ISSN 1512-1232

УДК 681.3

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ НЕЧЕТКОЙ ИДЕНТИФИКАЦИИ ДЛЯ СИНТЕЗА ТОЧНЫХ, КОМПАКТНЫХ И ИНТЕРПРЕТАБЕЛЬНЫХ БАЗ ЗНАНИЙ

Штовба С.Д.[1], Мазуренко В.В., Тылец Р.О.[1]

[1] Кафедра компьютерных систем управления, Винницкий национальный технический университет, Украина, Винница

Аннотация

Рассматривается задача синтеза из результатов наблюдений за многофакторными зависимостями их математических моделей в форме нечетких баз знаний. В работе приведены основные критерии оценки качества нечетких баз знаний (точность, компактность и интерпретабельность) и поставлены задачи структурной и параметрической идентификации по этому набору критериев. Разработана информационная технология, обеспечивающая синтез точных, компактных и интерпретабельных нечетких баз знаний. Приведены примеры применения этой информационной технологии для решения 7 задач идентификации с реальными экспериментальными данными.

Ключевые слова: *нечеткая идентификация, нечеткий база знаний, интепретабельность, компактность, нечеткий вывод, генетические алгоритмы.*

1. ВВЕДЕНИЕ

В технике, агрономии, экономике, экологии, медицине, биологии, политике, спорте и в других областях накоплены результаты наблюдений за различными многофакторными зависимостями. Задача идентификации состоит в построении по этим наблюдениям математических моделей рассматриваемых многофакторных зависимостей. Идентификацию зависимостей осуществляют в 2 этапа [1]. На первом этапе – этапе структурной идентификации, определяют структуру (тип, вид) модели исследуемой зависимости. На втором этапе – этапе параметрической идентификации подстраивают модель, изменяя ее параметры таким образом, чтобы ее поведение наилучшим образом описало экспериментальные данные. Этот этап наиболее формализованный и, как правило, сводится к решению задачи оптимизации.

В последнее время идентификацию сложных зависимостей в условиях неопределенности все чаще проводят с помощью нечетких баз знаний [2]. Нечеткой базой знаний называется совокупность продукционных правил <Если – то>, которые описывают взаимосвязь между входами и выходами с использованием нечетких термов "Низкий", "Средний", "Высокий" и т.п. При этом термы формализуют нечеткими множествами [3]. Идентификацию, в результате которой получают модель зависимости в форме нечеткой базы знаний, обычно называют нечеткой.

При использовании нечетких баз знаний этап структурной идентификации заключается в определении входных и выходных переменных моделей, иерархической организации входных переменных, формировании терм-множеств лингвистических переменных и описании зависимости нечеткими продукционными правилами [4]. Первые две процедуры являются общими для любого метода идентификации. Последние две процедуры являются специфическими для нечеткой идентификации. В результате структурной идентификации получаем грубую модель, которая в общих чертах описывает исследуемую зависимость. На этапе параметрической идентификации настраивают модель – изменяют функции принадлежности нечетких термов, а также веса правил в нечетких базах знаний Мамдани и коэффициенты в заключениях правил базы знаний Сугено [2].

Качество нечетких баз знаний оценивают по критериям точности, компактности и интерпретабельности [5–9]. Целью статьи является разработка информационной технологии, обеспечивающей синтез из экспериментальных данных точных, компактных и интерпретабельных нечетких баз знаний. Отличительной особенностью технологии является следующая четырехэтапная схема функционирования: 1) генерация списка адекватных правил; 2) селекция правил; 3) редукция антецедентов отобранных правил; 4) параметрическая настройка нечеткой базы знаний. Первые 3 этапа соответствуют структурной идентификации, а четвертый этап – параметрической. Критерии точности и интерпретабельности задействованы на первом и четвертом этапах, а критерии точности и компактности – на втором и третьем. Также предлагаются новые модели сохранения интерпретабельности нечетких баз знаний во время параметрической идентификации. Предлагаемая информационная технология протестирована на серии задач из репозитария по машинному обучению Калифорнийского университета в Ирвине [10]. Эксперименты проводились для нечетких баз знаний Мамдани.

2. НЕЧЕТКАЯ БАЗА ЗНАНИЙ МАМДАНИ

В базе знаний Мамдани антецеденты и консеквенты правил заданы нечеткими множествами. Пример такого правила для прогнозирования потребления природного газа: если температура наружного воздуха – низкая и теплоизоляция дома – низкая, то расход газа – высокий. Базу знаний Мамдани можно трактовать как разбиение факторного пространства на зоны с нечеткими границами, в каждой из которых функция отклика принимает нечеткое значение. Количество нечетких зон равно числу правил.

Нечеткую базу знаний Мамдани, связывающую входы $X=(x_1, x_2, \ldots, x_n)$ с выходом y исследуемой зависимости, запишем следующим образом [2]:

$$\textit{Если}\ \left(x_1=\tilde{a}_{i1}\ \textit{и}\ x_2=\tilde{a}_{i2}\ \textit{и} \ldots \textit{и}\ x_n=\tilde{a}_{in}\right),\ \textit{то}\ y=\tilde{d}_i,\ \textit{с весом}\ w_i,\ i=\overline{1,N}, \qquad (1)$$

где $\tilde{a}_{ij}$ – нечеткий терм, который используется для лингвистической оценки фактора x_j в i-ом правиле, $i=\overline{1,N}$, $j=\overline{1,n}$;

N – количество правил базы знаний;

$\tilde{d}_i$ – консеквент i-го правила в форме нечеткого терма;

$w_i \in [0;1]$ – вес i-го правила, который отражает уверенность эксперта в его достоверности.

Каждый нечеткий терм $\tilde{a}_{ij}$ выбирается из терм-множества $\left\{l_{j1}, l_{j2}, ..., l_{jk_j}\right\}$, причем $k_j \le N$, $j=\overline{1,n}$. Возможны и короткие правила, в антецедентах которых некоторые переменные могут принимать произвольные лингвистические значения. В этом случае, соответствующая переменная исключается из антецедента правила или ей присваиваться терм *“Don’t care”*, функция принадлежности которого принимает единичные значения на всем входном диапазоне.

ISSN 1512-1232

Обозначим функцию принадлежности термов $\tilde{a}_{ij}$ через $\mu_{ij}(x_j)$, $x_j \in [\underline{x}_j; \overline{x}_j]$, а термов $\tilde{d}_i$ через $\mu_{d_i}(y)$, $y \in [\underline{y}; \overline{y}]$. Для формализации нечетких множеств воспользуемся гауссовой функцией принадлежности:

$$\mu(x) = \exp\left(-\frac{(x-b)^2}{2c^2}\right), \tag{2}$$

где b – ядро нечеткого множества;

$c > 0$ – коэффициент концентрации (рис. 1).

