

УДК658.71.08, 519.87

*ГЛАДКИХ АНАТОЛИЙ АФАНАСЬЕВИЧ,
ВОЛКОВ АНДРЕЙ КОНСТАНТИНОВИЧ,
ВОЛКОВ АЛЕКСАНДР КОНСТАНТИНОВИЧ,
ПИСАРЕВ СЕРГЕЙ ГРИГОРЬЕВИЧ*

ПРОЕКТИРОВАНИЕ НЕЧЕТКИХ ПРАВИЛ В БАЗЕ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ОЦЕНКИ УРОВНЯ КОМПЕТЕНТНОСТИ ОПЕРАТОРОВ ДОСМОТРА НА ОСНОВЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

АННОТАЦИЯ

В статье предложен новый подход к оценке уровня компетентности операторов досмотра. Предлагается применить нейросетевые технологии, а также аппаратные методы психофизиологического мониторинга состояния оператора в процессе тренажерной подготовки. В качестве аппаратных методов обосновано использование системы регистрации движения глаз и частоту сердечных сокращений для оценки психофизиологической напряженности. Проведены экспериментальные исследования и разработаны нечеткие базы знаний Сугено и Мамдани. Анализ показал, что нечеткая база знаний Сугено по критерию средней квадратической ошибки точнее чем база знаний Мамдани.

Ключевые слова: транспортная безопасность; оператор досмотра; компетентность; нечеткая база знаний Сугено; нечеткая база знаний Мамдани; нейросетевая модель; субтрактивная кластеризация.

*GLADKIH A. A.,
VOLKOV A. K.,
VOLKOV A. K.
PISAREV S.G.*

DESIGNING OF FUZZY RULES IN THE BASIS OF KNOWLEDGE OF THE INTELLECTUAL SYSTEM OF EVALUATION OF THE LEVEL OF COMPETENCE OF OPERATORS ON THE BASIS OF NEURAL NETWORK TECHNOLOGIES

ABSTRACT

The article proposes a new approach to assessing the level of competence of inspectors. It is proposed to apply neural network technologies, as well as instrumental methods of psycho-physiological monitoring of the operator's state in the process of simulator training. As an instrumental method, the use of an eye movement detection system and heart rate to assess psycho-physiological tensions is justified. Experimental studies were conducted and fuzzy knowledge bases of Sugeno and Mamdani developed. The analysis showed that Sugeno's fuzzy knowledge base by the criterion of the mean square error is more accurate than the Mamdani knowledge base.

Keywords: transport security; screening operator; competence; fuzzy knowledge base of Sugeno; Mamdani fuzzy knowledge base; neural network model; subtractive clustering.

Безопасность на транспорте имеет жизненно важное значение для экономики и общества Российской Федерации, и она в значительной степени зависит от качественной работы подразделений транспортной безопасности. Ключевой мерой обеспечения безопасности на объектах транспортной инфраструктуры является процедура досмотра

пассажиров и их багажа. Для обнаружения запрещенных предметов применяется специальная рентгеновская аппаратура. Главной задачей операторов досмотра является интерпретация рентгеновских снимков в целях ответа на вопрос, является ли багаж безвредным или же он содержит запрещенный предмет и, следовательно, он требует

вторичного досмотра (обычно с использованием ручного досмотра и/или детектора обнаружения взрывчатых веществ). Несмотря на постоянное совершенствование рентгеновской аппаратуры, уровень компетентности оператора досмотра является основным фактором, определяющим уровень транспортной безопасности. В связи с этим, актуальной задачей является разработка предложений по совершенствованию процесса профессиональной подготовки данных специалистов. Целью данной работы является обоснование возможности применения нейросетевых технологий, а также аппаратурных методов психофизиологического мониторинга состояния оператора в процессе тренажерной подготовки.

