УДК 658.012

С. Д. Штовба, д. т. н., проф.; В. В. Мазуренко; Д. А. Савчук

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ВЫБОРА ПРАВИЛ НЕЧЕТКОЙ БАЗЫ ЗНАНИЙ, СБАЛАНСИРОВАННОЙ ПО КРИТЕРИЯМ ТОЧНОСТИ И КОМПАКТНОСТИ

Предлагается генетический алгоритм поиска набора правил для формирования нечеткой базы знаний, сбалансированной по критериям точности и компактности. Отличием алгоритма является введение в постановку задачи оптимизации линейного ограничения, которое задает уровень компенсации точности модели ее компактностью. Это приближает область допустимых решений к парето-фронту.

Ключевые слова: *нечеткая база знаний, точность, компактность, выбор правил, парето-фронт, генетическая оптимизация.*

Введение

Нечеткой базой знаний называется совокупность нечетких правил “Если – то”, описывающих взаимосвязь между входами и выходами некоторого объекта с использованием лингвистических термов [1]. Одной из задач проектирования нечеткой базы знаний является выбор правил из некоторого, наперед определенного, множества кандидатов. Правила-кандидаты могут быть сформированы экспертом или получены путем обработки соответствующих экспериментальных данных.

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

Выбор правил нечеткой базы знаний можно свести к бинарной задачи о рюкзаке. Правилу базы знаний соответствует предмет, который может быть помещен в рюкзак, точности базы знаний – полезность рюкзака, а количество правил – суммарный объем выбранных предметов. Отличие между задачами заключается в разных типах функции полезности, которая является линейной в задаче о рюкзаке и нелинейной в задаче выбора правил базы знаний. По аналогии с классическими постановками задач о рюкзаке [2] сформирована и задача выбора правил нечеткой базы знаний. Основными работами в этой области являются статьи [3, 4] по формированию множества баз знаний нечеткого классификатора, которые принадлежат парето-фронту недоминирующих альтернатив в координатах “количество правил – безошибочность”. Для этого применяют задачи оптимизации: с целью 1) максимизации безошибочности при ограниченном количестве правил; 2) минимизации количества правил при заданном уровне безошибочности; 3) минимизации интегрального критерия качества базы знаний в форме линейной свертки безошибочности и количества правил [4] или безошибочности, количества правил и суммарной длины антецедентов правил [5]. Для получения парето-фронта оптимизацию проводят многократно при разных пороговых значениях в ограничениях задач 1 и 2 и весовых коэффициентов целевой функции в задаче 3. Аналогичные подходы применяют, выбирая правила нечетких баз знаний для объектов с непрерывным выходом [6].

Задача выбора правил нечеткой базы знаний, как и задача о рюкзаке, является NP-полной. Соответственно, алгоритм точного решения этой задачи имеет экспоненциальную вычислительную сложность, и поэтому будет приемлемым только при небольшом количестве правил-кандидатов. На практике для решения этой задачи обычно применяют генетические алгоритмы. Кодирование вариантов осуществляют по Питтсбургскому методу [7], представляя вариант решения хромосомой, каждый ген которой задает принадлежность соответственного правила базе знаний [6].

Пороговое ограничение на сложность базы знаний или на точность нечеткой модели [3, 4] формирует довольно большую область допустимых решений, значительная часть которой размещена вдали от парето-фронта. Это замедляет нахождение оптимальных решений, которые находятся на парето-фронте. **Целью статьи** является сокращение вычислительной сложности выбора правил нечеткой базы знаний за счет разработки нового метода поиска оптимальных решений в окрестности парето-фронта. Эту окрестность зададим линейным ограничением, которое описывает компенсацию точности модели ее компактностью. 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 для задачи о рюкзаке [2]. Вычислительная сложность этой процедуры является квадратичной, поэтому она существенно не увеличит время оптимизации. Поиск оптимальных решений выполним генетическим алгоритмом.

1. Математические постановки задач

Будем считать известными:

- выборку из M пар экспериментальных данных, о влиянии факторов $X = (x_1, x_2, ..., x_n)$ на непрерывный выход y исследуемой зависимости:

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

где X_r – входной вектор в r-ой строчке выборки; y_r – соответствующее выходное значение;

- множество R из N правил-кандидатов в нечеткую базу знаний, $N = |R|$.