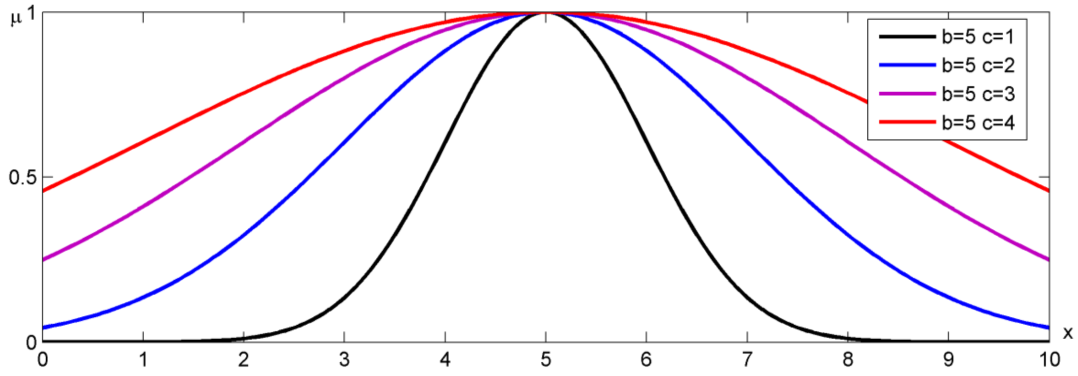

Рисунок 1 – Графики гауссовых функций принадлежности

Логический вывод для входного вектора $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ осуществляется согласно рис. 2. Вначале по формуле (2) рассчитываются степени принадлежности $\mu_{ij}(x_j^*)$ входных значений x_j^* нечетким термам $\tilde{a}_{ij}$ из базы знаний (1). В результате получаем входной вектор в форме бинечеткого множества:

$$\tilde{\tilde{X}}^* = \left(\frac{\mu_{i1}(x_1^*)}{\tilde{a}_{i1}}, \frac{\mu_{i2}(x_2^*)}{\tilde{a}_{i2}}, \ldots, \frac{\mu_{in}(x_n^*)}{\tilde{a}_{in}}\right), \quad i = \overline{1, N}. \tag{3}$$

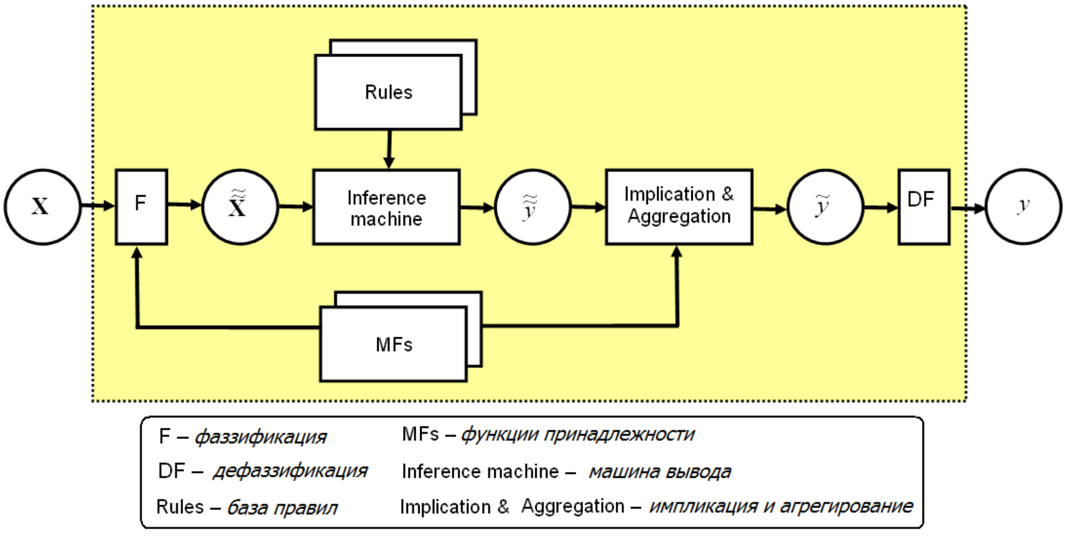

Рисунок 2 – Логический вывод по нечеткой базе знаний Мамдани

Особенность бинечеткого множества заключается в том, что каждый элемент его носителя задан нечетким множеством [11]. В бинечетком множестве (3) носитель задан

ISSN 1512-1232

термами антецедентов правил из базы знаний (1). Степень принадлежности входного вектора к консеквентам правил рассчитывается следующим образом:

$$\mu_{d_i}(X^*) = w_i \cdot \min\left(\mu_{i1}(x_1^*), \mu_{i2}(x_2^*), ..., \mu_{in}(x_n^*)\right), \quad i = \overline{1, N}. \tag{4}$$

В результате получаем выходное бинечеткое множество $\tilde{y}^* = \left(\frac{\mu_{d_1}(X^*)}{\tilde{d}_1}, \frac{\mu_{d_2}(X^*)}{\tilde{d}_2}, ..., \frac{\mu_{d_N}(X^*)}{\tilde{d}_N}\right)$, носителем которого является множество нечетких термов $\left\{\tilde{d}_1, \tilde{d}_2, ..., \tilde{d}_N\right\}$. Переход на носитель $\left[\underline{y}; \overline{y}\right]$ осуществляется посредством импликации:

$$\tilde{d}_i^* = \mathrm{imp}\left(\tilde{d}_i,\ \mu_{d_i}(X^*)\right), \quad i = \overline{1, N}, \tag{5}$$

где imp обозначает импликацию, которую реализуют операцией минимума.

Геометрическая интерпретация этой импликации заключается в срезании графика функции принадлежности $\mu_{d_i}(y)$ на уровне $\mu_{d_i}(X^*)$, $i = \overline{1, N}$ (рис. 3).

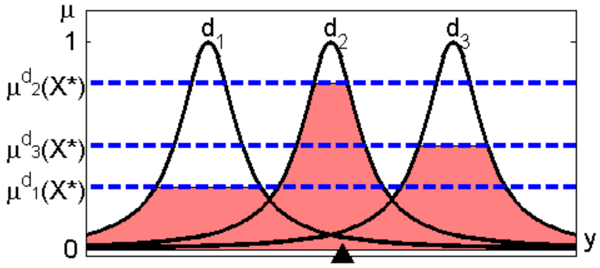

Рисунок 3 – Импликация, агрегирование и дефаззификация в алгоритме Мамдани

Результирующее нечеткое множество получают агрегированием нечетких множеств (5):

$$\tilde{y}^* = agg\left(\tilde{d}_1^*, \tilde{d}_2^*, ..., \tilde{d}_N^*\right) \tag{6}$$

где agg обозначает агрегирование, которое реализуют операцией максимума над функциями принадлежности.

Четкое значение выхода y^*, соответствующее входному вектору X^*, находят посредством дефаззификации нечеткого множества $\tilde{y}^*$ по методу центра тяжести:

$$y^* = \frac{\int_{\underline{y}}^{\overline{y}} y \cdot \mu_y(y)\, dy}{\int_{\underline{y}}^{\overline{y}} \mu_y(y)\, dy}. \tag{7}$$

Для предотвращения эффекта сужения выходного диапазона через дефаззификацию по центру тяжести расширим носитель нечетких термов переменной y в соответствии с рекомендациями из [12].