Одним из перспективных направлений совершенствования тренажерных комплексов подготовки операторов досмотра является внедрения механизмов психофизиологического мониторинга состояния обучающихся. В качестве аппаратурных методов предлагается использовать систему регистрации движения глаз (СРДГ) и частоту сердечных сокращений (ЧСС) для оценки психофизиологической напряженности. В работе [1] представлены некоторые результаты применения СРДГ при оценке компетентности операторов. При этом не в полной мере затронут вопрос учета психофизиологической напряженности. В связи с этим

в условиях возрастания количества обрабатываемой информации необходимо применять методы интеллектуального анализа данных. В данной работе предлагается использовать нечеткие базы знаний (НБЗ), спроектированные с применением нейросетевых технологий.

Нечеткая логика – это мощная методология решения проблем со многими приложениями в задачах управления и обработки информации. В общем смысле, нечеткая логика отражает механизм принятия решений человеком, способного работать с приближительными данными и при этом получать точные решения. В отличие от классической логики, которая требует глубокого понимания системы, точных уравнений и точных численных значений, нечеткая логика предполагает альтернативный подход, который позволяет моделировать сложные системы, используя более высокий уровень абстракции, исходящий из наших знаний и опыта. Известны два основных вида НБЗ: тип Сугено и тип Мамдани [2]. Проектирование НБЗ Сугено включает два основных этапа. Первый заключается в нахождении основной структуры НБЗ с помощью субтрактивной кластеризации (рисунок 1). Второй – оптимизация параметров функций принадлежности (ФП) и весов правил с применением ANFIS-алгоритма (ANFIS – Adaptive Network Based Fuzzy Inference System).

