ИНФОРМАТИКА

2011 июль-сентябрь № 3

УДК 591.2/.6(082)

В.В. Анищенко[1], Д.А. Вятченин[1], А.В. Доморацкий[2]

ОБУЧЕНИЕ КОНСЕКВЕНТОВ НЕЧЕТКИХ ПРАВИЛ, ПОСТРОЕННЫХ НА ОСНОВЕ ЭВРИСТИЧЕСКОЙ ВОЗМОЖНОСТНОЙ КЛАСТЕРИЗАЦИИ

Предлагается метод построения функций принадлежности нечетких множеств, соответствующих консеквентам нечетких правил, которые сгенерированы на основе результатов обработки обучающей выборки эвристическим алгоритмом возможностной кластеризации. Эффективность метода демонстрируется на иллюстративном примере.

Введение

В задачах нечеткого управления, нечеткой классификации и принятия решений в нечеткой среде центральное место отводится системам нечеткого вывода, представляющим собой алгоритм получения нечетких заключений на основе нечетких предпосылок [1]. Главным элементом систем нечеткого вывода является база нечетких продукционных правил, в наиболее общем случае имеющих вид

$$l\text{: ЕСЛИ } \hat{x}^1 \text{ есть } B_l^1 \text{ И} \ldots \text{И } \hat{x}^m \text{ есть } B_l^m \text{ ТО } y_1 \text{ есть } C_1^l \text{ И} \ldots \text{И } y_c \text{ есть } C_c^l, \qquad (1)$$

где $l \in \{1,\ldots,c\}$ – номер правила; $\hat{x}^t \in \hat{X}^t$, $t \in \{1,\ldots,m\}$, – входные переменные; $\hat{X}^t$ – область определения соответствующей переменной; $y_l \in Y_l$, $l \in \{1,\ldots,c\}$, – нечеткие выходные переменные, причем Y_l – область определения соответствующего заключения, а B_l^t, C_l^l – нечеткие множества с функциями принадлежности $\gamma_{B_l^t}(\hat{x}^t)$ и $\gamma_{C_l^l}(y_l)$, определенные на соответствующих универсумах. Следует отметить, что в правиле вида (1) m посылкам соответствует c заключений, так что структура правил вида (1) в специальной литературе именуется MIMO-структурой (от англ. Multi Inputs – Multi Outputs) [1]. Вместе с тем любое правило вида (1) может быть представлено c правилами вида

$$l\text{: ЕСЛИ } \hat{x}^1 \text{ есть } B_l^1 \text{ И} \ldots \text{И } \hat{x}^m \text{ есть } B_l^m \text{ ТО } y_l \text{ есть } C_l^l, \qquad (2)$$

структура которых именуется MISO-структурой (от англ. Multi Inputs – Single Output). Правила вида (1) обладают тем преимуществом, что их использование позволяет значительно сократить размерность базы правил. Необходимо также указать, что условная часть некоторого нечеткого продукционного правила в специальной литературе называется антецедентом, а заключение – консеквентом [1].

На формирование базы правил систем нечеткого вывода часто оказывает влияние ряд факторов, определяемых спецификой решаемой задачи или используемого алгоритма нечеткого вывода, наиболее известными из которых являются алгоритмы Мамдани, Цукамото, Ларсена и Такаги – Сугэно [1].

База нечетких правил может формироваться, с одной стороны, экспертным путем, а с другой – на основе обработки данных обучающей выборки, для чего чаще всего используются оптимизационные методы нечеткой или возможностной кластеризации [2]. При этом исследуемая совокупность $X = \{x_1,\ldots,x_n\}$ объектов обучающей выборки обрабатывается каким-либо методом нечеткой или возможностной кластеризации с последующим проецированием значений принадлежности μ_{li} или значений типичности υ_{li} того или иного нечеткого кластера A^l, $l = 1,\ldots,c$, на координатные оси признакового пространства $I^m(X)$. Так, при обращении к методам нечеткой кластеризации функции принадлежности $\tilde{\gamma}_{B_l^t}(\hat{x}^t)$ соответствующих проекций вычисляются по формуле

$$\tilde{\gamma}_{B_l^t}(\hat{x}^t) = \sup\left\{ \frac{1}{\left(\sum_{a=1}^{c} \left(\frac{d(x_i, \overline{\tau}^l)}{d(x_i, \overline{\tau}^a)} \right)^{1/(\gamma-1)} \right)} \middle| x_i = (\hat{x}_i^1, \ldots, \hat{x}_i^{t-1}, \hat{x}_i^t, \hat{x}_i^{t+1}, \ldots, \hat{x}_i^m) \in \Re^m \right\}, \tag{3}$$

где x_i, $i = 1, \ldots, n$, – объекты исследуемой совокупности; $\overline{\tau}^l$, $l = 1, \ldots, c$, – центры нечетких кластеров – элементов нечеткого c-разбиения $P(X)$; γ, $1 < \gamma < \infty$, – показатель нечеткости классификации. В свою очередь, при обращении к методам возможностной кластеризации функции принадлежности $\tilde{\gamma}_{B_l^t}(\hat{x}^t)$ вычисляются в соответствии с выражением

$$\tilde{\gamma}_{B_l^t}(\hat{x}^t) = \sup\left\{ \frac{1}{1 + \left(d(x_i, \overline{\tau}^l) / \eta_l \right)^{1/(\psi-1)}} \middle| x_i = (\hat{x}_i^1, \ldots, \hat{x}_i^{t-1}, \hat{x}_i^t, \hat{x}_i^{t+1}, \ldots, \hat{x}_i^m) \in \Re^m \right\}, \tag{4}$$

где $1 < \psi < \infty$ – возможностный аналог показателя нечеткости классификации; $\eta_l > 0$, $l = 1, \ldots, c$, – параметр, характеризующий каждый нечеткий кластер, являющийся элементом возможностного разбиения $\Upsilon(X)$.

Полученные с помощью выражений (3) или (4) нечеткие множества с дискретными функциями принадлежности $\tilde{\gamma}_{B_l^t}(\hat{x}^t)$ аппроксимируются параметрическими непрерывными функциями принадлежности $\tilde{\gamma}_{B_l^t}(\hat{x}^t)$ (как правило, треугольной или трапециевидной формы), для чего целесообразно воспользоваться алгоритмом, предложенным М. Сугэно и Т. Ясукавой [3].

Данный подход обладает довольно существенным недостатком: в силу того что в оптимизационных методах нечеткой и возможностной кластеризации первоначальное разбиение формируется, как правило, случайным образом, зачастую для получения приемлемого результата классификации в виде нечеткого c-разбиения или возможностного разбиения необходимым является проведение серии вычислительных экспериментов. Указанного недостатка лишен предложенный в [4] метод прототипирования систем нечеткого вывода, основанный на обработке данных об объектах обучающей выборки эвристическим D-AFC(c)-алгоритмом возможностной кластеризации [5]. Вместе с тем в изложенном в [4] подходе рассмотрен случай полностью разделенных нечетких кластеров, вследствие чего функции принадлежности заключений строятся только для нечетких множеств C_l^l с совпадающими значениями нижнего и верхнего индексов $l \in \{1, \ldots, c\}$. Целью предпринятого исследования является обобщение изложенного в [4] метода построения функций принадлежности $\gamma_{C_l^l}(y_l)$, $l = 1, \ldots, c$, нечетких множеств, соответствующих консеквентам продукционных правил вида (2), или, иными словами, обучение консеквентов в случае частично разделенных нечетких кластеров.