ISSN 1512-1232

3. КРИТЕРИИ КАЧЕСТВА НЕЧЕТКОЙ ИДЕНТИФИКАЦИИ

Будем считать известными выборку из M пар экспериментальных данных о влиянии факторов $X=\left(x_1,x_2,\ldots,x_n\right)$ на выход y исследуемой зависимости:

$$\left(X_r,\, y_r\right),\ r=\overline{1,M}\,, \tag{8}$$

где X_r – входной вектор в r-ой строчке выборки;

y_r – соответствующее выходное значение.

Задача нечеткой идентификации состоит в том, чтобы на основе данных (8) найди базу знаний (1), которая обеспечит наилучшее качество модели исследуемой зависимости. Качество нечеткой модели будем оценивать по критериям точности, компактности и интерпретабельности. Точность и компактность модели – это традиционные критерии оценки качества идентификации. Для нечетких моделей важна и интепретабельность, т.е. способность объяснить заказчикам на естественном языке как функционирует модель. Заказчики модели – врачи, экономисты, агрономы, финансисты, педагоги, политики и другие специалисты с невысокой математической квалификацией, не доверяют принятие важных решений набору непонятных им чисел. Объяснение процесса моделирования должно быть простым, буквально “на пальцах”. Именно возможность содержательной интерпретации является важным преимуществом нечетких моделей, позволяющим им конкурировать с другими технологиями идентификации сложных зависимостей.

Точность нечеткой модели $F(X)$ опишем средней квадратичной ошибкой

$$RMSE=\sqrt{\frac{1}{M}\sum_{r=\overline{1,M}}\left(y_r-F\left(X_r\right)\right)^2} \tag{9}$$

или нормированной ошибкой

$$NRMSE=\frac{RMSE}{\overline{y}-\underline{y}}\,. \tag{10}$$

Для оценки компактности модели воспользуемся такими показателями: N – суммарное количество правил и A – суммарная длина всех антецедентов, т. е. количество термов во всех антецедентах.

Под интерпретабельностю понимается возможность содержательного объяснения структуры и параметров модели. Нечеткая модель является интерпретабельной, если выполняются следующие условия [12]:

- база знаний не является противоречивой или избыточной, т. е. не содержит правил с одинаковыми антецедентами;
- база знаний согласована с количеством термов, т. е. каждый терм фигурирует хотя бы в одном нечетком правиле;
- для произвольного входного вектора на выходе получаем непустое нечеткое множество;
- изолированно функция принадлежности каждого терма содержательно интерпретируется, т. е. соответствующее нечеткое множество является нормальным и выпуклым [2];
- каждое терм-множество содержательно интерпретируется.

Терм-множество является интерпретируемым, если размещение нечетких множеств на носителе не противоречит здравому смыслу. Например, терм “Средний” находится между термами “Низкий” и “Высокий”. При этом, высоты пересечения графиков функций принадлежности термов “Низкий” и “Высокий” ниже, чем для термов “Низкий” и “Средний”, а также ниже, чем для термов “Средний” и “Высокий”. Пример неинтерпретабельного терм-множества приведен на рис. 4. На этом рисунке показаны нарушения интерпретабельности следующих типов:

a) сильная схожесть функций принадлежности соседних нечетких множеств “Низкий” и “Ниже среднего”, что может внести противоречие в базу знаний;
b) потеря линейной упорядоченности терм-множества через разную размазанность функций принадлежностей – на интервале (65; 82) нечеткое множество “Средний” больше нечеткого множества “Выше среднего”, а на интервале (0; 3) нечеткое множество “Ниже среднего” больше нечеткого множества “Низкий”, хотя должно быть наоборот;
c) неинтерпретабельность крайнего терма – уменьшение значений переменной x от 8 до 0 снижает степень принадлежности нечеткому терму “Низкий”, хотя должно быть наоборот;
d) неполное накрытие нечеткими множествами интервала возможных значений входной переменной – числа из (84; 88) не принадлежат нечетким множествам, поэтому при любом наборе правил в базе знаний результатом логического вывода для $x \in (84; 88)$ будет пустое нечеткое множество.

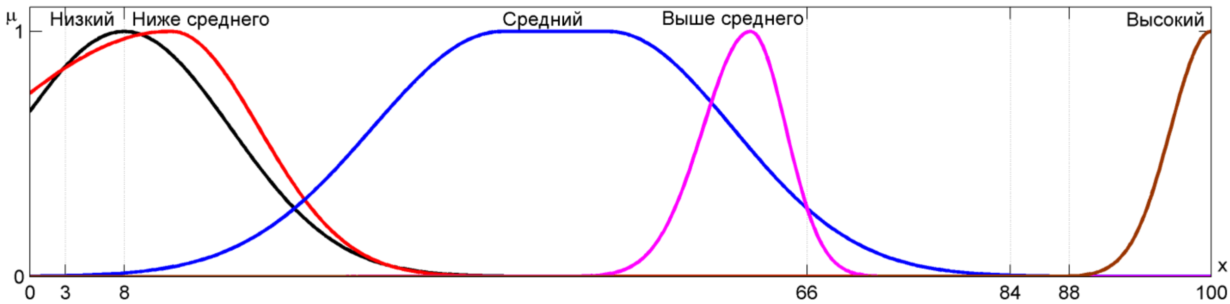

Рисунок 4 – Типовые нарушения интерпретабельности терм-множества

Схематично интерпретируемость терм-множества можно записать так: “Низкий” < “Средний” < “Высокий”. В более формализованном виде интерпретируемость терм-множества свяжем с выполнением следующих трех условий.

1. Количество термов должно быть не слишком большим, чтобы эксперт мог каждому нечеткому множеству сопоставить осмысленную лингвистическую оценку. В [13] мощность терм-множества ограничивают сверху “магическим” числом 7 ± 2 [14]. Исходя из нашего опыта проектирования нечетких систем целесообразнее использовать терм-множества с мощностью до 7.

2. Нечеткие множества различных термов не должны быть эквивалентными или почти эквивалентными. Отсюда, графики функций принадлежности соседних термов, например, “Низкий” и “Ниже среднего”, должны отличаться визуально.

3. Терм-множество должно быть линейно упорядоченным. Обозначим через l_i – терм с номером i, $i = \overline{1, K}$. Этот терм используется для лингвистической оценки переменной x на интервале $[\underline{x}; \overline{x}]$. Пусть терму l_i соответствует нечеткое множество $\tilde{l}_i$ с функцией принадлежности $\mu_i(x)$. Условие линейной упорядоченности терм-множества $\{l_1, l_2, ..., l_K\}$ запишем так:

ISSN 1512-1232

$$\begin{cases} \forall x \in \left[\underline{x}, \sup\left(\mathrm{core}(\tilde{l}_i)\right)\right] & : \mu_i(x) \geq \mu_{i+1}(x) \\ \exists x \in \left[\underline{x}, \sup\left(\mathrm{core}(\tilde{l}_i)\right)\right] & : \mu_i(x) > \mu_{i+1}(x) \\ \forall x \in \left[\inf\left(\mathrm{core}(\tilde{l}_{i+1})\right), \overline{x}\right] & : \mu_i(x) \leq \mu_{i+1}(x) \\ \exists x \in \left[\inf\left(\mathrm{core}(\tilde{l}_{i+1})\right), \overline{x}\right] & : \mu_i(x) < \mu_{i+1}(x) \end{cases}, \ i = \overline{1, K-1} \qquad (11)$$

где core – ядро нечеткого множества.