Обозначим через $y = F(R', X)$ модель на основе нечетких правил $R' \subseteq R$, связывающую входы X с выходом y исследуемой зависимости. Критерием точности нечеткой модели выберем среднюю квадратическую ошибку на выборке (1):

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

В общем случае задача состоит в определении такого множества правил R', которое обеспечивает:

$$\begin{cases} RMSE(R') \to \min \\ C(R') \to \min \end{cases}, \tag{3}$$

где $C(R')$ – сложность нечеткой модели, которую определим количеством правил $C(R') = |R'|$ или в общем случае уровнем наполненности базы знаний $C(R') = \frac{|R'|}{N}$.

Многокритериальную задачу оптимизации (3) преобразуют в такие скалярные задачи [2, 3]:

$$\begin{cases} RMSE(R') \to \min \\ C(R') \le C^* \end{cases}, \tag{4}$$

$$\begin{cases} C(R') \to \min \\ RMSE(R') \le RMSE^* \end{cases}, \tag{5}$$

где C^* и $RMSE^*$ – максимально допустимые значения сложности и ошибки.

Постановки (4) и (5) формируют большую область допустимых решений, причем значительная ее часть находиться далеко от парето-фронта (рис. 1а и 1б). Мы предлагаем ограничение в задаче оптимизации записать таким образом:

$$RMSE(R') \le a \cdot C(R') + b\,, \tag{6}$$

где $a < 0$ и $b > 0$ – параметры, подбирая которые можно сформировать область допустимых решений в окрестности парето-фронта (рис. 1в).

Используя ограничение (6), сформулируем следующие задачи выбора правил нечеткой базы знаний:

$$\begin{cases} RMSE(R') \to \min \\ RMSE(R') \le a \cdot C(R') + b \end{cases}, \tag{7}$$

$$\begin{cases} C(R') \to \min \\ RMSE(R') \le a \cdot C(R') + b \end{cases}. \tag{8}$$