1. Основные понятия эвристического метода возможностной кластеризации

Сущность эвристического метода возможностной кластеризации заключается в построении так называемого распределения по априори задаваемому числу c нечетких α-кластеров, удовлетворяющих введенному определению. Такое распределение является, как продемонстрировано в [6], частным случаем возможностного разбиения [7], в общем случае определяемого условием

$$\upsilon_{li} \geq 0, \ \sum_{l=1}^{c} \upsilon_{li} > 0, \ i = 1, \ldots, n, \ l = 1, \ldots, c, \tag{5}$$

где $X = \{x_1,...,x_n\}$ – совокупность объектов, на которой определена нечеткая толерантность T с функцией принадлежности $\mu_T(x_i, x_j)$, $i, j = 1,\ldots,n$, являющаяся матрицей исходных данных для D-AFC(c)-алгоритма, так что строки или столбцы этой матрицы представляют собой нечеткие множества $\{A^1,..., A^n\}$. В таком случае для некоторого значения α, $\alpha \in (0,1]$, нечеткое множество уровня α, определяемое условием $A^l_{(\alpha)} = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \ge \alpha\}$, $l \in [1, n]$, такое, что $A^l_{(\alpha)} \subseteq A^l$, $A^l \in \{A^1,\ldots, A^n\}$, будет именоваться нечетким α-кластером с функцией принадлежности υ_{li} объекта $x_i \in X$ нечеткому α-кластеру $A^l_{(\alpha)}$, определяемой выражением

$$\upsilon_{li} = \begin{cases} \mu_{A^l}(x_i), & \text{если } x_i \in A^l_\alpha; \\ 0 & \text{в противном случае,} \end{cases} \tag{6}$$

где $A^l_\alpha = \{x_i \in X \mid \mu_{A^l}(x_i) \ge \alpha\}$ – α-уровень A^l, $l \in \{1,\ldots,n\}$. Объект $x_i \in X$, обладающий наибольшим значением функции принадлежности υ_{li} некоторому нечеткому α-кластеру $A^l_{(\alpha)}$ и определяемый выражением

$$\tau^l = \arg\max_{x_i} \upsilon_{li},\ \forall x_i \in A^l_\alpha,\ A^l_{(\alpha)} \in \{A^1_{(\alpha)},..., A^c_{(\alpha)}\}, \tag{7}$$

именуется его типичной точкой и обозначается τ^l, а функция принадлежности, определяемая выражением (5), показывает степень сходства i-го объекта множества X с типичной точкой τ^l соответствующего нечеткого α-кластера. Как указывается в [8], нечеткий α-кластер $A^l_{(\alpha)}$, $\alpha \in (0,1]$, $l \in \{1,\ldots,n\}$, может обладать более чем одной типичной точкой τ^l_e, $e = 1, 2, \ldots$, и индекс e обозначает номер типичной точки нечеткого кластера $A^l_{(\alpha)}$, а множество $K(A^l_{(\alpha)}) = \{\tau^l_1,\ldots,\tau^l_{|e|}\}$ типичных точек нечеткого α-кластера $A^l_{(\alpha)}$ в дальнейшем будет именоваться его ядром, так что количество $|e|$ типичных точек τ^l_e, $e = 1,\ldots,|e|$, определяет мощность ядра нечеткого α-кластера $A^l_{(\alpha)}$, $card\left(K(A^l_{(\alpha)})\right) = |e|$. Если условие (5) выполняется для всех $A^l_{(\alpha)} \in R^\alpha_z(X)$, где $R^\alpha_z(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, 2 \le c \le n\}$ – семейство c нечетких α-кластеров для некоторого значения α, порожденных заданной на X нечеткой толерантностью T, то это семейство является распределением множества классифицируемых объектов X по c нечетким α-кластерам. Условие (5) в рассматриваемом случае требует, чтобы все объекты совокупности X были распределены по c нечетким α-кластерам $\{A^1_{(\alpha)},..., A^c_{(\alpha)}\}$ с положительными значениями υ_{li}, $l = 1,\ldots,c$, $i = 1,\ldots,n$.

Распределение, удовлетворяющее условиям задачи классификации в конкретном случае, именуется адекватным распределением. В общем случае для нечетких α-кластеров $A^l_{(\alpha)}$, $l = 1,\ldots,c$, некоторого распределения $R^\alpha_z(X) = \{A^l_{(\alpha)} \mid l = \overline{1,c}, 2 \le c \le n\}$ требуется выполнение условия

$$card(A^l_\alpha \cap A^m_\alpha) \le w,\ \forall A^l_{(\alpha)}, A^m_{(\alpha)},\ l \ne m,\ \alpha \in (0,1], \tag{8}$$

где $w \in \{0,\ldots,n\}$ – максимальное число объектов в области пересечения носителей любых двух нечетких α-кластеров некоторого распределения $R^\alpha_z(X)$. Очевидно, что при $w = 0$ в (8) выполняется равенство, так что имеет место условие

$$\sum_{l=1}^{c} card(A_{\alpha}^{l}) \geq card(X)\,,\ \forall A_{(\alpha)}^{l} \in R_{z}^{\alpha}(X)\,,\ card(R_{z}^{\alpha}(X)) = c\,, \tag{9}$$

и при выполнении равенства в (9) в (8) при $w = 0$ будет иметь место равенство, а выполнение (8) при $w > 0$ влечет выполнение строгого неравенства в (9). Таким образом, при $w = 0$ условия (8) и (9) принимают вид

$$card(A_{\alpha}^{l} \cap A_{\alpha}^{m}) = 0\,,\ \forall A_{(\alpha)}^{l}, A_{(\alpha)}^{m}\,,\ l \neq m\,,\ \alpha \in (0,1] \tag{10}$$

и

$$\bigcup_{l=1}^{c} A_{\alpha}^{l} = X \tag{11}$$

соответственно, где (10) выражает необходимость принадлежности каждого элемента множества классифицируемых объектов X какому-либо одному нечеткому α-кластеру $A_{(\alpha)}^{l}$, $l \in [1, c]$, с положительным значением функции принадлежности υ_{li}, а условие (11) формализует требование, согласно которому все объекты исследуемой совокупности $X = \{x_1,..., x_n\}$ должны быть расклассифицированы. Если любые два различных нечетких α-кластера распределения $R_{z}^{\alpha}(X)$ удовлетворяют условиям (8) и (9), то их носители имеют общие элементы и соответствующие кластеры именуются частично разделенными, а если же все нечеткие α-кластеры распределения $R_{z}^{\alpha}(X)$ удовлетворяют условиям (10) и (11), то их носители образуют четкое разбиение множества классифицируемых объектов X, а нечеткие α-кластеры именуются полностью разделенными.