Нарушение интерпретабельности могут привести к неожиданному поведению нечеткой модели, например, к локальным экстремумам на поверхностях, которые должны быть гладкими. Обнаружить такие недостатки практически невозможно без просчета значений в каждой точке многофакторного пространства с последующей визуализацией поверхностей “входы – выход”. Это приводит к значительным сложностям уже при трех входных переменных. Сегодня не существует приемлемого формального критерия интерпретабельности, поэтому мы будем использовать набор условий, выполнение которых гарантирует интерпретабельность модели.

Для сохранения линейной упорядоченности терм-множества $\{l_1, l_2, ..., l_K\}$ переменной x на интервале $\left[\underline{x}; \overline{x}\right]$ наложим такие ограничения на параметры функций принадлежности:

$$\begin{cases} b_1 = \underline{x} \\ b_K = \overline{x} \\ b_i < b_{i+1}, \ i = \overline{1, K-1} \\ \forall s_i \in (\underline{x}; \overline{x}) \to s_i \in (b_i; b_{i+1}), \ i = \overline{1, K-1}, \end{cases} \qquad (12)$$

где (b_i, c_i) – параметры гауссовой функции принадлежности (2) нечеткого множества $\tilde{l}_i$, $i = \overline{1, K-1}$;

s_i – абсцисса пересечения графиков функций принадлежности соседних нечетких множеств $\tilde{l}_i$ и $\tilde{l}_{i+1}$, $i = \overline{1, K-1}$.

Первые два условия в (12) определяют, что ядра крайних термов b_1 и b_K зафиксированы на границах интервала значений переменной x. Эти условия не допускают нарушения интерпретабельности типа "c". Третье условие задает линейную упорядоченность ядер нечетких множеств $\{\tilde{l}_1, \tilde{l}_2, ..., \tilde{l}_K\}$. Совместное применение первых трех условий обеспечивает попадание ядра любого нечеткого множества в диапазон $[\underline{x}; \overline{x}]$. Четвертое условие означает, что если функции принадлежности соседних нечетких множеств пересекаются внутри диапазона возможных значений переменной x, то пересечение должна быть между ядрами этих нечетких множеств. Выполнение этого условия защищает от нарушений интерпретабельности типа "b".

Точки пересечения соседних нечетких множеств рассчитываются по следующим простым формулам:

$$s_i = \frac{b_i c_{i+1} + b_{i+1} c_i}{c_{i+1} + c_i}, \qquad (13)$$

$$s_i = \frac{b_i c_{i+1} - b_{i+1} c_i}{c_{i+1} - c_i}. \qquad (14)$$

ISSN 1512-1232

Обозначим через q_i ту точку пересечения s_i, которая попадает в диапазон $(b_i; b_{i+1})$. Тогда, $b_1 < q_1 < b_2 < q_2 < ... < q_{k-1} < b_K$. Вторую точку s_i обозначим через v_i. Тогда, четвертое условие в (12) можно переписать так: $v_i \notin (\underline{x}; \overline{x})$.

Чтобы избежать нарушения интерпретабельности типа "d", ограничим снизу высоту h_i пересечения соседних нечетких множеств $\tilde{l}_i$ и $\tilde{l}_{i+1}$, $i = \overline{1, K-1}$. Эта высота рассчитывается подстановкой в формулу (2) значения аргумента q_i. После преобразований получаем:

$$h_i = \max\left(\exp\left(-\frac{1}{2}\left(\frac{b_{i+1} - b_i}{c_i \pm c_{i+1}} \right)^2 \right) \right). \tag{15}$$

В результате система (12) преобразуется к виду:

$$\begin{cases} b_1 = \underline{x} \\ b_K = \overline{x} \\ b_1 < q_1 < b_2 < q_2 < ... < q_{k-1} < b_K \\ v_i \notin (\underline{x}; \overline{x}), \ i = \overline{1, K-1} \\ h_i \geq h^*, \ i = \overline{1, K-1}, \end{cases} . \tag{16}$$

где h^* – минимально допустимая высота пересечения соседних нечетких множеств.

4. СТРУКТУРНАЯ ИДЕНТИФИКАЦИЯ

Одной из наиболее важных задач структурной идентификации является выбор правил нечеткой базы знаний из некоторого заранее сформированного множества кандидатов. Правила-кандидаты могут быть сформированы экспертом или получены путем обработки соответствующих экспериментальных данных. Предполагается, что соответствующие правилам терм-множества удовлетворяют требованиям прозрачности из раздела 3.

В идеальном случае нечеткая база знаний должна быть интерпретируемой, компактной и адекватной. Достичь этого в реальных задачах невозможно, потому на практике пытаются выбрать интерпретабельную базу знаний с корректным балансом между противоречивыми критериями компактности и точности. Необходимым условием такого баланса является попадание интерпретабельной базы знаний на парето-фронт в координатах “компактность модели – точность модели”.

Введем следующие обозначения:

R – множество правил-кандидатов;

$y = F(R', X)$ – модель на основе нечетких правил $R' \subseteq R$, связывающая входы X с выходом y исследуемой зависимости.

Типовой поход к выбору правил заключается в использовании порогового ограничения на компактность базы знаний или на точность нечеткой модели [15,16]:

$$\begin{cases} RMSE(R') \to \min \\ N(R') \leq N^* \end{cases}, \tag{17}$$

$$\begin{cases} N(R') \to \min \\ RMSE(R') \leq RMSE^* \end{cases}, \tag{18}$$

где N^* и $RMSE^*$ – предельно-допустимые значения компактности и точности.

Типовой подход формирует довольно большую область допустимых решений, значительная часть которой размещена вдали от парето-фронта (рис. 5а и 5б). Это замедляет нахождение оптимальных решений из парето-фронта. Для сокращения зоны поиска предлагается использовать новый метод выбора правил нечеткой базы знаний в окрестности парето-фронта [17]. Эту окрестность зададим линейным ограничением, которое описывает компенсацию точности модели ее компактностью:

$$RMSE(R') \le k_0 + k_1 \cdot N(R'), \tag{19}$$

где $k_0 > 0$ и $k_1 < 0$ – параметры, подбирая которые можно сформировать область допустимых решений в окрестности парето-фронта.