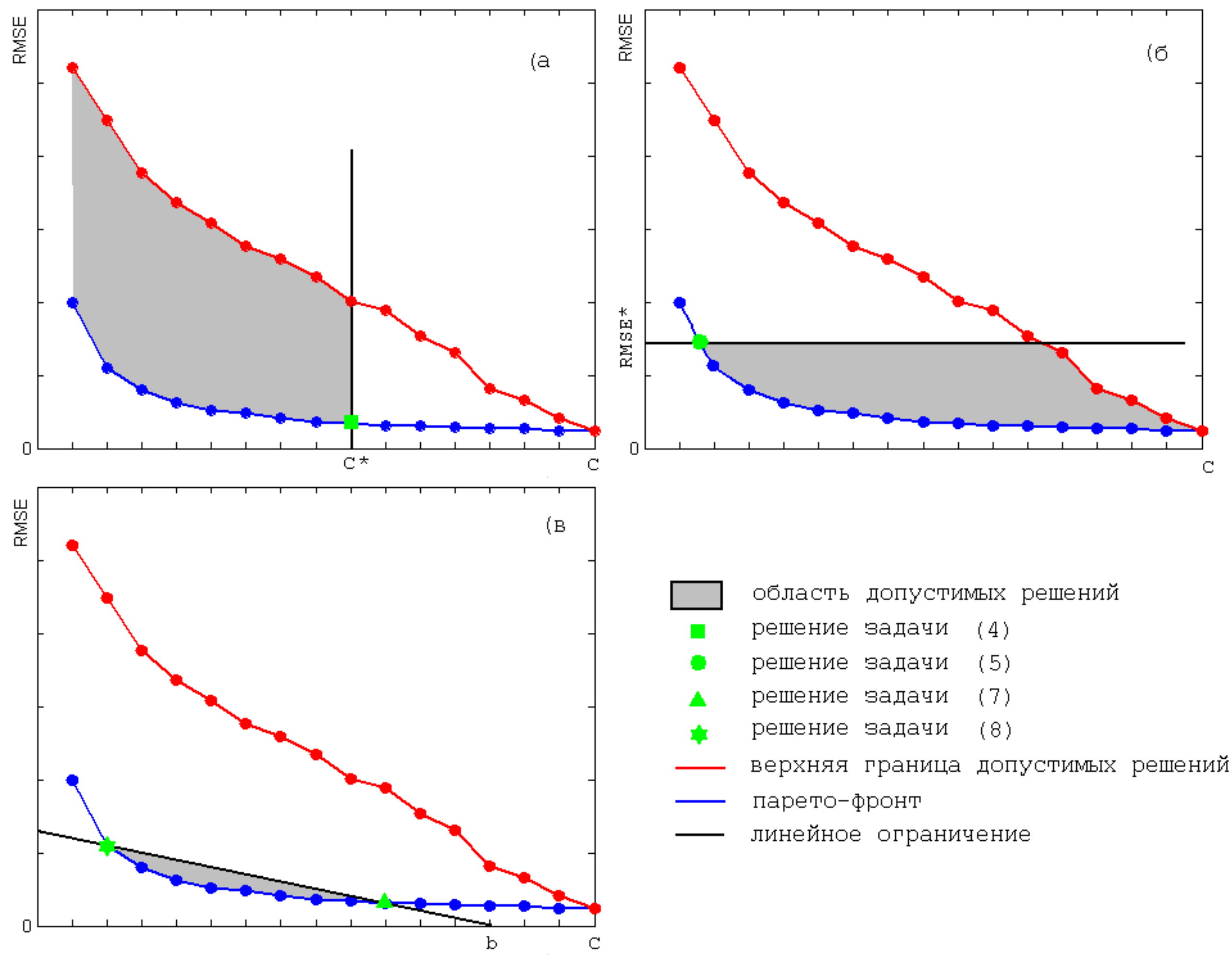

Рис. 1. Область допустимых решений:
а) задача (4); б) задача (5); в) задачи (7) и (8)

2. Оценка параметров линейного ограничения

Для определения параметров a и b ограничения (6) достаточно знать соответствующие характеристики двух баз знаний, которые удовлетворяют пользователя. Обозначим их $(C_1, RMSE_1)$ и $(C_2, RMSE_2)$. Тогда:

$$\begin{cases} a = \dfrac{RMSE_2 - RMSE_1}{C_2 - C_1} \\ b = RMSE_1 - a \cdot C_1 \end{cases}. \qquad (9)$$

Параметр a можно трактовать как коэффициент компенсации точности за счет компактности. Его можно рассчитать из ответа пользователя на вопрос «Насколько можно уменьшить точность модели за счет сокращения числа правил на 1?». Тогда для определения второго параметра b достаточно знать характеристики одной приемлемой базы знаний.

Параметры a и b можно определить, проведя линейное ограничение через какие-либо две крайние точки парето-фронта. Одна из крайних точек должна быть слева, а другая – справа (рис. 2). Вычислительная сложность полного перебора для определения 5-ти крайних точек парето-фронта для баз знаний из 1-го, 2-х, *N*–2, *N*–1 и *N* правил является квадратичной $O(N^2)$, поэтому такой подход можно применять и для задач большой размерности.

Линейное ограничение можно провести и через две точки кривой обучения в форме зависимости невязки от сложности нечеткой базы знаний. Кривую обучения предлагается построить по результатам выполнения жадного алгоритма выбора правил. Этот алгоритм заключается в добавлении на каждом шаге к базе знаний одного правила, которое максимально уменьшает невязку. Полученная кривая обучения всегда будет не ниже парето-фронта (см. рис. 2). Исходной базой знаний для жадного алгоритма можно выбрать базу знаний из парето-фронта, содержащую 2 или *N*–2 правила. Вычислительная сложность жадного алгоритма является квадратичной.

