ОБ АДАПТИВНОМ АЛГОРИТМЕ НЕЧЕТКОЙ КЛАСТЕРИЗАЦИИ ДАННЫХ

В статье рассмотрена задача кластеризации на основе вероятностного и возможностного подходов. Предложен новый адаптивный алгоритм нечеткой кластеризации без обучающего сигнала для обработки данных в реальном времени. Предложенный алгоритм основан на комбинации вероятностного и возможностного подходов.

### Введение

Задача кластеризации многомерных наблюдений, поступающих на обработку в реальном времени, достаточно часто встречается во многих приложениях, связанных с интеллектуальным анализом данных. Традиционный подход к решению этих задач предполагает, что каждое наблюдение может относиться только к одному кластеру [1, 2], хотя более естественной представляется ситуация, когда обрабатываемый вектор признаков с различными уровнями вероятностей или возможностей может принадлежать сразу нескольким классам. Данная ситуация является предметом рассмотрения нечеткого (фаззи) кластерного анализа, интенсивно развивающегося в настоящее время в двух направлениях: вероятностном подходе и подходе, основанном на возможностях [3–8]

Исходной информацией для обоих подходов является выборка наблюдений, сформированная из  $N$   $n$ -мерных векторов признаков  $X = \{x(1), x(2), \dots, x(N)\}$ ,  $x(k) \in X$ ,  $k = 1, 2, \dots, N$ . Результат работы алгоритма представляет собой разбиение исходного массива данных на  $m$  классов с некоторым уровнем  $w_j(k)$  принадлежности  $k$ -того вектора признаков  $j$ -му кластеру.

Поступающие на обработку данные предварительно центрируются и стандартизируются по всем признакам так, чтобы все наблюдения принадлежали гиперкубу  $[-1, 1]^n$ . Центрирование осуществляется либо относительно среднего, вычисляемого с помощью соотношения

$$m_i(k) = m_i(k-1) + \frac{1}{k}(x_i(k) - m_i(k-1)),$$

либо с целью придания процедуре центрирования робастных свойств (защита от аномальных наблюдений) оно осуществляется относительно медианы, которая вычисляется согласно рекуррентному соотношению

$$me_i(k) = me_i(k-1) + \eta_m \operatorname{sign}(x_i(k) - me_i(k-1)), \quad i = 1, 2, \dots, n,$$

где  $\eta_m$  — параметр шага поиска, выбираемый в стационарном случае в соответствии с условиями Дворецкого.

© Е.В. Бодянский, Е.В. Горшков, И.В. Кокшенев, В.В. Колодяжный, 2002

В настоящей работе предпринята попытка синтеза адаптивного вычислительно простого алгоритма нечеткой кластеризации для рекуррентной обработки данных в реальном времени по мере поступления данных, сочетающего в себе свойства как вероятностного, так и возможностного подходов.

### Пакетный алгоритм нечеткой кластеризации

Алгоритмы, основанные на целевых функциях [3], предназначены для решения задачи кластеризации путем оптимизации некоторого наперед заданного критерия качества кластеризации и являются наиболее строгими с математической точки зрения.

Целевая функция, подлежащая минимизации имеет вид

$$E(w_j(k), c_j) = \sum_{k=1}^N \sum_{j=1}^m w_j^\beta(k) d^2(x(k), c_j) \quad (1)$$

при ограничениях

$$\sum_{j=1}^m w_j(k) = 1, \quad k = 1, \dots, N, \quad (2)$$

$$0 < \sum_{k=1}^N w_j(k) < N, \quad j = 1, \dots, m. \quad (3)$$

Здесь  $w_j(k) \in [0, 1]$  — уровень принадлежности вектора  $x(k)$  к  $j$ -му кластеру,  $c_j$  — прототип (центр)  $j$ -того кластера,  $\beta$  — неотрицательный параметр, именуемый “фазсификатором” (обычно  $\beta = 2$ ),  $d^2(x(k), c_j)$  — расстояние между  $x(k)$  и  $c_j$  в принятой метрике. Результатом кластеризации предполагается  $N \times m$  матрица  $W = \{w_j(k)\}$ , называемая “матрицей нечеткого разбиения”.

Заметим, что поскольку элементы матрицы  $W$  могут рассматриваться как вероятности гипотез принадлежности векторов данных определенным кластерам, процедуры, порождаемые (1) при ограничениях (2), (3), называются “вероятностными алгоритмами кластеризации”.