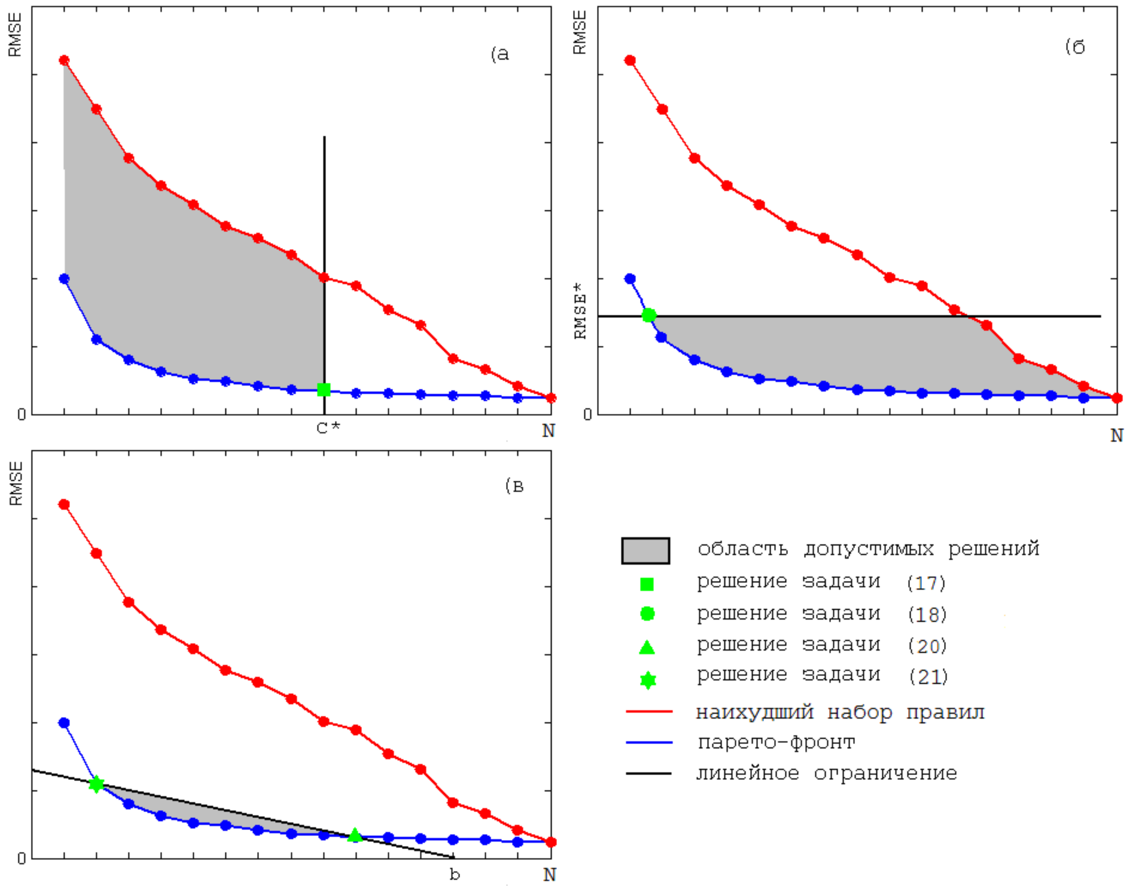

Рисунок 5 – Область допустимых решений:
а) задача (17); б) задача (18); в) задачи (20) и (21)

С учетом (19) задачи (17) и (18) преобразуются к виду:

$$\begin{cases} RMSE(R') \to \min \\ RMSE(R') \le k_0 + k_1 \cdot N(R') \end{cases}; \tag{20}$$

$$\begin{cases} N(R') \to \min \\ RMSE(R') \le k_0 + k_1 \cdot N(R') \end{cases}. \tag{21}$$

Сокращение области поиска оптимальных решений для задач (20) и (21) иллюстрирует рис. 5в. Коэффициенты ограничения оценим по крайним точкам парето-фронта, которые соответствуют почти пустым и почти заполненным базам знаний, а также по его верхней границе, которую найдем жадным алгоритмом на основе идей приближенного метода Сахни

ISSN 1512-1232

для задачи о рюкзаке [18]. Вычислительная сложность этой процедуры является квадратичной, поэтому она существенно не увеличит время оптимизации.

Поиск оптимальных решений выполним генетическим алгоритмом с кодированием хромосом по питтсбургской схеме [19]. Каждая хромосома популяции задает нечеткую базу знаний с собственным набором правил R'. Каждый ген этой хромосомы соответствует одному правилу-кандидату. Если правило попадает в базу знаний, тогда соответствующий ген принимать значение 1, а если не попадает, тогда – 0. Например, хромосома $(1, 0, 1, 0, 0, 1, 0)$ кодирует нечеткую базу знаний с тремя правилами под номерами 1, 3 и 6. Начальная популяция генерируется случайно, но с включением субоптимальных решений, найденных жадным алгоритмом (рис. 6).

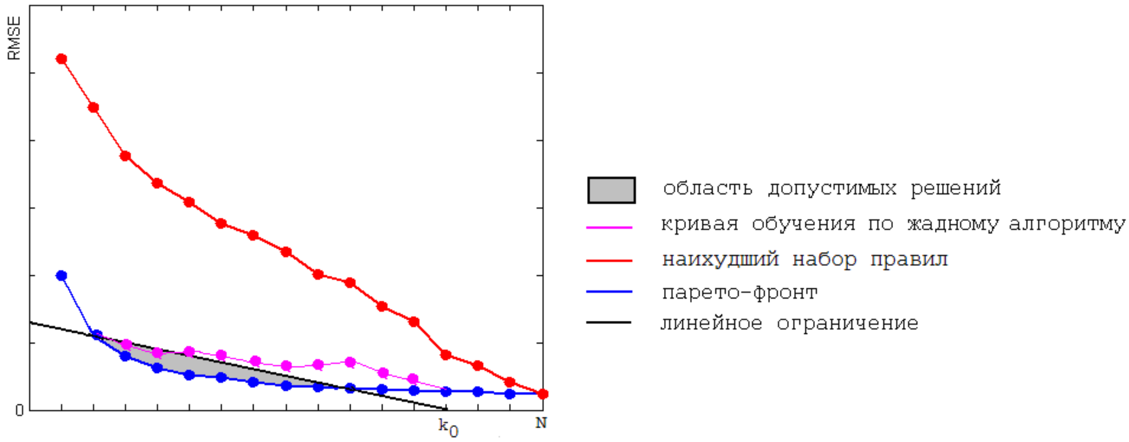

Рисунок 6 – Кривая обучения по жадному алгоритму

После получения оптимального набора правил проведем редукцию антецедентов. Смысл этой процедуры заключается в замене одного или нескольких термов в антецеденте на терм *“Don’t care”*. Такая замена эквивалента вычеркиванию соответствующих термов из антецедентов правил, т.е. переходу от длинных правил к коротким. Например, к переходу от правила $\textit{Если } \left(x_1 = \textit{Низкий и } x_2 = \textit{Средний и } x_3 = \textit{Высокий}\right), \textit{ то } y = \textit{Низкий}$ к правилу $\textit{Если } x_2 = \textit{Средний}, \textit{ то } y = \textit{Низкий}$. В результате улучшается компактность базы знаний по критерию A – суммарная длина всех антецедентов. Иногда это приводит к сокращению суммарной мощности терм-множеств, а также к объединению нескольких правил в одно.