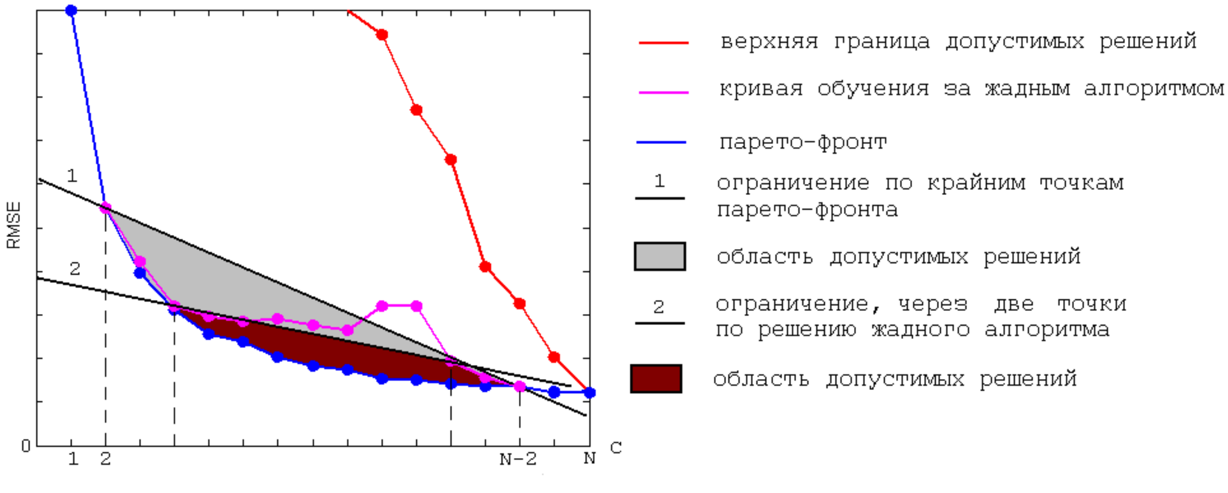

Рис. 2. К расчету параметров линейного ограничения

3. Генетический алгоритм решения задачи

Для решения задачи оптимизации (7) и (8) воспользуемся генетическим алгоритмом на основе Питтсбургского подхода. Каждая хромосома популяции задает нечеткую базу знаний с собственным набором правил R'. Каждый из N генов этой хромосомы может принимать такие значения: 1 (если соответствующее правило попадает в базу знаний) и 0 (если не попадает).

Начальная популяция генерируется случайно, но с включением субоптимальных решений, найденных жадным алгоритмом.

Вероятность выбора хромосомы для скрещивания определяется следующим образом:

$$p = \frac{n-j}{2n}, \tag{10}$$

где n – размер популяции; j – ранг хромосомы, который определяется фитнес-функцией.

Мутации подвергается β-доля хромосом, полученных в результате скрещивания.

Селекция осуществляется детерминированным выбором n лучших хромосом.

4. Компьютерные эксперименты

Эксперименты проведем для синглтонных нечетких баз знаний, в которых антецеденты правил задаются нечеткими множествами, а консеквенты – числовыми значениями [1]. Как и в наших предыдущих статьях по многокритериальному формализированному проектированию нечетких баз знаний [8 – 11] эксперименты осуществим на трёх эталонных зависимостях (рис. 3):

$$\text{возрастающей} - y = a\sqrt{b} \quad a \in [2;22],\ b \in [2;14]; \tag{11}$$

$$\text{унимодальной} - \ y = -a^2 - b^2,\ a \in [-7;3],\ b \in [-5;5]; \tag{12}$$

$$\text{многоэкстремальной} - \ y = (1+\sin(a)^2)^b,\ a \in [0;5],\ b \in [0.5;2]. \tag{13}$$

Для каждой зависимости (11) – (13) создана полная синглтонная нечеткая база знаний из $N = 16$ правил (табл. 1). Фаззификация входных переменных осуществлена гауссовыми функциями принадлежности [1] (рис. 4). Консеквенты правил рассчитаны по функциям (11) – (13) с аргументами, равными ядрам нечетких множеств из антецедентов правил.

Таблица 1

Полные наборы правил (R) для каждой зависимости

№	a	b	y, для зависимости (11)	y, для зависимости (12)	y, для зависимости (13)
1	Very low	Very low	5,04	-71,91	0,95
2	Low	Very low	14,04	-48,94	0,81
3	Medium	Very low	24,84	-45,27	0,94
4	High	Very low	33,84	-62,14	0,79
5	Very low	Low	7,59	-46,08	1,04
6	Low	Low	21,14	-23,11	1,23
7	Medium	Low	37,4	-19,44	1,06
8	High	Low	50,95	-36,3	1,26
9	Very low	Medium	9,82	-31,08	1,17
10	Low	Medium	23,37	-8,1	2,02
11	Medium	Medium	48,42	-4,44	1,25
12	High	Medium	65,96	-21,3	2,17
13	Very low	High	11,36	-31,91	2,29
14	Low	High	31,64	-8,94	3,04
15	Medium	High	55,97	-5,27	1,45
16	High	High	76,25	-22,14	3,40