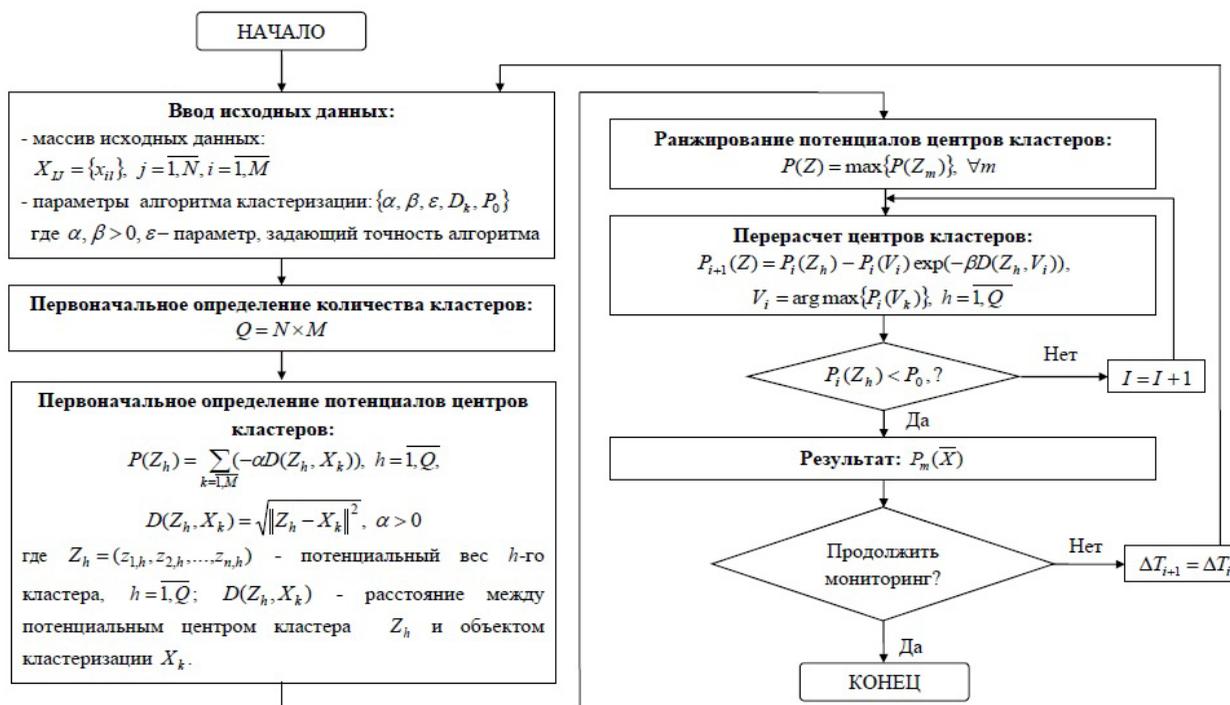


Рисунок 1 – Алгоритм субтрактивной кластеризации

Проектирование НБЗ Мамдани осуществляется на основе алгоритма нечетких *c*-средних (fuzzy *c*-means, FCM), который представлен на рисунке 2.

Таким образом, проектирование НБЗ с применением алгоритмов кластеризации заключается в нахождении кластеров в исходных данных и представлении этих кластеров в виде нечетких правил [3].

В работах [4] и [5] представлены результаты проектирования НБЗ для идентификации зависимостей при решении задач прогнозирования электропотребления и оценки неисправной главного судового дизельного двигателя соответственно.

Для получения экспериментальных данных были протестированы 35 курсантов ФГБОУ ВО «Ульяновский институт гражданской авиации»