Сущность D-AFC(c)-алгоритма заключается в построении множества допустимых решений $B(c) = \{R_{z}^{\alpha}(X)\}$ для c классов с последующим выбором в качестве решения задачи классификации некоторого единственного распределения $R^{*}(X) \in B(c)$. Выбор $R^{*}(X)$ основывается на вычислении для всех $R_{z}^{\alpha}(X) \in B(c)$ критерия

$$F(R_{z}^{\alpha}(X), \alpha) = \sum_{l=1}^{c} \frac{1}{n_l} \sum_{i=1}^{n_l} \upsilon_{li} - \alpha \cdot c\,, \tag{12}$$

определяющего качество каждого $R_{z}^{\alpha}(X) \in B(c)$, где $n_l = card(A_{\alpha}^{l})$ – мощность носителя нечеткого множества $A_{(\alpha)}^{l} \in R_{z}^{\alpha}(X)$, $l \in \{1,\ldots,c\}$, $\alpha \in (0,1]$, так что (12) определяет среднюю суммарную принадлежность объектов множества X нечетким α-кластерам $\{A_{(\alpha)}^{1},..., A_{(\alpha)}^{c}\}$ распределения $R_{z}^{\varepsilon}(X)$ за вычетом величины $\alpha \cdot c$, регуляризующей число классов в $R_{z}^{\alpha}(X)$. При этом оптимальному распределению $R^{*}(X)$ соответствует максимальное значение (12) и решение состоит в нахождении такого распределения, при котором

$$R^{*}(X) = \arg \max_{R_{z}^{\alpha}(X) \in B(c)} F(R_{z}^{\alpha}(X), \alpha)\,. \tag{13}$$

Результатом работы D-AFC(c)-алгоритма является не только распределение $R^{*}(X)$ объектов совокупности X по заданному числу c нечетких α-кластеров, но и значение порога сходства α.

D-AFC(c)-алгоритм представляет собой базовую версию кластер-процедуры, в работе [9] предлагается его модификация, использующая аппарат частичного обучения. В силу этого указанная модификация получила обозначение D-AFC-PS(c)-алгоритма (от англ. partial supervision –

частичное обучение). В работе [10] предложен D-PAFC-алгоритм, строящий так называемое главное распределение по априори неизвестному наименьшему числу c полностью разделенных нечетких α-кластеров. Кроме того, в [11] предложены модификации D-AFC(c)-алгоритма, использующие транзитивное замыкание $\breve{T}$ нечеткой толерантности T и получившие названия D-AFC-TC-алгоритма, D-PAFC-TC-алгоритма и D-AFC-TC(α*)-алгоритма, матрицей исходных данных для которых является матрица нормированных данных «объект – признак». Следует указать, что эвристические возможностные кластер-процедуры, использующие транзитивное замыкание $\breve{T}$ нечеткой толерантности T, не требуют априорного задания числа c полностью разделенных нечетких α-кластеров в искомом распределении $R^*(X)$.

2. Обучение антецедентов нечетких правил на основе результатов кластеризации

В работе [4] предложен метод извлечения нечетких правил, основанный на обработке данных обучающей выборки, которые представлены матрицей «объект – признак» $X_{n\times m} = [\hat{x}_i^t]$, $i = 1,\ldots,n$, $t = 1,\ldots,m$, где n – число объектов классифицируемого множества $X = \{x_1,\ldots,x_n\}$, а m – размерность признакового пространства, с помощью эвристических алгоритмов возможностной кластеризации. Необходимо отметить, что система нечеткого вывода может генерироваться на основе результатов кластеризации, полученных с помощью любого из вышеуказанных алгоритмов. Однако поскольку операция транзитивного замыкания искажает геометрию исходных данных, так что соответствующие алгоритмы оказываются удобным средством лишь для проведения разведочного анализа данных, генерирование системы нечеткого вывода лучше проводить на основе результатов, полученных с помощью реляционных алгоритмов. В связи с этим дальнейшее изложение метода извлечения нечетких правил без нарушения общности будет производиться исходя из предположения, что данные обрабатываются D-AFC(c)-алгоритмом.

Сущность предложенного в [4] метода заключается в следующем. Нечеткие множества B_l^t, $l \in \{1,\ldots,c\}$, $t \in \{1,\ldots,m\}$, условных частей продукционных правил вида (1) и (2) с, в общем случае, трапециевидными функциями принадлежности $\gamma_{B_l^t}(\hat{x}^t)$ могут быть представлены в параметрической форме $B_l^t = (\underline{a}_{(l)}^t, \underline{m}_{(l)}^t, \overline{m}_{(l)}^t, \overline{a}_{(l)}^t)$ (рис. 1), где $\underline{m}_{(l)}^t$ – нижнее модальное значение нечеткого интервала B_l^t; $\overline{m}_{(l)}^t$ – его верхнее модальное значение, а величины $(\underline{m}_{(l)}^t - \underline{a}_{(l)}^t)$ и $(\overline{a}_{(l)}^t - \overline{m}_{(l)}^t)$ представляют собой значения левого и правого коэффициентов нечеткости соответственно [12].

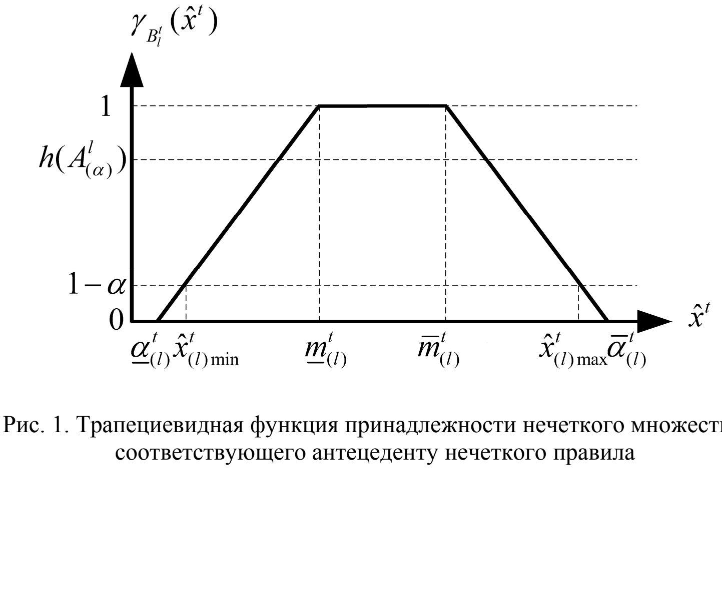


Рис. 1. Трапециевидная функция принадлежности нечеткого множества, соответствующего антецеденту нечеткого правила

Необходимо отметить, что в случае совпадения нижнего и верхнего модальных значений, $m^t_{(l)} = \underline{m}^t_{(l)} = \overline{m}^t_{(l)}$, нечеткий интервал B^t_l будет представлять собой треугольное нечеткое число $B^t_l = (\underline{a}^t_{(l)}, m^t_{(l)}, \overline{a}^t_{(l)})$ [12]. При построении условных частей нечетких правил задача заключается в определении соответствующих параметров нечетких множеств B^t_l, $l \in \{1,\ldots,c\}$, $t \in \{1,\ldots,m\}$.

После предварительной обработки матрицы исходных данных $X_{n\times m} = [\hat{x}^t_i]$, $i = 1,\ldots,n$, $t = 1,\ldots,m$, заключающейся в нормировке значений $\hat{x}^t_i$, к примеру, с помощью нормализации

$$x^t_i = \frac{\hat{x}^t_i}{\max\limits_i \hat{x}^t_i}, \tag{14}$$

так что каждая строка x^t_i матрицы $X_{n\times m} = [x^t_i]$, $i = 1,\ldots,n$, $t = 1,\ldots,m$, может интерпретироваться как нечеткое множество на универсуме признаков с соответствующей функцией принадлежности $x^t_i = \mu_{x_i}(x^t) \in [0,1]$, $i = 1,\ldots,n$, $t = 1,\ldots,m$, и применения к полученному множеству n нечетких множеств какого-либо расстояния между нечеткими множествами [13], например квадрата нормализованного евклидова расстояния

$$\varepsilon(x_i, x_j) = \frac{1}{m}\sum_{t=1}^{m}\Big(\mu_{x_i}(x^t) - \mu_{x_j}(x^t)\Big)^2, \tag{15}$$

получается матрица нечеткого отношения несходства $I = [\mu_I(x_i, x_j)]$, $i, j = 1,\ldots,n$. После применения к ней, в свою очередь, операции дополнения

$$\mu_T(x_i, x_j) = 1 - \mu_I(x_i, x_j), \ \forall i, j = 1,\ldots,n, \tag{16}$$

строится матрица нечеткой толерантности $T = [\mu_T(x_i, x_j)]$, $i, j = 1,\ldots,n$, на $X = \{x_1,\ldots,x_n\}$.