Вводя функцию Лагранжа

$$\begin{aligned} L(w_j(k), c_j, \lambda(k)) &= \sum_{k=1}^N \sum_{j=1}^m w_j^\beta(k) d^2(x(k), c_j) + \sum_{k=1}^N \lambda(k) \left( \sum_{j=1}^m w_j(k) - 1 \right) = \\ &= \sum_{k=1}^N \left( \sum_{j=1}^m w_j^\beta(k) d^2(x(k), c_j) + \lambda(k) \left( \sum_{j=1}^m w_j(k) - 1 \right) \right) \end{aligned} \quad (4)$$

(здесь  $\lambda(k)$  — неопределенный множитель Лагранжа) и решая систему

уравнений Куна-Таккера

$$\begin{cases} \frac{\partial L(w_j(k), c_j, \lambda(k))}{\partial w_j(k)} = 0, \\ \nabla_{c_j} L(w_j(k), c_j, \lambda(k)) = 0, \\ \frac{\partial L(w_j(k), c_j, \lambda(k))}{\partial \lambda(k)} = 0, \end{cases}$$

несложно получить искомое решение в виде

$$w_j^{pr}(k) = \frac{(d^2(x(k), c_j))^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (d^2(x(k), c_l))^{\frac{1}{1-\beta}}}, \quad (5)$$

$$c_j^{pr} = \frac{\sum_{k=1}^N w_j^\beta(k)x(k)}{\sum_{k=1}^N w_j^\beta(k)}, \quad (6)$$

$$\lambda(k) = - \left( \sum_{l=1}^m (\beta d^2(x(k), c_l))^{\frac{1}{1-\beta}} \right)^{1-\beta}. \quad (7)$$

Уравнения (5)–(7) порождают широкий класс процедур кластеризации. Выбирая  $\beta = 2$  и принимая евклидово расстояние  $d^2(x(k), c_j) = \|x(k) - c_j\|^2$ , получаем простой и эффективный алгоритм нечеткой кластеризации Бездека [3]:

$$w_j^{pr}(k) = \frac{\|x(k) - c_j\|^{-2}}{\sum_{l=1}^m \|x(k) - c_l\|^{-2}}, \quad (8)$$

$$c_j^{pr} = \frac{\sum_{k=1}^N w_j^2(k)x(k)}{\sum_{k=1}^N w_j^2(k)}, \quad (9)$$

$$\lambda(k) = - \sum_{l=1}^m \left( \frac{\|x(k) - c_l\|^{-2}}{2} \right)^{-1}.$$

К вероятностным алгоритмам кластеризации относятся также алгоритмы Густафсона-Кесселя [4], Гата-Джевы [9], и ряд других. Основные недостатки вероятностного подхода связаны с ограничениями (2). В простейшем случае двух кластеров ( $m = 2$ ) несложно видеть, что наблюдения  $x(k)$ , равноправно принадлежащее обоим кластерам, и наблюдения  $x(p)$ , не принадлежащее ни одному из них, имеют одинаковые уровни принадлежности  $w_1^{pr}(k) = w_2^{pr}(k) = w_1^{pr}(p) = w_2^{pr}(p) = 0.5$ . Естественно, что данное обстоятельство, ухудшающее точность классификации, привело к появлению возможных подходов к нечеткой классификации [10, 11].

В возможных алгоритмах кластеризации критерий имеет вид

$$E(w_j(k), c_j) = \sum_{k=1}^N \sum_{j=1}^m w_j^\beta(k) d^2(x(k), c_j) + \sum_{j=1}^m \mu_j \sum_{k=1}^N (1 - w_j(k))^\beta, \quad (10)$$

где скалярный параметр  $\mu_j > 0$  определяет расстояние, на котором уровень принадлежности принимает значение 0.5, то есть если  $d^2(x(k), c_j) = \mu_j$ , то  $w_j(k) = 0.5$ .