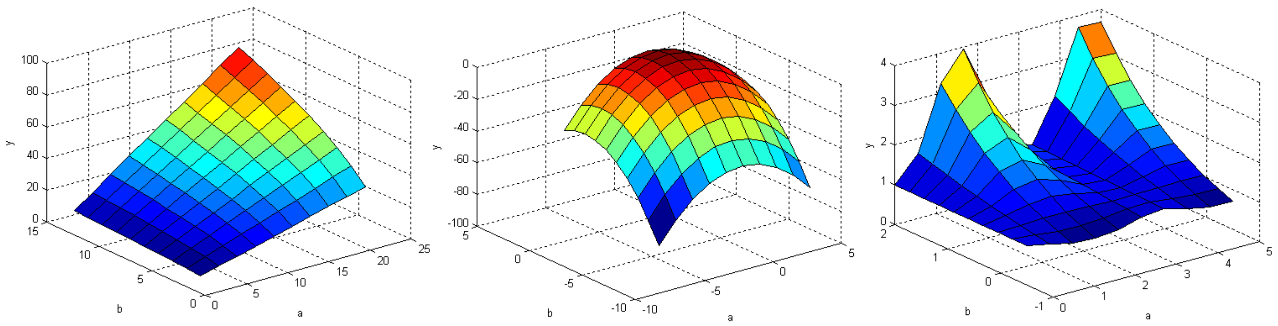

Рис. 3. Поверхности зависимостей (11) – (13)