Для построения базы правил системы нечеткого вывода типа Мамдани [14] данные об объектах исследуемой совокупности, представленные в форме матрицы нечеткой толерантности $T = [\mu_T(x_i, x_j)]$, $i, j = 1,\ldots,n$, обрабатываются D-AFC(c)-алгоритмом для заданного числа классов c, и для каждого нечеткого α-кластера $A^l_{(\alpha)}$, $l \in \{1,\ldots,c\}$, полученного распределения $R^*(X) = \{A^1_{(\alpha)},\ldots,A^c_{(\alpha)}\}$ выделяются его ядро $K(A^l_{(\alpha)})$ и носитель A^l_α. Кроме того, как указывалось выше, дополнительным результатом классификации является значение порога сходства $\alpha \in (0,1]$, соответствующего полученному результату классификации $R^*(X)$. Интервал $[\hat{x}^t_{(l)\min}, \hat{x}^t_{(l)\max}]$ значений каждого признака $\hat{x}^t$, $t \in \{1,\ldots,m\}$, вычисляется для носителя A^l_α каждого нечеткого α-кластера $A^l_{(\alpha)}$, $l \in \{1,\ldots,c\}$. В частности, значение $\hat{x}^t_{(l)\min}$, $t \in \{1,\ldots,m\}$ может быть получено по формуле

$$\hat{x}^t_{(l)\min} = \min_{x_i \in A^l_\alpha} \hat{x}^t, \ \forall t \in \{1,\ldots,m\}, \ \forall l \in \{1,\ldots,c\}, \tag{17}$$

а значение $\hat{x}^t_{(l)\max}$, $t \in \{1,\ldots,m\}$, – по формуле

$$\hat{x}^t_{(l)\max} = \max_{x_i \in A^l_\alpha} \hat{x}^t, \ \forall t \in \{1,\ldots,m\}, \ \forall l \in \{1,\ldots,c\}. \tag{18}$$

В свою очередь, параметр $\underline{a}^t_{(l)}$ вычисляется путем решения линейных уравнений при условиях

$$\gamma_{B^t_l}(\hat{x}^t_{(l)\min}) = (1-\alpha),\ \gamma_{B^t_l}(\underline{a}^t_{(l)}) = 0, \tag{19}$$

а параметр $\overline{a}^t_{(l)}$ – при условиях

$$\gamma_{B^t_l}(\hat{x}^t_{(l)\max}) = (1-\alpha),\ \gamma_{B^t_l}(\overline{a}^t_{(l)}) = 0. \tag{20}$$

Значение $\underline{\hat{x}}^t_{(l)}$ вычисляется для всех типичных точек $\tau^l_e \in K(A^l_{(\alpha)})$ нечеткого α-кластера $A^l_{(\alpha)}$, $l \in \{1,\ldots,c\}$, в соответствии с выражением

$$\underline{\hat{x}}^t_{(l)} = \min_{\tau^l_e \in K(A^l_{(\alpha)})} \hat{x}^t,\ \forall e \in \{1,\ldots,|l|\}, \tag{21}$$

а значение $\hat{\overline{x}}^t_{(l)}$ – по формуле

$$\hat{\overline{x}}^t_{(l)} = \max_{\tau^l_e \in K(A^l_{(\alpha)})} \hat{x}^t,\ \forall e \in \{1,\ldots,|l|\}. \tag{22}$$

Таким образом, параметр $\underline{m}^t_{(l)}$ может быть вычислен путем решения линейных уравнений исходя из условий

$$\gamma_{B^t_l}(\underline{\hat{x}}^t_{(l)}) = h(A^l_{(\alpha)}),\ \gamma_{B^t_l}(\underline{m}^t_{(l)}) = 1, \tag{23}$$

а параметр $\overline{m}^t_{(l)}$ вычисляется аналогичным образом при условиях

$$\gamma_{B^t_l}(\hat{\overline{x}}^t_{(l)}) = h(A^l_{(\alpha)}),\ \gamma_{B^t_l}(\overline{m}^t_{(l)}) = 1, \tag{24}$$

где $h(A^l_{(\alpha)}) = \sup_{x_i \in A^l_\alpha} \mu_{A^l_{(\alpha)}}(x_i)$ – высота соответствующего нечеткого α-кластера $A^l_{(\alpha)} \in R^*(X)$. Следует указать, что необходимость учета высоты $h(A^l_{(\alpha)})$ каждого нечеткого α-кластера при построении функций принадлежности $\gamma_{B^t_l}(\hat{x}^t)$ (см. рис. 1) обусловлена возможностью обработки D-AFC(c)-алгоритмом матриц слабых толерантностей, которые представляют структуру сходства динамических объектов, т. е. объектов, признаки которых принимают различные значения в различные моменты времени [15] либо представляют собой интервалы значений. Полученные в результате классификации нечеткие α-кластеры $A^l_{(\alpha)}$, $l \in \{1,\ldots,c\}$, будут представлять собой субнормальные нечеткие множества и именоваться слабыми нечеткими α-кластерами [8]. Очевидно, что если матрица исходных данных $X_{n\times m} = [\hat{x}^t_i]$, $i = 1,\ldots,n$, $t = 1,\ldots,m$, будет представлять собой обычную матрицу вида «объект – признак», где каждому признаку x^t, $t = 1,\ldots,m$, соответствует только одно значение, все полученные нечеткие α-кластеры будут представлять собой нормальные нечеткие множества, т. е. $h(A^l_{(\alpha)}) = 1$, $\forall A^l_{(\alpha)} \in R^*(X)$, и именоваться сильными нечеткими α-кластерами [8].

Таким образом, число входных переменных $\hat{x}^t$ генерируемых правил вида (1) соответствует размерности признакового пространства $I^m(X)$, а число выходных переменных y_l, которым в соответствие могут быть поставлены метки классов, как и число продукционных правил, эквивалентно числу нечетких α-кластеров $A^l_{(\alpha)}$, $l \in \{1,\ldots,c\}$, полученного распределения $R^*(X)$. Необходимо указать, что правила построенной вышеизложенным способом базы нечетких продукций будут иметь вид (1) как в случае, когда нечеткие α-кластеры $A^l_{(\alpha)}$, $l \in \{1,\ldots,c\}$, являются частично разделенными, так и в случае полностью разделенных нечетких α-кластеров.