Минимизация (10) по  $w_j(k)$ ,  $c_j$ , и  $\mu_j$  дает очевидное решение

$$w_j^{pos}(k) = \left( 1 + \left( \frac{d^2(x(k), c_j)}{\mu_j} \right)^{\frac{1}{\beta-1}} \right)^{-1}, \quad (11)$$

$$c_j^{pos} = \frac{\sum_{k=1}^N w_j^\beta(k) x(k)}{\sum_{k=1}^N w_j^\beta(k)}, \quad (12)$$

$$\mu_j = \frac{\sum_{k=1}^N w_j^\beta(k) d^2(x(k), c_j)}{\sum_{k=1}^N w_j^\beta(k)}. \quad (13)$$

Видно, что возможные и вероятностные алгоритмы очень похожи и переходят друг в друга заменой выражения (11) на формулу (5), и наоборот. Общим недостатком рассмотренных алгоритмов является их вычислительная сложность и невозможность работы в реальном времени.

Работа алгоритма (5)–(7) начинается с задания начальной (обычно случайной) матрицы разбиения  $W^0$ . На основе ее значений рассчитывается начальный набор прототипов  $c_j^0$ , которые затем используются для вычисления новой матрицы  $W^1$ . Затем в пакетном режиме пересчитываются  $c_j^1, W^2, \dots, W^t, c_j^t, W^{t+1}$  и т.д., пока разность  $\|W^{t+1} - W^t\|$  не станет меньше некоторого наперед заданного порога  $\varepsilon$ . Таким образом, вся имеющаяся выборка данных обрабатывается многократно.

Решение, полученное с помощью вероятностного алгоритма, рекомендуется использовать в качестве начальных условий для возможностного алгоритма (11)–(13) [11, 12]. Параметры расстояния  $\mu_j^t$  инициализируются в соответствии с (15) по результатам работы вероятностного алгоритма.

### Адаптивные алгоритмы кластеризации

Анализ уравнения (5) показывает, что для расчета уровней принадлежности  $w_j(k)$  вместо лагранжиана (4) можно использовать его локальную модификацию:

$$L_k(w_j(k), c_j, \lambda(k)) = \sum_{j=1}^m w_j^\beta(k) d^2(x(k), c_j) + \lambda(k) \left( \sum_{j=1}^m w_j(k) - 1 \right). \quad (14)$$

Оптимизация выражения (14) с помощью процедуры Эрроу-Гурвица-Удзавы дает следующий алгоритм:

$$w_j^{pr}(k) = \frac{(d^2(x(k), c_j(k)))^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (d^2(x(k), c_l(k)))^{\frac{1}{1-\beta}}}, \quad (15)$$

$$\begin{aligned} c_j^{pr}(k+1) &= c_j^{pr}(k) - \eta(k) \nabla_{c_j} L_k(w_j(k), c_j^{pr}(k), \lambda(k)) = \\ &= c_j^{pr} - \eta(k) w_j^\beta(k) d(x(k+1), c_j^{pr}(k)) \nabla_{c_j} d(x(k+1), c_j^{pr}(k)), \end{aligned} \quad (16)$$

где  $\eta(k)$  — параметр скорости обучения,  $c_j^{pr}(k)$  — прототипы  $j$ -го кластера, вычисленные на выборке из  $k$  наблюдений.

Процедура (15), (16) довольно похожа на алгоритм обучения Чанга-Ли [13] и для  $\beta = 2$  совпадает с градиентной процедурой кластеризации Парка-Дэггера [14]:

$$w_j^{pr}(k) = \frac{\|x(k) - c_j(k)\|^{-2}}{\sum_{i=1}^m \|x(k) - c_i(k)\|^{-2}},$$

$$c_j^{pr}(k+1) = c_j^{pr}(k) + \eta(k)w_j^2(k)(x(k+1) - c_j^{pr}(k)).$$

В рамках возможностного подхода локальный критерий принимает форму

$$E_k(w_j(k), c_j) = \sum_{j=1}^m w_j^\beta(k)d^2(x(k), c_j) + \sum_{j=1}^m \mu_j(1 - w_j(k))^\beta.$$

а результат его оптимизации имеет вид

$$w_j^{pos} = \left( 1 + \left( \frac{d^2(x(k), c_j(k))}{\mu_j(k)} \right)^{\frac{1}{\beta-1}} \right)^{-1}, \quad (17)$$

$$c_j^{pos}(k+1) = c_j^{pos}(k) - \eta(k)w_j^\beta(k)d(x(k+1), c_j^{pos}(k)) \cdot \nabla_{c_j} d(x(k+1), c_j^{pos}(k)), \quad (18)$$

$$\mu_j(k+1) = \frac{\sum_{p=1}^k w_j^\beta(p)d^2(x(p), c_j(k+1))}{\sum_{p=1}^k w_j^\beta(p)}. \quad (19)$$