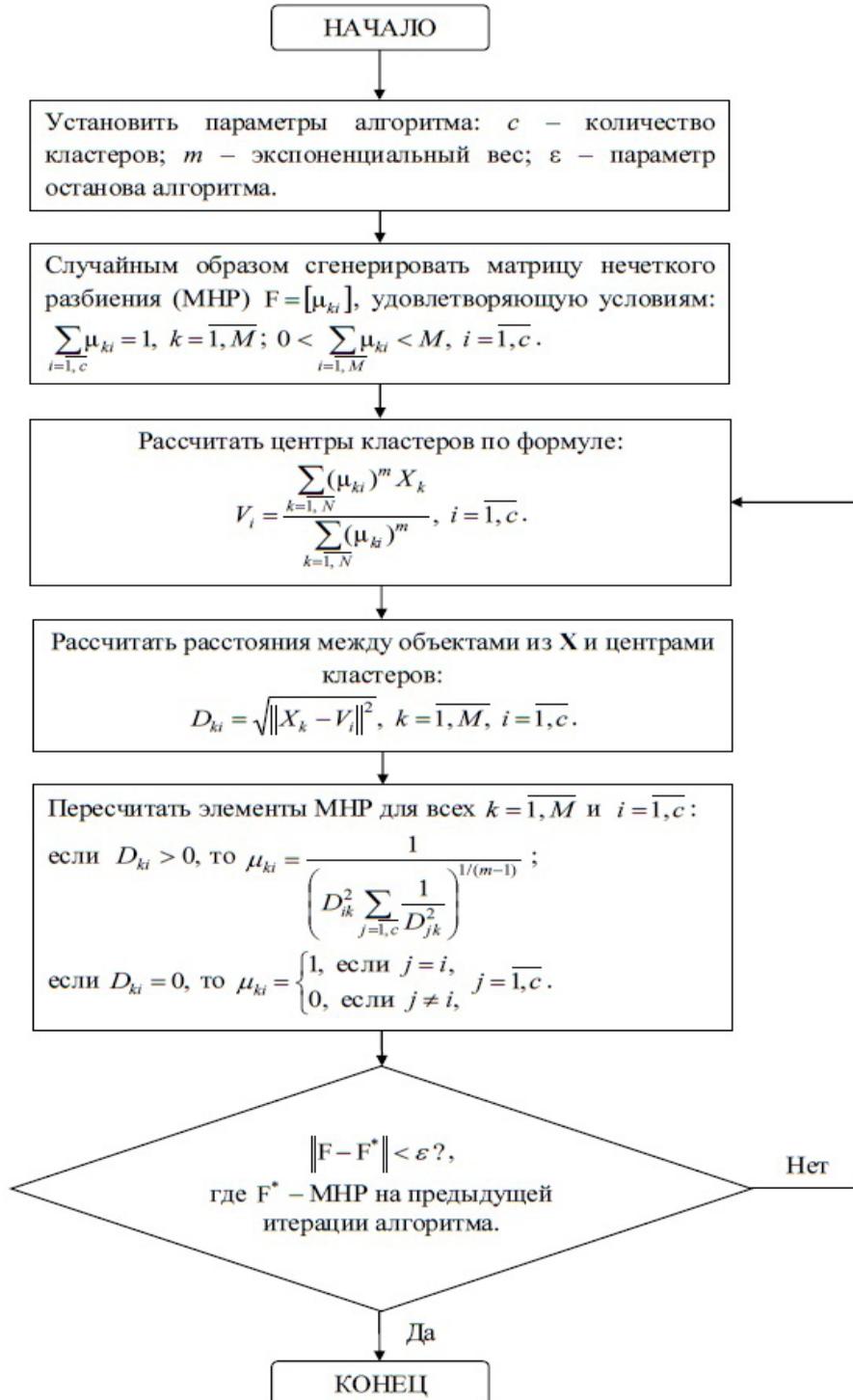


Рисунок 2 – Алгоритм нечетких *c*-средних

имени Главного маршала авиации Б.П. Бугаева» очной формы обучения, направления подготовки 25.03.03 Аэронавигация, профиля подготовки 9. Обеспечение авиационной безопасности. Курсанты прошли необходимое обучение в Авиационном учебном центре и имели сертификат авиационного специалиста. В качестве СРДГ применялся мобильный ай-трекер Eye Tracking Glasses 2.0 фирмы «SensoMotoric Instruments». Для мониторинга ЧСС использовался комплект устройства психофизиологического тестирования Устройство психофизиологического тестирования УПФТ-1/30 «Психофизиолог» (далее «Психофизиолог»).

Экспериментальное тестирование заключалось в следующем:

Во-первых: создание блока тестовых рентгеновских изображений багажа и ручной клади, включающего 20 изображений;

Во-вторых: предобработка изображений с использованием программного обеспечения «SMI BeGaze 3.7»;

В-третьих: тестирование курсантов при решении ими задачи зрительного поиска запрещенных предметов. Рентгеновское изображение предъявлялось с экспозицией в 10 секунд. Во время тестирования регистрировалась глазодвигательная активность и ЧСС;

В-четвертых: формирование набора данных для проектирования НБЗ.

Исходная выборка была поделена на обучающую (30 курсантов) и тестовую (5 курсантов).

Глазодвигательная активность курсантов характеризовалась следующими 3 показателями:

Первый показатель (in1), рассчитывается по формуле (1):

$$in1 = \frac{\sum_{j=1}^k DT_{оп} \cdot \frac{1}{FC_{оп}}}{\sum_{l=1}^m DT_{неоп} \cdot \frac{1}{FC_{оп}} + \sum_{j=1}^k DT_{оп} \cdot \frac{1}{FC_{оп}}}, (1)$$

где $DT_{оп}$ – среднее время удержания взгляда испытуемого в АОИ, содержащую запрещенный к провозу предмет, мс.;

$DT_{неоп}$ – среднее время удержания взгляда испытуемого в АОИ, не содержащую запрещенный к провозу предмет, мс.;

$1/FC_{оп}$ – частота наблюдения j -го опасного предмета;

$1/FC_{неоп}$ – частота наблюдения l -го неопасного предмета;

k – количество тестовых рентгеновских изображений, шт.

n – количество АОИ, содержащих запрещенный к провозу предмет, шт.;

m – количество АОИ, не содержащих запрещенный к провозу предмет, шт.

Второй показатель (in2), рассчитывается по формуле (2):

$$П_2 = \frac{\sum_{i=1}^k SE_{оп}}{k}, (2)$$

где $SE_{оп}$ – индекс, характеризующий схему фиксации внимания испытуемого, равный 0 или 1, в том случае если испытуемый начал схему поиска с АОИ, содержащую опасный или запрещенный к провозу предмет или вещество, показатель равен 1 в ином другом случае показатель равен 0;

k – количество тестовых рентгеновских изображений, шт.