Задачу редукции антецедентов также решим с помощью генетического алгоритма. Хромосому зададим строчкой из $N \times n$ генов. Каждый ген соответствует одному терму в антецеденте правила наилучшей нечеткой базы знаний (рис. 7). Ген принимает значение 1, если в правиле используется терм из наилучшей нечеткой базы знаний, и значение 0, если используется терм *“Don’t care”*. Чем больше количество термов *“Don’t care”* в правиле, тем оно короче и тем компактнее нечеткая база знаний. Оптимизация проводиться по критерию *RMSE* с ограничением на недопустимость появления противоречивых правил. После оптимизации база знаний подчищается – удаляются правила, антецеденты которых содержат только термы *“Don’t care”*, а также корректируются терм-множества.

$$\begin{array}{ccc} x_1 & x_2 & x_3 \\ \textit{Low} & \textit{Don't care} & \textit{High} \\ \textit{Average} & \textit{Don't care} & \textit{Low} \\ \textit{High} & \textit{High} & \textit{Don't care} \end{array} \Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \Rightarrow (1, 0, 1,\ \ 1, 0, 1,\ \ 1, 1, 0)$$

Исходная база антецедентов правил	Закодированая база правил	Хморомосома базы правил

Рисунок 7– Формирование хромосомы для задачи редукции антецедентов

ISSN 1512-1232

5. ПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ

Задача параметрической идентификации заключается в подборе таких весов правил и параметров функций принадлежностей, которые обеспечивают минимальное значение невязки $RMSE$ на тестовой выборке. В ходе параметрической идентификации структура модели не изменяется, поэтому критерии компактности не будут задействованы во время оптимизации. Однако необходимо следить, чтобы оптимизация не привела к неинтерпретабельной базе знаний.

Вектор управляемых переменных P сформируем из весов правил (W) а также параметров $(P_1, P_2, ..., P_n)$ и P_y, определяющих функции принадлежности термов входных переменных $(x_1, x_2, ..., x_n)$ и выходной переменной y. Таким образом, $P = (W, P_1, P_2, ..., P_n, P_y)$.

Для терм-множества $\{l_1, l_2, ..., l_K\}$ переменной x на интервале $\left[\underline{x}; \overline{x}\right]$ настраиваемыми параметрами будут:

$(b_2, b_3, ..., b_{K-1})$ - ядра некрайних термов $l_2, l_3, ..., l_{K-1}$;

$(q_1, q_2, ..., q_{K-1})$ - точки пересечения соседних нечетких множеств на интервале $\left[\underline{x}; \overline{x}\right]$;

c_1 - коэффициент концентрации функции принадлежности терма l_1.

Таким образом, для переменной x вектор настраиваемых параметров задается так: $P_x = (q_1, b_2,\ q_2, b_3, ..., b_{K-1}, q_{K-1}, c_1)$. Это позволяет записать ограничения для сохранения прозрачности в удобной форме (16), в отличие от более громоздких и субъективных ограничений в [2, 12]. Общее количество настраиваемых параметров составляет:

$$|P| = N + (2T_y - 2) + \sum_{j=\overline{1,\, n}} (2k_j - 2), \tag{22}$$

где T_y – мощность терм-множеств выходной переменной y.

На основе параметров P_x коэффициенты концентраций функций нечетких множеств $\tilde{l}_i$ рассчитываються так:

$$c_i = \pm c_{i-1} \frac{q_{i-1} - b_i}{q_{i-1} - b_{i-1}},\ i = \overline{2,\, K}. \tag{23}$$

6. ИНФОРМАЦИОНАЯ ТЕХНОЛОГИЯ НЕЧЕТКОЙ ИДЕНТИФИКАЦИИ

Выше описанные модели и методы реализованы в форме информационной технологии нечеткой идентификации. Нечеткая идентификация осуществляется поэтапно, согласно концепции “генерация – селекция - редукция – настройка” (рис. 8).

На первом этапе "Fuzzy Rules Generation" происходит генерация нечетких правил из экспериментальных данных методом прямого прохода. Данный метод базируется на идеях генерации базы знаний методом Ванга–Менделя [20], с тем лишь отличаем, что консеквентом правила выбирается терм не с максимальной принадлежностью для какой-то одной строчки выборки данных, а терм с максимальной средней принадлежности по всем данных из соответствующей зоны факторного пространства. Если аналогом метода Ванга–Менделя является алгоритм нечеткой классификации с единственным правилом-победителем, то аналогом нашего метода является алгоритм нечеткой классификации с голосованием правил [21]. Обычно нечеткие классификаторы с алгоритмом голосования правил обеспечивают лучшую безошибочность, поэтому аналогична схема выбрана нами для генерации правил.

Второй этап "Fuzzy Rules Selection" реализован с помощью бинарного генетического алгоритма с условиями (20) или (21). Опционально, параметры ограничений могут быть оценены из кривых обучения жадного алгоритма.

Третий этап "Antecedents Reduction" реализован тем же генетическим алгоритмом с применением альтернативного кодирования.

Последний этап "Tuning" настраивает внутреннюю структуру нечеткой базы знаний с использованием градиентных и квазиньтоновских методов оптимизации. Для сохранения интерпретабельности оптимизация осуществляется с ограничениями (12).