В квадратичном случае, алгоритм (17)–(19) преобразуется в достаточно простую конструкцию

$$w_j^{pos}(k) = \frac{\mu_j(k)}{\mu_j(k) + \|x(k) - c_j(k)\|^2}, \quad (20)$$

$$c_j^{pos}(k+1) = c_j^{pos}(k) + \eta(k)w_j^2(k)(x(k+1) - c_j^{pos}(k)), \quad (21)$$

$$\mu_j(k+1) = \frac{\sum_{p=1}^k w_j^2(p)\|x(p) - c_j(k+1)\|^2}{\sum_{p=1}^k w_j^2(p)}. \quad (22)$$

Параллельное применение адаптивных вероятностного и возможностного алгоритмов приводит к объединенной процедуре

$$\left\{ \begin{array}{l} c_j^{pr}(k) = c_j^{pos}(k-1) - \eta(k)w_j^{pos}(k-1)d(x(k), c_j^{pos}(k-1)) \cdot \nabla_{c_j} d(x(k), c_j^{pos}(k-1)), \\ w_j^{pr}(k) = (d^2(x(k), c_j^{pr}(k)))^{\frac{1}{1-\beta}} \left( \sum_{l=1}^m (d^2(x(k), c_l^{pr}(k)))^{\frac{1}{1-\beta}} \right)^{-1}, \\ c_j^{pos}(k) = c_j^{pr} - \eta(k)w_j^{pr \beta}(k)d(x(k), c_j^{pr}(k))\nabla_{c_j} d(x(k), c_j^{pr}(k)), \\ \mu_j(k) = \left( \sum_{p=1}^k w_j^{pr \beta}(p)d^2(x(p), c_j^{pos}(k)) \right) \left( \sum_{p=1}^k w_j^{pr \beta}(p) \right)^{-1}, \\ w_j^{pos}(k) = \left( 1 + \left( \frac{d^2(x(k), c_j^{pos}(k))}{\mu_j(k)} \right)^{\frac{1}{1-\beta}} \right)^{-1}, \quad j = 1, 2, \dots, m. \end{array} \right. \quad (23)$$

Признаком правильного восстановления прототипов (а следовательно и корректной кластеризации) с помощью алгоритма (23) является выполнение неравенства

$$\sum_{l=1}^m d^2(c_l^{pr}(k), c_l^{pos}(k)) \leq \varepsilon,$$

где  $\varepsilon$  определяет приемлемую точность кластеризации.

Для евклидовой метрики значение параметра  $\mu_j(k)$  может вычисляться согласно рекуррентному соотношению

$$\left\{ \begin{array}{l} \beta_q(k) = \beta_q(k-1) + w_j^{pr \beta}(k)s_q(x(k)), \quad q = 0, 1, 2, \\ \mu_j(k) = \frac{\beta_2(k-1) - 2c_j^{pos T}(k)\beta_1(k-1) + \|c_j^{pos}(k)\|^2\beta_0(k-1)}{\beta_0(k-1)}, \end{array} \right. \quad (24)$$

где

$$s_q(x(k)) = \begin{cases} 1, & \text{если } q = 0, \\ x(k), & \text{если } q = 1, \\ \|x(k)\|^2, & \text{если } q = 2. \end{cases}$$

Параметр  $\beta_q(k)$  инициализируется как

$$\beta_q(N) = \sum_{p=1}^N w_j^{pr \beta}(p)s_q(x(p)), \quad q = 0, 1, 2.$$

Предложенный адаптивный алгоритм может использоваться как в пакетном режиме для итеративной обработки заданной выборки, так и в режиме реального времени, где количество наблюдений  $k$  определяется текущим дискретным временем  $k = 1, 2, \dots, N, N+1, \dots$ . В этом случае алгоритм последовательно обрабатывает наблюдения, поступающие на вход, настраивая уровни принадлежности и прототипы кластеров под новые данные.