3. Метод обучения консеквентов нечетких продукционных правил

Нечеткие множества C^l_l, $l = 1,\ldots,c$, соответствующие консеквентам нечетких продукционных правил, определяются на интервале значений принадлежности $[0,1]$. В [4] для построения функций принадлежности y_l, $l \in \{1,\ldots,c\}$, предлагается следующий метод: нечеткие множества C^l_l, $l \in \{1,\ldots,c\}$, консеквентов продукционных правил вида (1) представляются в параметрической форме $C^l_l = (\alpha, \underline{\mu}_l, \overline{\mu}_l, 1)$, где α – порог сходства, при котором строится решение $R^*(X)$ задачи классификации, также являющийся, как указывалось выше, результатом работы D-AFC(c)-алгоритма; $\underline{\mu}_l$ – наименьшее значение принадлежностей объектов $x_i \in A^l_\alpha$ соответствующему нечеткому α-кластеру $A^l_{(\alpha)} \in R^*(X)$, $\underline{\mu}_l = \min_{x_i \in A^l_\alpha} \mu_{li}$, а $\overline{\mu}_l$ – наибольшее значение принадлежности объектов $x_i \in A^l_\alpha$ нечеткому α-кластеру $A^l_{(\alpha)} \in R^*(X)$, $l \in \{1,\ldots,c\}$, $\overline{\mu}_l = \max_{x_i \in A^l_\alpha} \mu_{li}$. В связи с этим соответствующие функции принадлежности $\gamma_{C^l_l}(y_l)$ будут иметь трапециевидную форму (рис. 2, *а*). В случае же когда носитель A^l_α некоторого нечеткого α-кластера $A^l_{(\alpha)}$ распределения $R^*(X)$ представляет собой пустое множество, выходной переменной y_l, $l \in \{1,\ldots,c\}$, будет соответствовать пустое нечеткое множество C^l_l, так что $\gamma_{C^l_l}(y_l) = 0$.

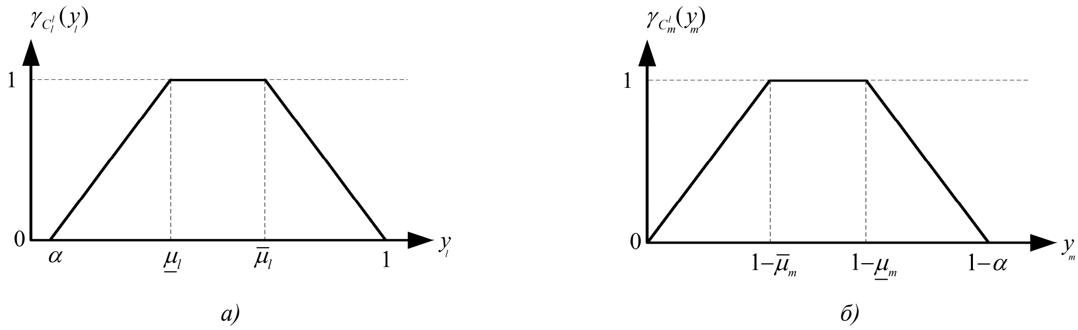


Рис. 2. Трапециевидная функция принадлежности консеквента в случае слабого нечеткого α-кластера: *а)* высокая степень принадлежности классу; *б)* низкая степень принадлежности классу

Подобный подход оказывается приемлемым лишь для случая, когда все нечеткие α-кластеры в искомом распределении $R^*(X)$ оказываются полностью разделенными. В свою очередь, если $A^l_{(\alpha)}$ и $A^m_{(\alpha)}$ – два нечетких α-кластера распределения $R^*(X)$, $l \neq m$, для которых в неравенстве (8) $w \neq 0$, то в соответствии с доказанной в [8] леммой их ядра не пересекаются, т. е. если некоторый объект $x_i \in X$ является типичной точкой $\tau^l_e \in K(A^l_{(\alpha)})$ нечеткого

α-кластера $A^l_{(\alpha)} \in R^*(X)$, то он не будет являться элементом ядра нечеткого α-кластера $A^m_{(\alpha)} \in R^*(X)$. В таком случае нечеткой выходной переменной y_m l-го нечеткого правила вида (1), соответствующей нечеткому α-кластеру $A^m_{(\alpha)}$, можно поставить в соответствие нечеткое множество $C^l_m = (0, 1-\overline{\mu}_m, 1-\underline{\mu}_m, 1-\alpha)$ (рис. 2, *б*).

Необходимо указать, что изображенный на рис. 2 случай является общей иллюстрацией метода обучения консеквентов. К примеру, если в результате обработки данных D-AFC(c)-алгоритмом полученные нечеткие α-кластеры $A^l_{(\alpha)} \in R^*(X)$, $l \in \{1,\ldots,c\}$, будут являться сильными нечеткими α-кластерами, то, очевидно, будет иметь место $\overline{\mu}_l = 1$, так что трапециевидная функция принадлежности $\gamma_{C^l_l}(y_l)$ соответствующего нечеткого множества C^l_l, $l \in \{1,\ldots,c\}$, примет вид, изображенный на рис. 3, *а*. Функция принадлежности $\gamma_{C^l_m}(y_m)$ нечеткого множества C^l_m, $m \in \{1,\ldots,c\}$, $m \neq l$, примет вид, изображенный на рис. 3, *б*.

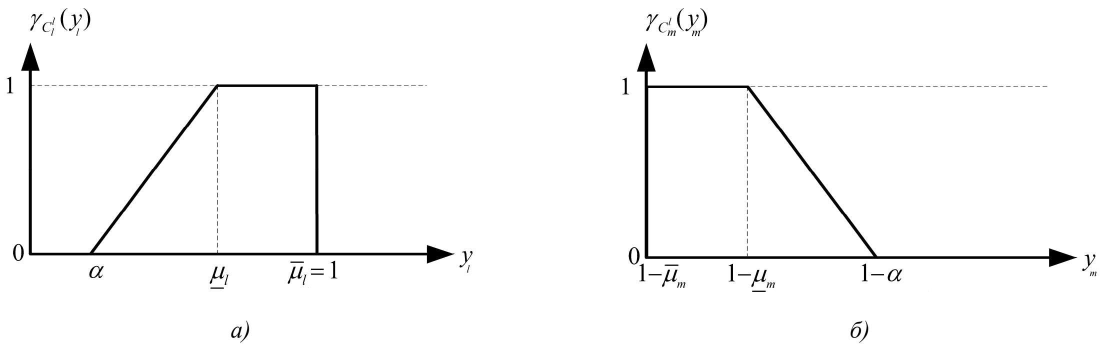


Рис. 3. Трапециевидная функция принадлежности консеквента в случае сильного нечеткого α-кластера: *а)* высокая степень принадлежности классу; *б)* низкая степень принадлежности классу

С содержательной точки зрения функциям принадлежности консеквентов нечетких правил (см. рис. 2, *а* и рис. 3, *а*) можно поставить в соответствие метки «высокая принадлежность», тогда как функции принадлежности консеквентов (см. рис. 2, *б* и рис. 3, *б*) совершенно естественно получают интерпретацию «низкая принадлежность».

4. Иллюстративный пример

Предложенный метод обучения консеквентов нечетких правил целесообразно проиллюстрировать на примере, для чего были выбраны четырехмерные данные Е. Андерсона по 150 ирисам [16]; каждый цветок описывается длиной и шириной лепестка и длиной и шириной чашелистика. Задача классификации заключается в разбиении исследуемой совокупности на три класса – SETOSA, VERSICOLOR и VIRGINICA, в каждом из которых содержится 50 ирисов. При проведении вычислительных экспериментов правила строились на основе результатов кластеризации всей совокупности объектов, т. е. обучающая и тестируемая выборки совпали. В первом эксперименте база нечетких правил строилась с помощью D-AFC(c)-алгоритма.