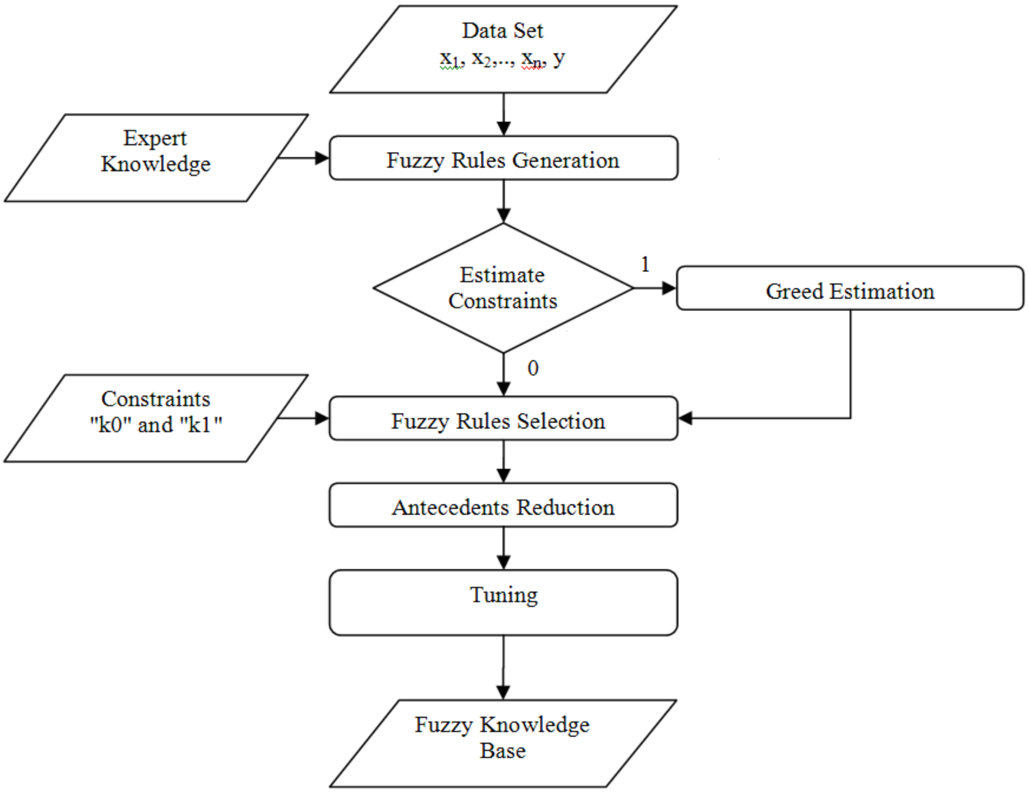

Рисунок 8 – Схема функционирования информационной технологии

7. КОМПЬЮТЕРНЫЕ ЭКСПЕРИМЕНТЫ

Эксперименты по нечеткой идентификации проведены для 7 реальных задач из UCI Machine Learning Repository [10], краткое описание которых приведено в табл. 1.

Таблица 1 – Тестовые задачи из UCI Machine Learning Repository

№	Задача	Количество записей	Количество входов
1	Auto-MPG	398	9
2	Boston Housing	506	13
3	Combined Cycle Power Plant	9568	4
4	Condition Based Maintenance of Naval Propulsion Plants	11934	16
5	Airfoil Self-Noise	1503	6
6	SkillCraft1 Master Table	3395	20
7	Physicochemical Properties of Protein Tertiary Structure	45730	9

Каждая выборка данных разбивалась на обучающую и тестовую одинакового объема. В обучающую выборку вошли нечетные строки, а в тестовую – четные. Для генерации правил использовалось лингвистическое разбиение с тремя термами на каждый вход и с пятью термами на выход. В качестве ориентира выбран метод генерации нечетких баз знаний на основе кластеризации по методу нечетких c-средних (FCM) в реализации MATLAB. В отличии от используемого метода генерации правил, FCM синтезирует компактную, но неинтепретабельную базу знаний.

Генетическая селекция правил осуществлялась по постановке задачи (20) с коэффициентами в линейном ограничении $k_0 = 0.3$ и $k_1 = -0.0075$. Эти значения заданы из априорных соображений, чтобы обнаружить нечеткую базу знаний объемом не более 40 правил с нормированной ошибкой ниже 0.3. Редукции антецедентов осуществлялась по постановке задачи (17) с минимизацией ошибки, но с ограничением на суммарную длину антецедентов всех правил $\frac{A}{A_{\max}} \le 0.9$. Обучение осуществлялось настройкой внутренних параметров нечеткой базы знаний квазиньютоновским методом оптимизации в двух режимах – с ограничениями на сохранение прозрачности и без них. На каждом этапе оптимизация длилась 150 итераций.

Вначале были сгенерированы большие базы знаний, количеством правил для некоторых задач превышало несколько сотен (рис. 9). После селекции количество правил существенно сократилось до уровня, сопоставимого с FCM-базами знаний. На этапе редукции антецедентов в некоторых случаях удалось уменьшить длину антецедентов вдвое (см. рис. 9, на котором отображены относительные значения $Rules' = \frac{A \cdot N}{A_{\max}}$). Для задачи SkillCraft метод FCM не применялся из-за пропусков некоторых данных в выборке.

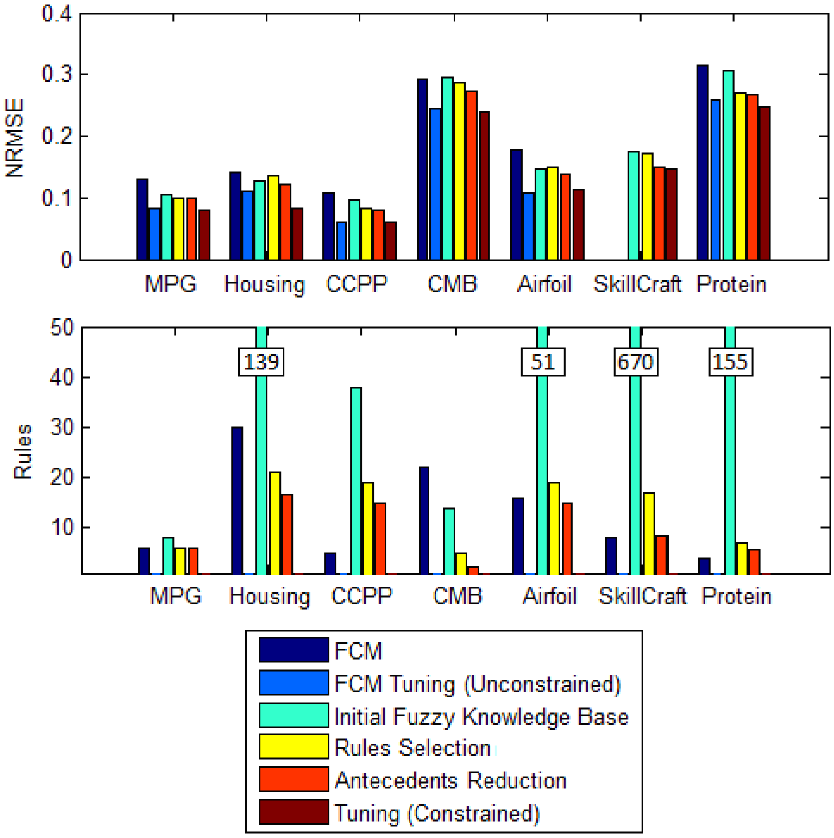

Рисунок 9 – Результаты нечеткой идентификации по методу FCM и по предлагаемому методу

В целом после настройки нечеткие базы знаний, синтезированные по нашему методу, оказались немного точнее, чем в случае FCM-метода. Причем, по нашему методу базы знаний получаются не только точными, но и интерпретабельными. Кроме того, в базах знаний, синтезированных по нашему методу, присутствуют короткие правила с термами

ISSN 1512-1232

“Don’t care”, Например, для задачи CMB длина антецедентов сокращена в 4 раза. Селекция правил и редукция антецедентов может снизить суммарное количества термов в базе знаний (рис. 10), а значит и уменьшить сложность задачи настройки. Существенное сокращение количества термов наблюдается в задачах с большим количеством факторов: SkillCraft, Housing и CMB.