## Экспериментальные исследования

Мы тестировали работу предложенного алгоритма (23) в задачах классификации данных и сравнивали результаты с результатами алгоритма Бездека, пакетным алгоритмом возможностной кластеризации и рекуррентным алгоритмом кластеризации Парка-Дэггера. Для тестирования использовались три широко известных набора данных: данные “Вина”, данные “Ирисы” и данные “Щитовидная железа” из репозитория UCI [15]. Данные “Ирисы” содержат описания 150 экземпляров цветов ириса, равномерно распределенных на 3 вида. Цветы описываются 4 атрибутами. Данные “Вина” содержат 178 результатов химического анализа вин, полученных из винограда трех разных сортов, выращенных в одном регионе Италии. Анализ определял количество 13 компонент, присутствующих в каждом из 3 типов вин. Данные “Щитовидная железа” содержат 215 результатов (разделенных на 3 класса) медицинских анализов по 5 параметрам.

Задача классификации состоит в отнесении каждой представленной комбинации признаков к определенному классу. Для простоты мы предполагаем, что количество кластеров равно количеству классов.

Вышеописанные алгоритмы кластеризации используются для построения классификатора. В наборе данных  $X$ , который является входной информацией для алгоритмов кластеризации, каждый вектор  $x(k) \in X$ ,  $k = 1, 2, \dots, N$  состоит из вектора признаков и из присоединенных векторов вида  $(0, \dots, 0, 1, 0, \dots, 0)^T$ , где индекс единицы определяет номер класса. Таким образом, классификация выполняется в пространстве “вход-выход” [16]. Центры полученных кластеров в пространстве входов представляют прототипы в пространстве признаков, а присоединенные координаты классов в пространстве выходов представляют метки соответствующих классов.

При классификации векторы  $x(k)$  соответствуют векторам признаков, а присоединенные координаты классов в векторах прототипов отбрасываются. Класс определяется по кластеру с наибольшим уровнем принадлежности данного объекта.

Наборы данных были поделены на обучающую и тестовую выборки, содержащие 70% и 30% данных соответственно. Для лучшей работы рекуррентных алгоритмов наборы данных были случайным образом перемешаны. Обучающие выборки использовались для инициализации классификатора с помощью нечеткой кластеризации, а тестовые выборки использовались для сравнения точности классификации.

Для сравнения качества работы алгоритмов было проведено два эксперимента. В первом эксперименте сравнивалась работа алгоритмов кластеризации в задачах классификации, когда в выборке данных, используемой для кластеризации, присутствуют экземпляры всех имеющих классов, то есть количество классов известно заранее.

Шаг обучения был принят  $\eta(k) = 0.01$  в рекуррентных процедурах Парка-Дэггера и (23), параметр “фазсификатора” —  $\beta = 2$  в пакетной

возможностной процедуре кластеризации (11)–(13). Процедуры инициализировались по результатам вероятностной кластеризации с помощью алгоритма Бездека.

Было выполнено по 10 итераций для пакетных процедур и по 10 проходов по всей выборке для рекуррентных процедур. Эксперимент был повторен 50 раз, а затем были вычислены усредненные результаты, которые представлены в табл. 1. Они показывают количество (в процентах) неправильно классифицированных объектов из тестовой выборки.

Табл. 1 – Классификация с известным количеством классов

Тест	Бездека	Пакетный возможно- стный	Парка- Дэггера	Предложен- ный алгоритм
Ирисы	7.4%	7.6%	6.9%	7.8%
Вина	3.8%	4.4%	3.9%	4.3%

Результаты 1-го эксперимента в табл. 1 довольно близки для всех протестированных алгоритмов кластеризации. Алгоритм (23) показал результаты, аналогичные полученным с помощью пакетного возможностного алгоритма (11)–(13).

Во втором эксперименте мы включили в обучающую выборку только экземпляры 2-х из 3-х имеющихся классов. Таким образом, количество классов в обучающих выборках было меньшим, чем в тестовых. Процедуры кластеризации использовались для создания кластеров только известных классов. В тестовые выборки включались объекты третьего неизвестного класса. Результаты 2-го эксперимента представлены в табл. 2.