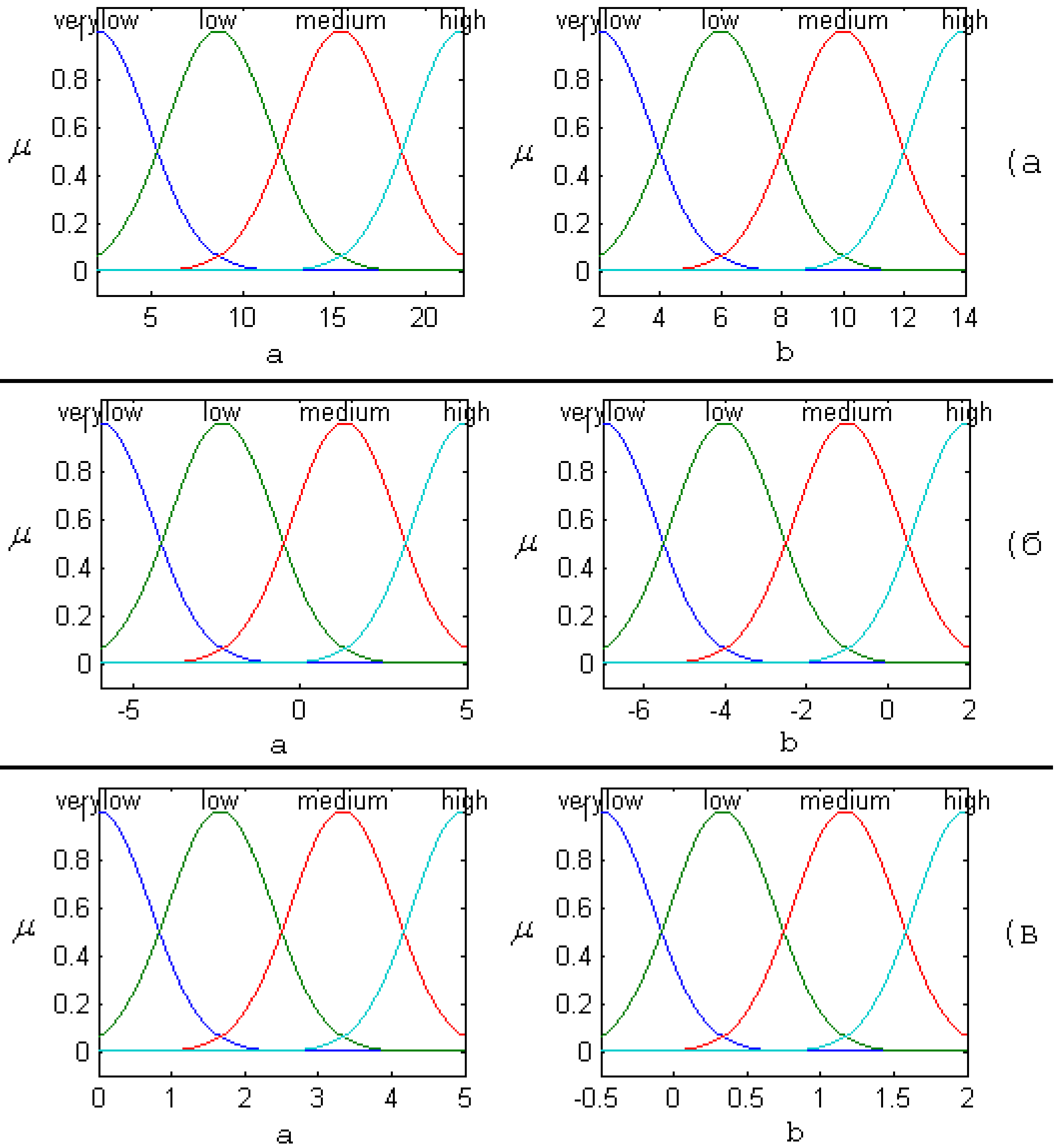

Рис. 4. Функции принадлежности термов входных переменных для:
а) зависимости (11); б) зависимости (12); в) зависимости (13)

Параметры линейного ограничения в (7) – (8) определим для каждого эксперимента отдельно. Для задачи с эталонной зависимостью (11) сначала жадным алгоритмом найдем лучшие базы знаний с разным количеством правил (рис. 5). Считая полученную кривую обучения ориентиром, установим желательное значение невязки для базы знаний из 4-х правил немного меньше, чем на рис. 5, например, на уровне $RMSE \le 0,55$. Другой точкой линейного ограничения выберем базу знаний из 10 правил с невязкой не больше, чем для полной базы знаний, то есть $RMSE \le 0,22$. Подставляя в (9), получаем $a = -0,0367$ и $b = 0,6968$. Для зависимости (12) будем считать приемлемыми базу знаний из 6 правил с невязкою $RMSE \le 0,75$ и базу знаний из 10 правил с невязкою $RMSE \le 0,58$. Подставляя эти значения в (9), получаем $a = -0,0425$ и $b = 1,005$. Для зависимости (13) приемлемой будем считать базу знаний из 9-ти правил с невязкой $RMSE \le 0,0365$ и таким уровнем компенсации невязки на одно дополнительное – $\Delta RMSE \le -0,00125$. Отсюда $a = -0,00125$ и $b = 0,0365 + 9 \cdot 0,00125 = 0,04775$.

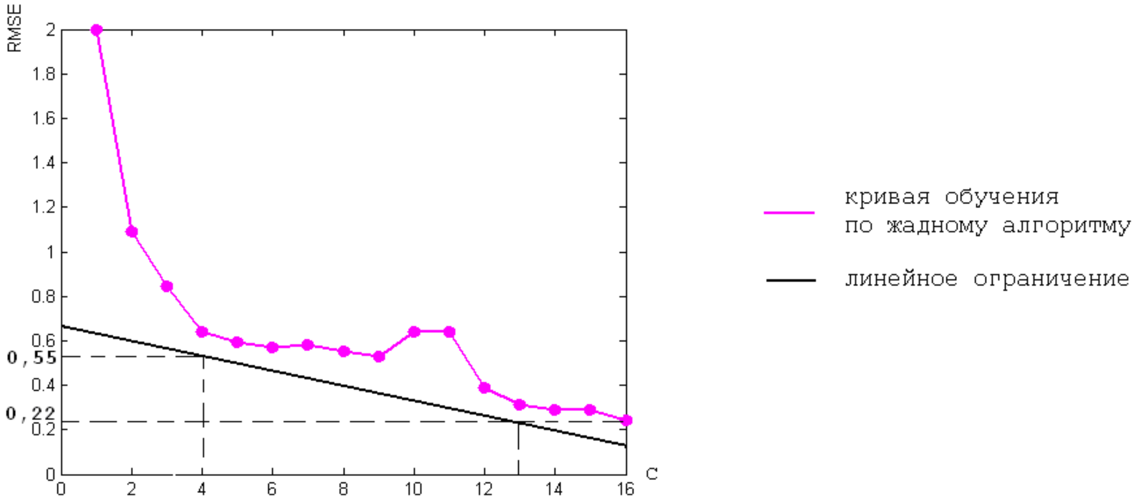

Рис. 5. Оценка параметров ограничения для экспериментов с эталонной зависимостью (11)

Эксперименты проведены при следующих параметрах генетического алгоритма: размер популяции $n = 160$, количество генов $N = 16$, давление мутации $\beta = 0,3$, количество эпох $k = 10$. Полученные решения задач (7) и (8) сведены в табл. 2. Во всех 6-ти случаях найденные нечеткие базы знаний находятся на парето-фронте, то есть имеют наименьшую невязку при фиксированном числе правил (рис. 6). Парето-фронт, а также верхняя граница области допустимых решений найдены в наших предыдущих работах [10, 11] с помощью полного перебора всех возможных комбинаций правил нечеткой базы знаний. Вычислительная сложность полного перебора является экспоненциальной $O(2^N)$, поэтому в тех работах для решения каждой тестовой задачи проверено 65536 вариантов нечеткой базы знаний. Предложенный генетический алгоритм нашел глобальное решение, перебрав для каждой задачи 1600 вариантов.