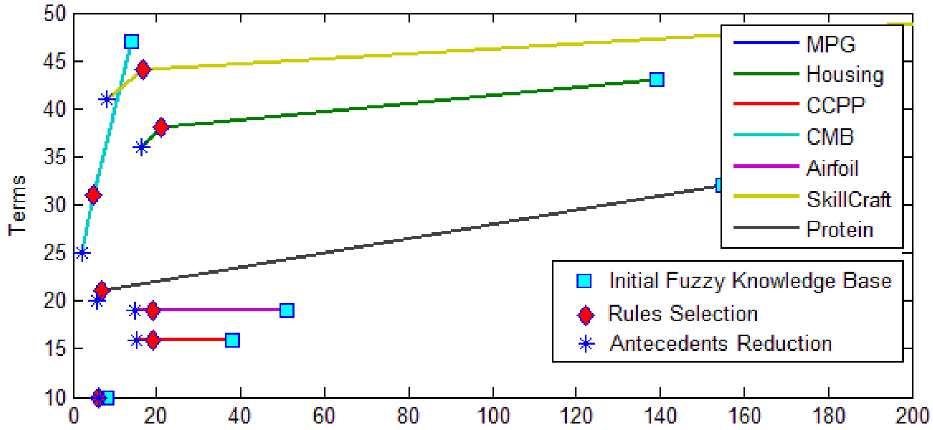

Рисунок 10 – Траектории изменения компактности нечетких баз знаний

Эксперименты показали, что ограничения для сохранения интерпретабельности не снижают точность настройки (рис. 11). Настройка осуществлялась после редукции антецедентов. Разница между точностью настройки в случае применения ограничений и без них составляет всего лишь несколько сотых для шести задач. Это подтверждает, что для правильно синтезированной нечеткой базы знаний интерпретабельность не вредит точности.

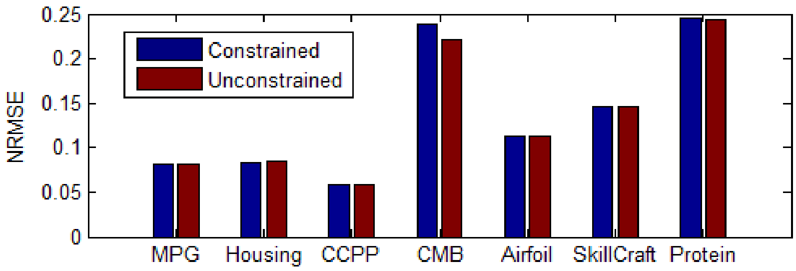

Рисунок 11 – Точность настройки нечетких баз знаний с ограничениями, сохраняющих интерпретабельность, и без них

ВЫВОДЫ

Предложена информационная технология нечеткой идентификации для синтеза точных, компактных и интерпретабельных баз знаний. Информационная технология включает 4 этапа: генерация достоверных нечетких правил-кандидатов, селекция правил, редукция антецедентов и настройка функций принадлежности и весов правил. Компьютерные эксперименты на 7 задачах из UCI Machine Learning Repository показали, что предлагаемая информационная технология синтезирует нечеткие базы знаний с показателями точности и компактности на уровне конкурентных технологий. При этом, в отличие от конкурентов, предлагаемая технология обеспечивает интерпретабельность нечетких баз знаний.

Статья содержит результаты исследований, выполненных при поддержке Государственного фонда фундаментальных исследований Украины по программе Ф62 «Гранты Президента Украины докторам наук для проведения научных исследований”.

ISSN 1512-1232

ЛИТЕРАТУРА

1. Цыпкин Я.З. Основы информационной теории идентификации. Москва: Наука, 1984.
2. Штовба С.Д. Проектирование нечетких систем средствами MATLAB. Москва: Горячая линия – Телеком, 2007.
3. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. Москва: Мир, 1976.
4. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткая логика, генетические алгоритмы, нейронные сети. Винница: УНІВЕРСУМ–Вінниця, 1999.
5. Ishibuchi H., Nakashima T., Murata T. Three-objective genetics-based machine learning for linguistic rule extraction. Information Science, 2001, Vol. 136, No. 1, P. 109–133.
6. Cordon O. A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems. International Journal of Approximate Reasoning, 2011, Vol. 52, P. 894–913.
7. Gacto, María José, Rafael Alcalá, and Francisco Herrera. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. Information Sciences, 2011, Vol. 181, No. 20, P 4340–4360.
8. Guillaume, Serge, and Brigitte Charnomordic. Learning interpretable fuzzy inference systems with FisPro. Information Sciences, 2011, Vol. 181, No. 20, P. 4409–4427.
9. Штовба С.Д., Штовба Е.В., Панкевич О.Д. Критерии точности и компактности для оценки качества нечетких баз знаний в задачах идентификации. Научные труды Винницкого национального технического университета, 2012, №4.
10. Lichman M. UCI Machine Learning Repository. Irvine (USA): School of Information and Computer Science, University of California, 2013. Режим доступа: http://www.ics.uci.edu/~mlearn/MLRepository.html .
11. Лю Б. Теория и практика неопределенного программирования. Москва: БИНОМ. Лаборатория знаний, 2005.
12. Штовба С.Д. Обеспечение точности и прозрачности нечеткой модели Мамдани при обучении по экспериментальным данным. Проблемы управления и информатики, 2007, №4, С. 102-114.
13. Ротштейн А.П. Медицинская диагностика на нечеткой логике. Винница: Континент–ПРИМ, 1996.
14. Miller G.A. The magic number seven plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, Vol. 63, P. 81–97.
15. Ishibuchi H., Nozaki K., Yamamoto N., Tanaka H., Selecting fuzzy if-then rules for classification problems using genetic algorithms. IEEE Transactions on Fuzzy Systems, 1995, Vol. 3, No. 3, P. 260–270.
16. Ishibuchi H., Murata T., Turksen I. B. Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. Fuzzy Sets and Systems, 1997, Vol. 89, No. 2, P. 135–150.
17. Штовба С.Д., Мазуренко В.В., Савчук Д.А. Генетический алгоритм выбора правил нечеткой базы знаний, сбалансированной по критериям точности и компактности. Научные труды Винницкого национального технического университета, 2012, №3.
18. Martello S., Toth P. Knapsack problems: algorithms and computer implementations. New York: John Wiley & Sons, Inc, 1990.
19. Cordon O., Gomide E., Herrera E., Homannc E. Magdalena L. Ten years of genetic fuzzy systems: current framework and new trends. Fuzzy Sets and Systems, 2004, Vol. 141, P. 5–31.
20. Wang, L. X., & Mendel, J. M. Generating fuzzy rules by learning from examples. Systems, Man and Cybernetics, IEEE Transactions on, 1992, Vol. 22, No. 6, P. 1414-1427.
21. Ishibuchi H., Nakashima T., Morisawa T. Voting in fuzzy rule-based systems for pattern classification problems. Fuzzy Sets and Systems, 1999, Vol. 103, №2, P. 223–238.

Количество рисунков – 11 шт.
Количество таблиц – 1 шт.

Статья получена: 2015-09-29