Табл. 2 – Классификация с неизвестным классом

Тест, не- известный класс	Бездека	Пакетный возможно- стный	Парка- Дэггера	Предложен- ный алгоритм
Ирисы, класс 3	33.3%	14.0%	33.3%	15.3%
Ирисы, класс 1	38.0%	6.0%	38.0%	6.7%
Вина, класс 3	29.0%	19.6%	29.0%	19.6%
Вина, класс 1	35.9%	23.6%	35.9%	23.4%
Тироиды, класс 2	18.1%	20.9%	18.6%	11.6%
Тироиды, класс 3	15.4%	17.2%	15.4%	5.6%

Как указание на неизвестный класс, мы использовали порог 0.2 для суммы уровней принадлежности. Естественно, классификаторы, основанные на вероятностных процедурах, были практически неспособны отличить объекты неизвестного класса от объектов двух других клас-

сов, поскольку вероятностные процедуры основаны на предположении о равенстве суммы уровней принадлежности единице. Это не так для возможностей процедур кластеризации, которые показали существенно лучшие результаты во втором эксперименте.

### Выводы

В данной статье рассмотрены вероятностные и возможностные алгоритмы кластеризации. Предложена рекуррентная модификация вероятностного и возможностного алгоритмов, выполняемых параллельно. Компьютерное моделирование демонстрирует применение предложенного алгоритма для решения задач классификации данных. Алгоритм может использоваться для диагностики неисправностей, интеллектуального анализа данных в сети Internet, распознавания образов и т.д., когда размер выборки данных неизвестен заранее и данные должны обрабатываться в реальном времени. Результаты экспериментальных исследований подтверждают эффективность предложенного алгоритма.

### Литература

1. *MacQueen J.* On convergence of k-means and partitions with minimum average variance // *Ann. Math. Statist.* – 1965. – 36. – p. 1084.
2. *Cover T. M.* Estimates by the nearest-neighbor rule // *IEEE Trans. on Information Theory.* – 1968. – 14. – p. 50–55.
3. *Bezdek J. C.* Pattern Recognition with Fuzzy Objective Function Algorithms. — N.Y.: Plenum Press, 1981. – p. 272.
4. *Gustafson E. E., Kessel W. C.* Fuzzy clustering with a fuzzy covariance matrix // In: *Proc. IEEE CDC.* – San Diego, California, 1979. – p. 761–766.
5. *Keller J. M., Gray M. R., Givens J. A., Jr.* A fuzzy k-nearest neighbor algorithm // *IEEE Trans. on Syst., Man and Cybern.* – 1985. – 3, N3. – p. 32–57.
6. *Klawonn F., Kruse R., Timm H.* Fuzzy shell cluster analysis // In: Della Riccia G., Lenz H. J., Kruse R. (eds.): *Learning, Networks and Statistics.* – Wien: Springer-Verlag, 1997. – p. 105–120.
7. *Yager R. R., Filev D. P.* Approximate clustering via the mountain method // *IEEE Trans. on Syst., Man and Cybern.* – 1994. – 24. – p. 1279–1284.
8. *Höppner F., Klawonn F., Kruse R.* Fuzzy Clusteranalyse. — Braunschweig: Vieweg, 1999. – p. 280.
9. *Gath I., Geva A. B.* Unsupervised optimal fuzzy clustering // *IEEE Trans. on Pattern Analysis and Machine Intelligence.* – 1989. – 11. – p. 773–781.
10. *Krishnapuram R., Keller J.* A possibilistic approach to clustering // *IEEE Trans. on Fuzzy Systems.* – 1993. – 1. – p. 98–110.
11. *Krishnapuram R., Keller J.* Fuzzy and possibilistic clustering methods for computer vision // *Neural Fuzzy Systems.* – 1994. – 12. – p. 133–159.

12. *Klawonn F., Kruse R.* Constructing a fuzzy controller from data // *Fuzzy Sets and Systems*. – 1997. – 85. – p. 117–193.
13. *Chung F. L., Lee T.* Fuzzy competitive learning // *Neural Networks*. – 1994. – 7, N3. – p. 539–552.
14. *Park D. C., Dagher I.* Gradient based fuzzy c-means (GBFCM) algorithm // *Proc. IEEE Int. Conf. on Neural Networks*. – 1984. – p. 1626–1631.
15. *Murphy P. M., Aha D. W.* UCI Repository of machine learning databases. – URL: <http://www.ics.uci.edu/~mlearn/MLRepository.html>. — CA: University of California, Department of Information and Computer Science, 1994.
16. *Abe S.* *Neural Networks and Fuzzy Systems*. — Boston: Kluwer Academic Publishers, 1996. – p. 257.

Получено: 03.11.2002