Таблица 2

Результаты экспериментов

Эталонная зависимость	(11)	(11)	(12)	(12)	(13)	(13)
Постановка задачи	(7)	(8)	(7)	(8)	(7)	(8)
Параметры ограничения	a=-0,0367 b=0,6968	a=-0,0367 b=0,6968	a=-0,0425 b=1,005	a=-0,0425 b=1,005	a=-0,00125 b=0,04775	a=-0,00125 b=0,04775
Решение (R')	(1; 2; 5; 6; 7; 9; 10; 12; 16)	(2; 3; 6; 7; 11; 16)	(1; 2; 4; 5; 6; 7; 8; 10; 11; 13; 16)	(1; 3; 6; 11; 12)	(2; 5; 7; 9; 10; 11; 12; 13; 14; 15; 16)	(1; 3; 11; 14; 16)
C(R')	10	5	11	5	11	5
RMSE(R')	0,3098	0,5128	0,5134	0,7235	0,0334	0,0387

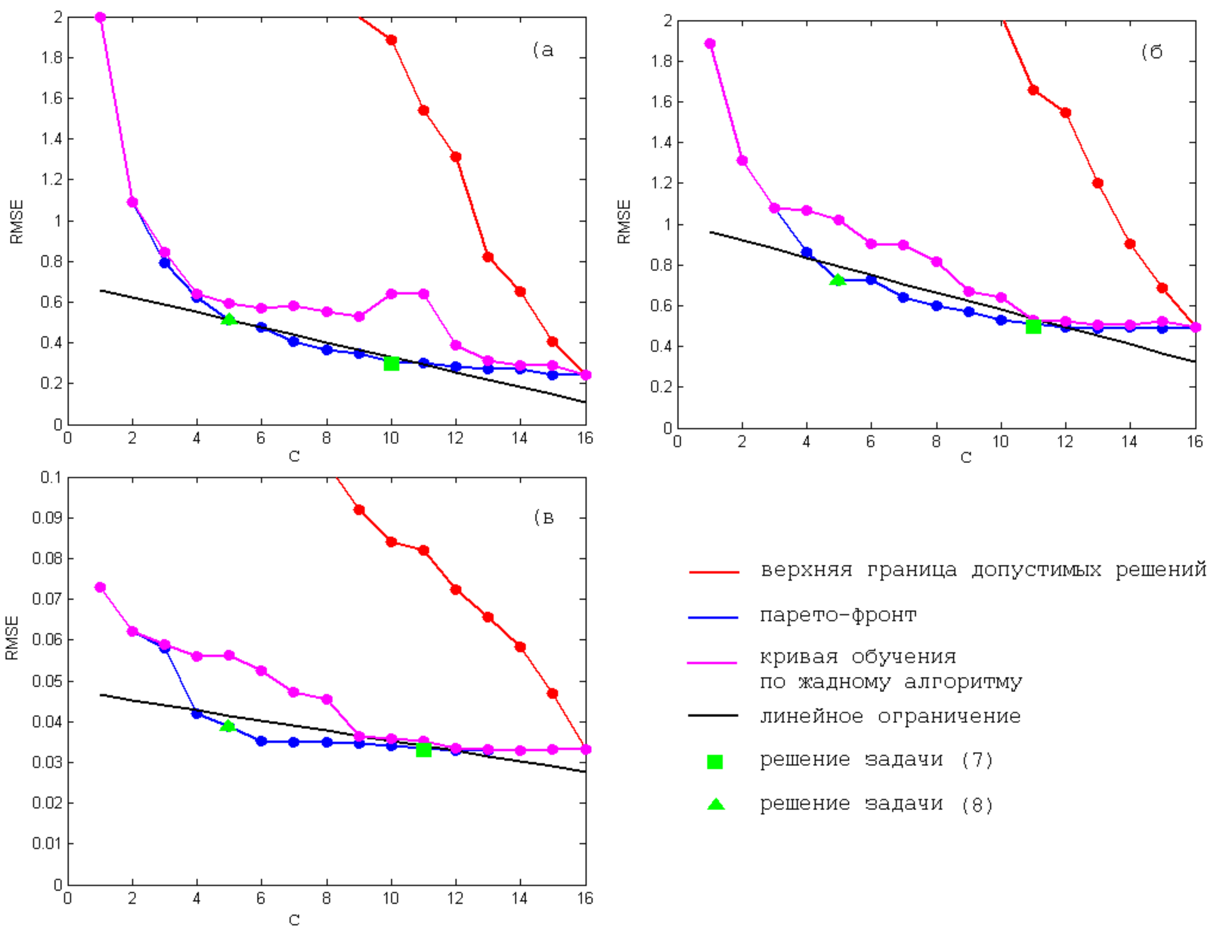

Рис. 6. Кривые обучения нечеткой базы знаний для:
а) зависимости (11); б) зависимости (12); в) зависимости (13)

Выводы

Разработан новый метод решения одной из задач нечеткой идентификации, а именно: выбора правил нечеткой базы знаний с учетом точности и компактности. Новизна метода состоит в использовании вместо типовых предельных ограничений на точность и сложность линейного ограничения, которое задает уровень компенсации между этими противоречивыми критериями. При новых ограничениях удается существенно сузить область допустимых решений, стянув ее к окрестности парето-фронта. С помощью компьютерных экспериментов установлено, что по новой постановке задачи генетический алгоритм находит глобальный оптимум, сгенерировав в десятки раз меньше вариантов нечетких баз знаний, чем в случае полного перебора.

СПИСОК ЛИТЕРАТУРИ

1. Штовба С. Д. Проектирование нечетких систем средствами MATLAB / С. Д. Штовба. – М.: Горячая линия. – Телеком, 2007. – 288 с.
2. Martello S. Knapsack problems: algorithms and computer implementations / S. Martello , P. Toth. – New York: John Wiley & Sons, Inc, 1990. – 296 p.
3. Ishibuchi H. Selecting fuzzy if-then rules for classification problems using genetic algorithms / H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka // IEEE Transactions on Fuzzy Systems. – 1995. – Vol. 3, No. 3. – P. 260 – 270.