Для предварительной обработки матрицы исходных данных $X_{150\times4} = [\hat{x}^t_i]$ была выбрана нормализация (14), а в качестве расстояния – квадрат нормализованного евклидова расстояния (15). После обработки полученной матрицы нечеткой толерантности $T_{150\times150} = [\mu_T(x_i, x_j)]$, $i, j = 1,\ldots,150$, для $c = 3$ результатом классификации оказалось распределение $R^*(X)$ по трем полностью разделенным нечетким α-кластерам при значении порога сходства $\alpha = 0{,}9642$. Типичными точками первого нечеткого α-кластера, соответствующего классу SETOSA, оказались объекты x_{95} и x_{106}, типичной точкой второго нечеткого α-кластера, которому в соответствие ставится класс VERSICOLOR, оказался объект x_{98}, а типичной точкой третьего нечеткого α-кластера, соответствующего классу VIRGINICA, – объект x_{73}. Значения принадлежности объектов первому классу обозначены на рис. 4 символом +, второму – символом ■ и третьему – символом □. Некорректно расклассифицированными оказались всего шесть объектов.

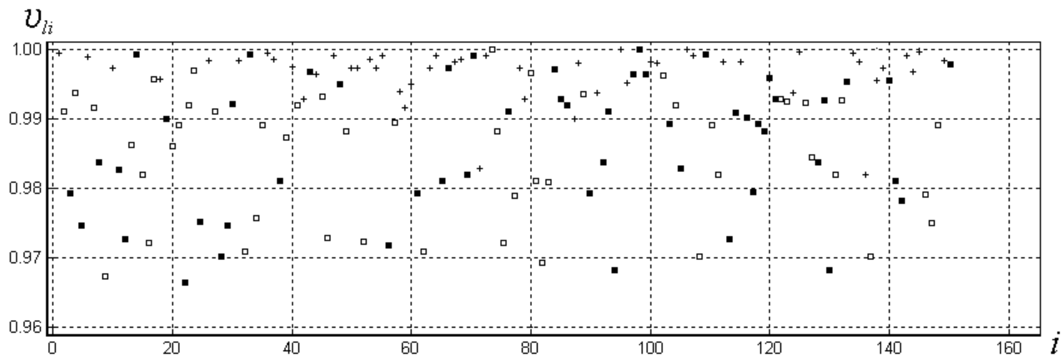


Рис. 4. Значения принадлежности объектов полностью разделенным нечетким кластерам распределения, полученного с помощью D-AFC(c)-алгоритма

После применения предложенного подхода к извлечению нечетких правил на основе результатов кластеризации была сгенерирована система нечеткого вывода (рис. 5), где символами SL, SW, PL и PW обозначены входные переменные, соответствующие признакам $\hat{x}^t$, $t = 1,\ldots,4$, объектов исследуемой совокупности, значения от 1 до 3 соответствуют номеру правила, а консеквентам нечетких правил в качестве меток были поставлены названия соответствующих классов. В результате классификации объектов исследуемой совокупности некорректно расклассифицированными оказались объекты x_5, x_9, x_{25}, x_{56} и x_{90}, а объекты x_{12} и x_{147} оказались расклассифицированными дважды, т. е. принадлежащими классам VERSICOLOR и VIRGINICA с одинаково высокой степенью.

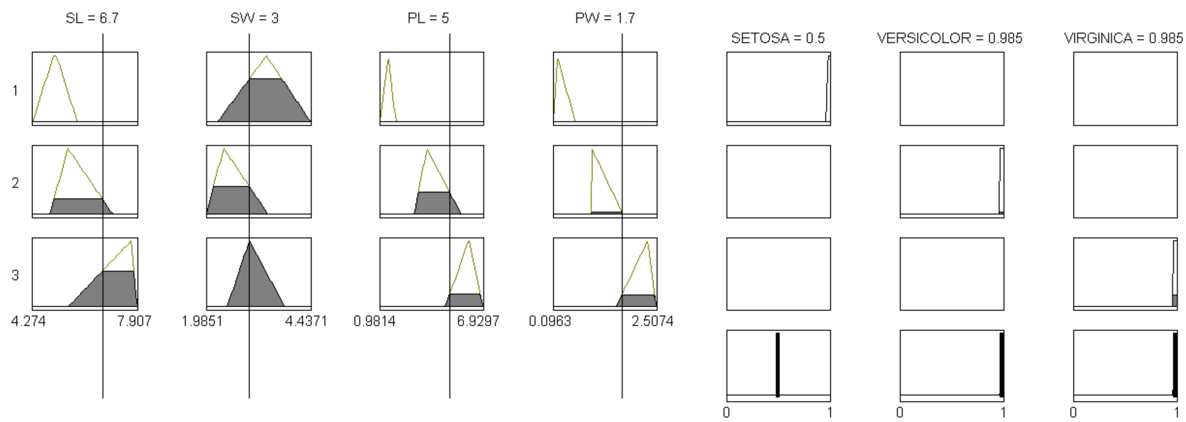


Рис. 5. Классификация объекта исследуемой совокупности, построенной с помощью D-AFC(c)-алгоритма, системой нечеткого вывода типа Мамдани

На рис. 5 приведен пример классификации объекта x_{147}, который, как указывалось выше, оказался принадлежащим классам VERSICOLOR и VIRGINICA с одинаково высокой степенью, что допускает двусмысленность в интерпретации результата классификации. В свою очередь, значение принадлежности указанного объекта классу SETOSA, полученное в результате его классификации с помощью построенной системы нечеткого вывода типа Мамдани, оказалось равным 0,5, что допускает такую интерпретацию, как «неопределенно».

Во втором эксперименте база нечетких правил была сгенерирована на основе результатов кластеризации исходных данных с помощью D-AFC-PS(c)-алгоритма, причем для построения подмножества помеченных объектов была использована методика, предложенная в [17], так что в качестве помеченных объектов были выбраны x_{95}, x_3, x_5 с общим для всех априорным значением принадлежности $y_{l(j)} = 0{,}9929$. После построения матрицы нечеткой толерантности методом, аналогичным рассмотренному выше, и обработки данных с помощью D-AFC-PS(c)-алгоритма результатом классификации оказалось распределение $R^*(X)$ по трем частично разделенным нечетким α-кластерам при значении порога сходства $\alpha = 0{,}9715$. Значения принадлежности объектов первому классу обозначены на рис. 4 символом +, второму – символом ■ и третьему – символом □. Число некорректно классифицированных объектов равно 10.

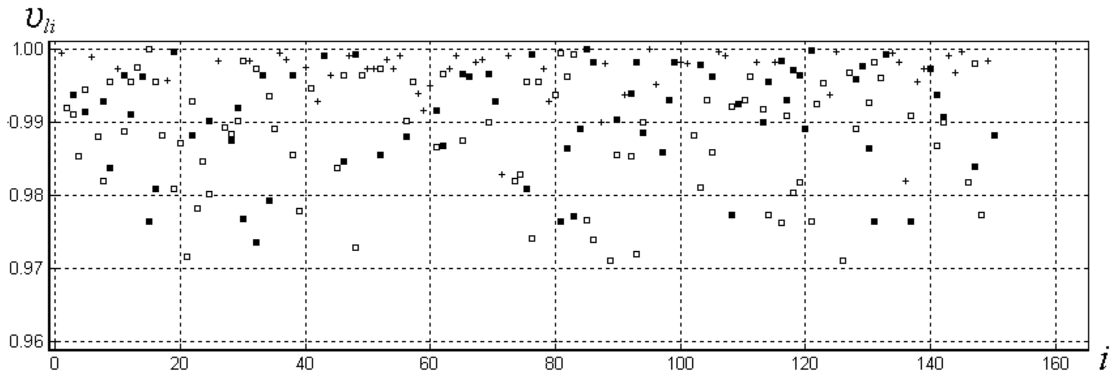


Рис. 6. Значения принадлежности объектов частично разделенным нечетким кластерам распределения, полученного с помощью D-AFC-PS(c)-алгоритма

Из рис. 7 видно, что принадлежность объекта x_{147} классу VIRGINICA оказалась высокой, а классу VERSICOLOR – низкой. Это позволило устранить двусмысленность при интерпретации результата классификации, имевшую место в предыдущем эксперименте.