Третий показатель (in3), рассчитывается по формуле (3):

$$П_3 = \sum_{i=1}^k \left(\frac{\sum_{j=1}^n ET_{оп} \cdot S_{оп}}{\sum_{l=1}^m ET_{неоп} \cdot S_{неоп} + \sum_{j=1}^n ET_{оп} \cdot S_{оп}} \right), (3)$$

где $ET_{оп}$ – средняя продолжительность времени от начала эксперимента до первой фиксации в АОИ, содержащей запрещенный к провозу предмет, мс.;

$ET_{неоп}$ – средняя продолжительность времени от начала эксперимента до первой фиксации в АОИ, содержащей запрещенный к провозу предмет, мс.;

$S_{оп}$ – доля площади j -го опасного предмета относительно общей площади стимула;

$S_{неоп}$ – доля площади l -го неопасного предмета относительно общей площади стимула;

k – количество тестовых рентгеновских изображений, шт.;

n – количество АОИ, содержащих запрещенный к провозу предмет, шт.;

m – количество АОИ, не содержащих запрещенный к провозу предмет, шт.

ПФН курсантов была представлено следующим относительным показателем:

Показатель (in4) рассчитывается следующим образом:

$$P_4 = \frac{Y_{\text{ЧСС}}^{\text{Тек}}}{Y_{\text{ЧСС}}^{\text{Фон}}}, \quad (4)$$

где $Y_{\text{ЧСС}}^{\text{Тек}}$ – текущее значение ЧСС оператора и $Y_{\text{ЧСС}}^{\text{Фон}}$ – фоновое значение ЧСС.

Выходная переменная (out1) была представлена частотой обнаружения запрещенных предметов.

Проектирование НБЗ осуществлялось по критерию средней квадратической ошибки (root-mean-square error, RMSE), согласно формуле (5):

$$RMSE = \sqrt{\frac{1}{M} \sum_{r=1}^M (y_r - F(P, W, X_r))} \quad (5)$$

где M – количество пар экспериментальных данных; P – вектор параметров функций принадлежности (ФП) входных переменных (x) и выходной (y); W – вектор весовых коэффициентов правил базы знаний; $F(P, W, X_r)$ – результат расчета нечеткой базой знаний.

В целях робастного оценивания исходных данных была использована программа «Statistica». На рисунке 3 показаны описательные статистики исходных параметров анализа.

Оценка робастности проводилась по следующим параметрам [6]:

- усеченное среднее (среднее значение после удаления выбросов);
- винсоризованное среднее (среднее значение после замены выбросов процентиллю, по которой сделано усечение);
- критерий Граббса для выбросов.

Исходя из рисунка 3 видно, что среднее арифметическое, усеченное среднее и винсоризованное среднее имеют приблизительно равные значения для каждого из оцениваемых параметров. Критическое значение критерия Граббса для уровня

значимости 0,01 составляет 3,33. Для анализируемых параметров критерии Граббса равны 2,207525; 2,006237; 2,603334; 3,036886; 2,001045 и не превышают критического значения; при этом уровни значимости равны 0,795209; 1; 0,219935; 0,038577; 1 соответственно выше выбранного уровня. Таким образом, проведенный анализ позволил обосновать отсутствие выбросов в исходных данных и пригодность их для последующего анализа.

Проведенный факторный анализ исходных данных (методом «варимакс») подтвердил, что все исследуемые показатели несут достаточную факторную нагрузку и могут быть использованы для анализа (рисунок 4).

Проектирование НБЗ осуществлялось в пакете Fuzzy Logic Toolbox системы Matlab. Формирование нечеткого вывода типа Сугено было реализовано с использованием функции genfis2. Были выбраны следующие параметры алгоритма субтрактивной кластеризации: радиус кластеров – 0,7 (значение устанавливается из диапазона [0, 1]); коэффициент подавления – 1,25; коэффициент принятия – 0,5; коэффициент отторжения – 0,15. На рисунке 5 показана схема спроектированной НБЗ Сугено.