4. Ishibuchi H. Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems / H. Ishibuchi, T. Murata, I. B. Turksen // Fuzzy Sets and Systems. – 1997. – Vol. 89, No. 2 – P. 135 – 150.
5. Ishibuchi H. Three-objective genetics-based machine learning for linguistic rule extraction / H. Ishibuchi, T. Nakashima, T. Murata // Inform. Sci. – 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 / O. Cordon // International Journal of Approximate Reasoning. – 2011. – Vol. 52. – P. 894 – 913.
7. Cordon O. Ten years of genetic fuzzy systems: current framework and new trends / O. Cordon, F. Gomideb, F. Herreraa, F. Homannc, L. Magdalenad // Fuzzy Sets and Systems. – 2004. – Vol. 141. – P. 5 – 31.
8. Штовба С. Д. Вплив кількості нечітких правил на точність бази знань Мамдані / С. Д. Штовба , В. В. Мазуренко , О. Д. Панкевич // Вісник Хмельницького національного університету. Технічні науки. – 2011. – № 2. – С. 185 – 188.
9. Штовба С. Д. Дослідження навчання компактних нечітких баз знань типу Мамдані / С. Д. Штовба, В. В. Мазуренко // Штучний інтелект. – 2011. – № 4. – С. 521 – 529.
10. Штовба С. Д. Дослідження навчання компактних нечітких синглтонних баз знань / С. Д. Штовба, В. В. Мазуренко // Вимірювальна та обчислювальна техніка в технологічних процесах. – Хмельницький: ХНУ., 2011 – № 1 – С. 133 – 139.
11. Штовба С. Д. Залежність точності ідентифікації від обсягу нечіткої синглтонної бази знань / С. Д. Штовба , О. Д. Панкевич, В. В. Мазуренко // Інформаційні технології та комп’ютерна інженерія. – 2011. – № 1. – С. 73 – 78.

Штовба Сергей Дмитриевич – профессор, д. т. н., профессор кафедры компьютерных систем управления.

Мазуренко Виктор Владимирович – аспирант кафедры компьютерных систем управления.

Савчук Дмитрий Анатольевич – студент институт автоматики, электроники и компьютерных систем управления.

Винницкий национальный техничный университет.