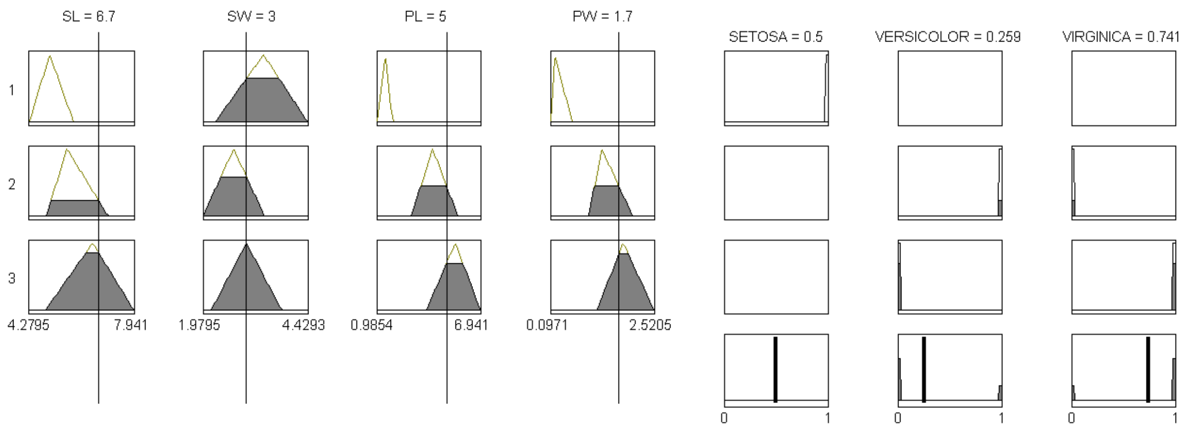
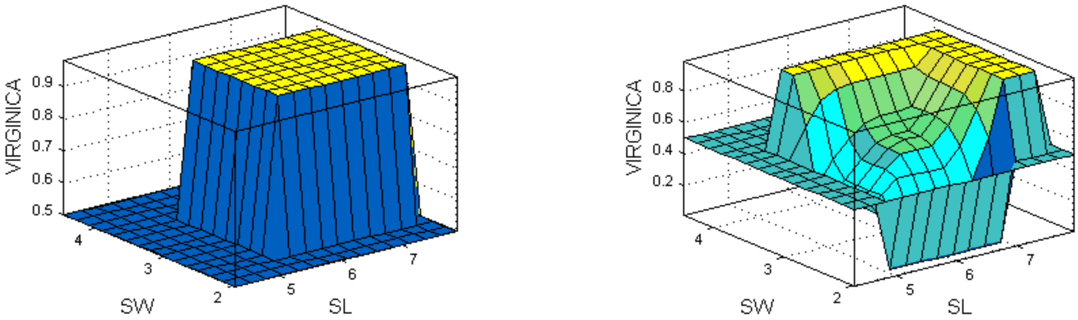


Рис. 7. Классификация объекта исследуемой совокупности, построенной с помощью D-AFC-PS(c)-алгоритма, системой нечеткого вывода типа Мамдани

Различие между консеквентами построенных в результате обоих экспериментов систем нечеткого вывода можно продемонстрировать на примере поверхности нечеткого вывода, визуализирующей зависимость выходных переменных от отдельных входных переменных. На рис. 8 изображены поверхности нечеткого вывода, показывающие зависимость значения выходной переменной VIRGINICA от первых двух входных переменных SL и SW при классификации объекта x_{147} для обоих построенных систем нечеткого вывода.



а) *б)*

Рис. 8. Вид поверхности нечеткого вывода для моделей, разработанных на основе результатов кластеризации обучающих данных: *а)* с помощью D-AFC(c)-алгоритма; *б)* с помощью D-AFC-PS(c)-алгоритма

Анализ поверхностей нечеткого вывода для остальных входных и выходных переменных позволяет сделать вывод об адекватности нечеткой модели, построенной на основе результатов кластеризации исходных данных с помощью D-AFC-PS(c)-алгоритма.

Устойчивость результатов кластеризации, свойственная эвристическому методу возможностной кластеризации, а также алгоритмическая простота предложенного метода быстрого прототипирования базы правил системы нечеткого вывода, с учетом вышеизложенных результатов, позволила разработать соответствующий программный модуль, получивший название FIS Generator [18]. Генерирование правил системы нечеткого вывода выполняется после предварительной обработки данных D-AFC(c)-алгоритмом с ручным выбором в главном окне программы параметров кластеризации, таких, как вид используемого расстояния между нечеткими множествами [13], метод нормировки исходных данных, количество классов, вид используемого в D-AFC(c)-алгоритме критерия качества классификации и значение показателя точности вычислений. На рис. 9 изображен выбор параметров классификации при генерировании системы нечеткого вывода по данным Е. Андерсона.

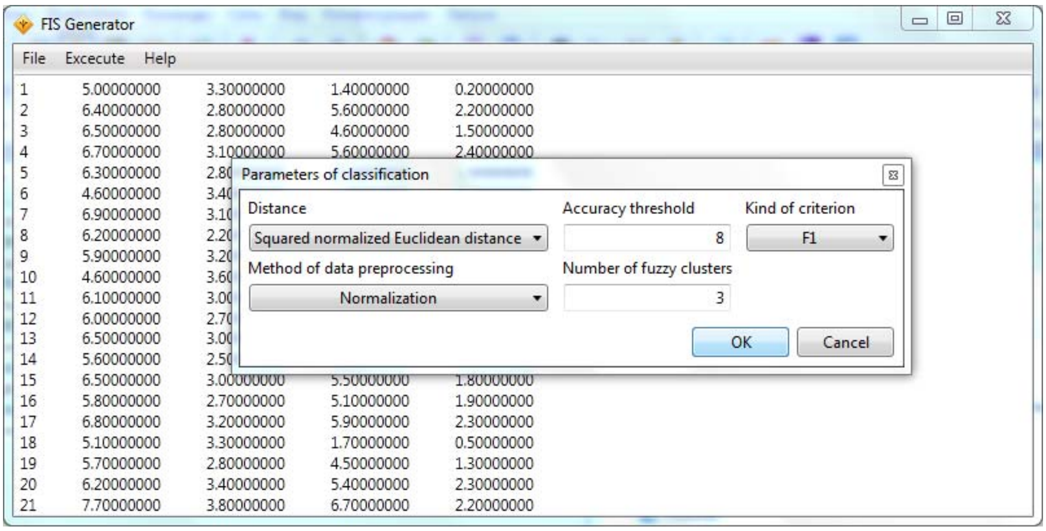


Рис. 9. Вид главного окна программного модуля FIS Generator и диалоговое окно выбора параметров классификации

Результат кластеризации выводится в рабочее окно программы, после чего производится непосредственное генерирование правил прямо в файл, указанный пользователем. Все вычисления выполняются в фоновом потоке и сопровождаются окном, информирующем о степени обработки данных. Генерируемая с помощью разработанного программного модуля FIS Generator система нечеткого вывода представляет собой внешний файл с расширением fis, сохраняемый пользователем с помощью стандартного диалогового окна и запускаемый с помощью среды MATLAB® версии не ниже 7.6.0 [19]. Так как каждое правило идентифицирует объекты определенного класса, то при обучении или переобучении системы нечеткого вывода выходным переменным сгенерированной системы необходимо ставить в соответствие названия классов, для чего пользователю следует содержательно интерпретировать результаты кластеризации.