Спроектированная НБЗ включает в себя 3 правила, которые характеризуют три найденных кластера:

ЕСЛИ in1=in1cluster1 И in2=in2cluster1 И in3=in3cluster1 И in4=in4cluster1, ТО out1=out1cluster1;

ЕСЛИ in1=in1cluster2 И in2=in2cluster2 И in3=in3cluster2 И in4=in4cluster2, ТО out1=out1cluster2;

ЕСЛИ in1=in1cluster3 И in2=in2cluster3 И in3=in3cluster3 И in4=in4cluster3, ТО out1=out1cluster3.

Переменная	Описательные статистики (Statistica_35.sta)								
	N набл.	Среднее	Усеченн. средн. 5,0000%	Винсориз. средн. 5,0000%	Кр. Граббса Статист.	p-значение	Минимум	Максим.	Ст. откл.
in1	35	2,981565	2,989415	2,984727	2,207525	0,795209	1,998964	3,781094	0,445114
in2	35	0,105714	0,100000	0,105714	2,006237	1,000000	0,000000	0,300000	0,096841
in3	35	2,285672	2,258083	2,293878	2,603334	0,219935	0,623319	4,095842	0,695328
in4	35	0,952817	0,957600	0,954731	3,036886	0,038577	0,805000	1,000000	0,048674
out1	35	0,924286	0,927419	0,924286	2,001045	1,000000	0,850000	0,950000	0,037123

Рисунок 3 – Описательные статистики робастной оценки

Workbook_35.stw - Фактор.нагрузки (Вар.исходн.) (...)

Фактор.нагрузки (Вар.исходн.) (Statistica_35.sta)
Выделение: Главные компоненты
(Отмечены нагрузки >.700000)

Перемен.	Фактор 1	Фактор 2	Фактор 3	Фактор 4
in1	0,938694	0,038410	0,181939	0,290303
in2	0,159900	0,182271	0,967456	-0,072372
in3	-0,032650	-0,953281	-0,190518	0,232149
in4	-0,336575	0,281046	0,091273	-0,894092
Общ. дис.	1,021064	1,022429	1,013701	0,942806
Доля общ.	0,255266	0,255607	0,253425	0,235702

Фактор.нагрузки (Вар.исходн.) (Statistica_35.sta)

Рисунок 4 – Матрица факторных нагрузок

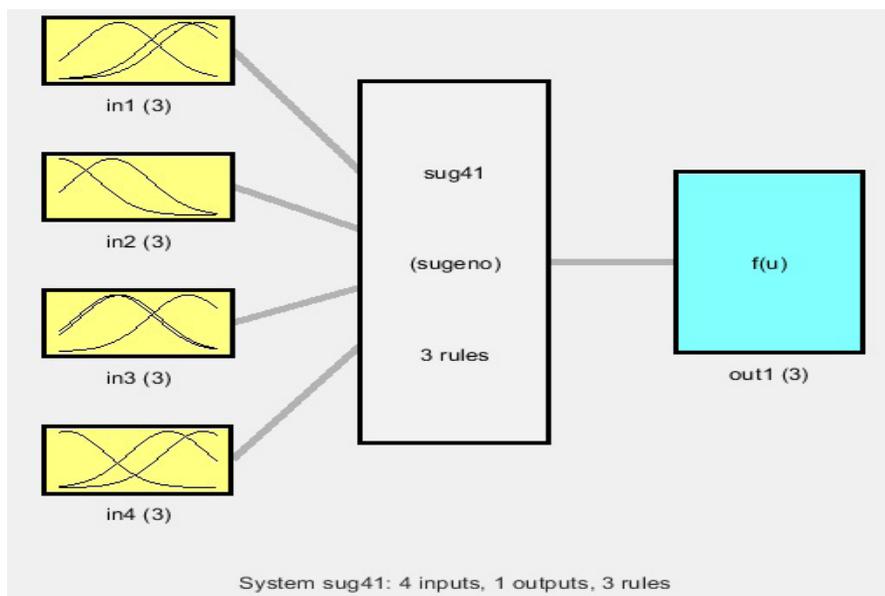


Рисунок 5 – Схема спроектированной нечеткой базы знаний Сугено

Ошибка на обучающей выборке равна – $trnRMSE1=0,0084$, на тестовой – $chkRMSE1=0,0445$.