Простота реализованного метода, интуитивно понятный интерфейс, а также возможность генерирования систем нечеткого вывода в режиме времени, близком к реальному, делают разработанный программный модуль FIS Generator удобным средством для моделирования нечеткого вывода при проектировании систем оценки защищенности критически важных объектов информационно-телекоммуникационной инфраструктуры, таких, как системы военного назначения [20].

Заключение

В статье предложен метод построения функций принадлежности нечетких множеств, соответствующих консеквентам продукционных правил механизма нечеткого вывода типа Мамдани, строящегося на основе результатов обработки обучающей выборки эвристическим D-AFC(c)-алгоритмом возможностной кластеризации. Предложенный метод отличается простотой и содержательной осмысленностью количественной оценки результатов классификации. С учетом предложенного в статье метода обучения консеквентов разработка основы подхода к проектированию базы правил системы нечеткого вывода, базирующегося на обработке данных эвристическим алгоритмом возможностной кластеризации, получает свое завершение.

Анализ результатов вычислительных экспериментов наглядно демонстрирует высокую эффективность предлагаемого подхода в сравнении с традиционным методом проектирования систем нечеткого вывода, основанным на использовании оптимизационных кластер-процедур нечеткой и возможностной кластеризации. Отличия предлагаемого подхода заключаются, прежде всего, в сокращении числа нечетких продукций, формирующих базу правил системы нечеткого вывода, а также в возможности быстрой разработки прототипа системы нечеткого вывода в автоматическом режиме, что является следствием устойчивости результатов обработки данных эвристическим алгоритмом возможностной кластеризации.

Список литературы

1. Борисов, В.В. Нечеткие модели и сети / В.В. Борисов, В.В. Круглов, А.С. Федулов. – М. : Горячая линия – Телеком, 2007. – 284 с.

2. Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition / F. Höppner [et al.]. – Chichester : Wiley Intersciences, 1999. – 289 p.

3. Sugeno, M. A fuzzy-logic-based approach to qualitative modeling / M. Sugeno, T. Yasukawa // IEEE Transactions on Fuzzy Systems. – 1993. – Vol. 1. – P. 7–31.

4. Viattchenin, D.A. Automatic generation of fuzzy inference systems using heuristic possibilistic clustering / D.A. Viattchenin // Journal of Automation, Mobile Robotics and Intelligent Systems. – 2010. – Vol. 4, № 3. – P. 36–44.

5. Viattchenin, D.A. A new heuristic algorithm of fuzzy clustering / D.A. Viattchenin // Control & Cybernetics. – 2004. – Vol. 33. – P. 323–340.

6. Вятченин, Д.А. О возможностной интерпретации значений принадлежности в методе нечеткой кластеризации, основанном на понятии распределения / Д.А. Вятченин // Вести Института современных знаний. – 2008. – № 3. – С. 85–90.

7. Krishnapuram, R. A possibilistic approach to clustering / R. Krishnapuram, J.M. Keller // IEEE Transactions on Fuzzy Systems. – 1993. – Vol. 1. – P. 98–110.

8. Вятченин, Д.А. Виды нечетких α-кластеров / Д.А. Вятченин // Вести Института современных знаний. – 2008. – № 4. – С. 95–101.

9. Viattchenin, D.A. A direct algorithm of possibilistic clustering with partial supervision / D.A. Viattchenin // Journal of Automation, Mobile Robotics and Intelligent Systems. – 2007. – Vol. 1. – P. 29–38.

10. Viattchenin, D.A. An algorithm for detecting the principal allotment among fuzzy clusters and its application as a technique of reduction of analyzed features space dimensionality / D.A. Viattchenin // Journal of Information and Organizational Sciences. – 2009. – Vol. 33. – P. 205–217.

11. Вятченин, Д.А. Прямые алгоритмы нечеткой кластеризации, основанные на операции транзитивного замыкания, и их применение к обнаружению аномальных наблюдений / Д.А. Вятченин // Искусственный интеллект. – 2007. – № 3. – С. 205–216.

12. Дюбуа, Д. Теория возможностей. Приложения к представлению знаний в информатике / Д. Дюбуа, А. Прад ; пер. В.Б. Тарасова. – М. : Радио и связь, 1990. – 288 с.

13. Кофман, А. Введение в теорию нечетких множеств / А. Кофман ; пер. В.Б. Кузьмина. – М. : Радио и связь, 1982. – 432 с.

14. Mamdani, E.H. An experiment in linguistic synthesis with a fuzzy logic controller / E.H. Mamdani, S. Assilian // International Journal of Man-Machine Studies. – 1975. – Vol. 7. – P. 1–13.

15. Viattchenin, D.A. An outline for a heuristic approach to possibilistic clustering of the three-way data / D.A. Viattchenin // Journal of Uncertain Systems. – 2009. – Vol. 3. – P. 64–80.

16. Anderson, E. The irises of the Gaspe Peninsula / E. Anderson // Bulletin of the American Iris Society. – 1935. – Vol. 59. – P. 2–5.

17. Damaratski, A. A novel technique of partial supervision for a heuristic algorithm of possibilistic clustering / A. Damaratski, A. Juodelis // Pattern Recognition and Information Processing : Proceedings of the 11th International Conference PRIP'2011, Minsk, 18–20 May 2011 / BSUIR ; ed. by R. Sadykhov [et al.]. – Minsk, 2011. – P. 121–126.

18. Доморацкий, А.В. Модуль генерирования систем нечеткого вывода / А.В. Доморацкий // Управление в социальных и экономических системах: материалы XX Междунар. науч.-практ. конф., Минск, 20 мая 2011 г. / Минский ин-т управления ; редкол. : Н.В. Суша [и др.]. – Минск, 2011. – С. 239–240.

19. Леоненков, А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А.В. Леоненков. – СПб. : БХВ-Петербург, 2003. – 736 с.

20. Анищенко, В.В. Оценка защищенности информационных систем военного назначения на основе применения систем нечеткого вывода / В.В. Анищенко, Д.А. Вятченин, А.М. Криштофик // Комплексная защита информации : материалы XVI науч.-практ. конф., Гродно, 17–20 мая 2011 г. / БелГИСС ; под общ. ред. А.Н. Курбацкого. – Минск, 2011. – С. 159–161.

Поступила 29.06.11

[1]*Объединенный институт проблем информатики НАН Беларуси,*
Минск, Сурганова, 6
e-mail: anishch@newman.bas-net.by

[2]*НИРУП «Геоинформационные системы» НАН Беларуси,*
Минск, Сурганова, 6

U.V. Anishchanka, D.A. Viattchenin, A.V. Damaratski

TRAINING CONSEQUENTS OF FUZZY RULES CONSTRUCTED BASED ON HEURISTIC POSSIBILISTIC CLUSTERING

A method of constructing of membership functions which correspond to consequents of fuzzy rules constructed based on results of the training data set processing by a heuristic algorithm of possibilistic clustering is proposed. The efficiency of the method is demonstrated on an illustrative example.