Рисунок 6 а отражает результаты тестирования НБЗ на обучающей выборке.

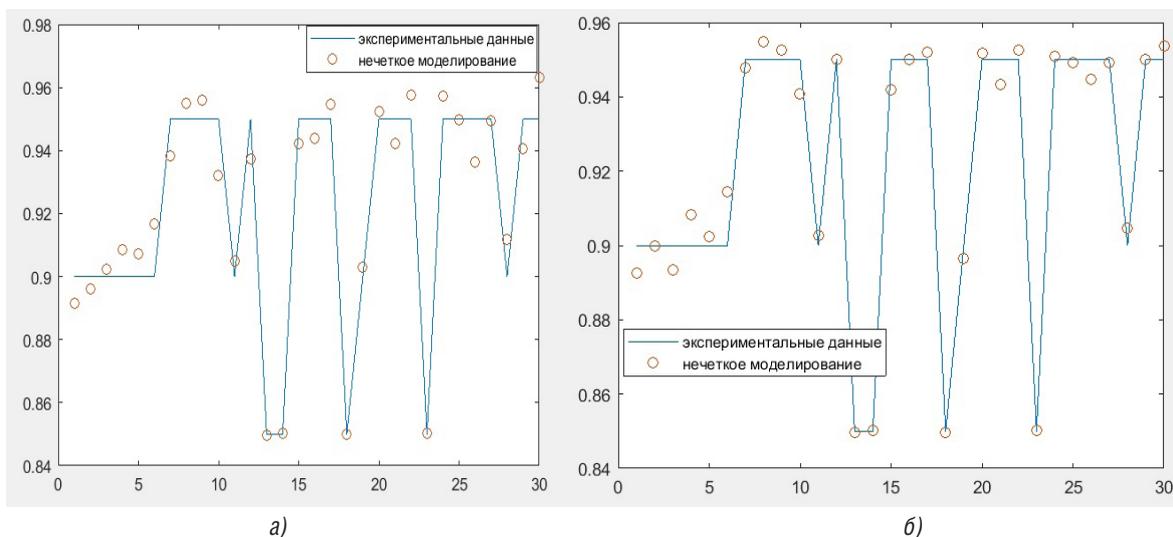


Рисунок 6 – Результаты тестирования модели
а) после субтрактивной кластеризации; б) после ANFIS-обучения

Спроектированная НБЗ типа Сугено на основе субтрактивной кластеризации не вполне точно описывает экспериментальными данными (рисунок 6а). Согласно описанному выше подходу обучим данную модель ее с помощью ANFIS-алгоритма. Количество итераций обучения зададим равное 18. Согласно рисунку 6б можно сделать вывод, что точность модели возросла. Ошибки после ANFIS-обучения равны $trnRMSE2= 0,0049$ и $chkRMSE2= 0,0275$.

На рисунке 7 представлена зависимость значений ошибок от количества итераций ANFIS-алгоритма.

Анализ графика на рисунке 7 позволяет сде-

лать вывод, что ошибка на тестовой выборке достигает наименьшего значения на 7-й итерации ($chkRMSE=0,0246$). При этом ошибка на обучающей выборке снижается на протяжении всех 18 итераций. Полученная нейро-нечеткая модель представлена на рисунке 8.

ФП в посылках правил описываются кривыми гауссовского типа, согласно формуле (6) [2]:

$$\mu(x) = e^{-\frac{(x-b)^2}{2c^2}}, \quad (6)$$

где b – координата максимума ФП;

c – коэффициент концентрации ФП.

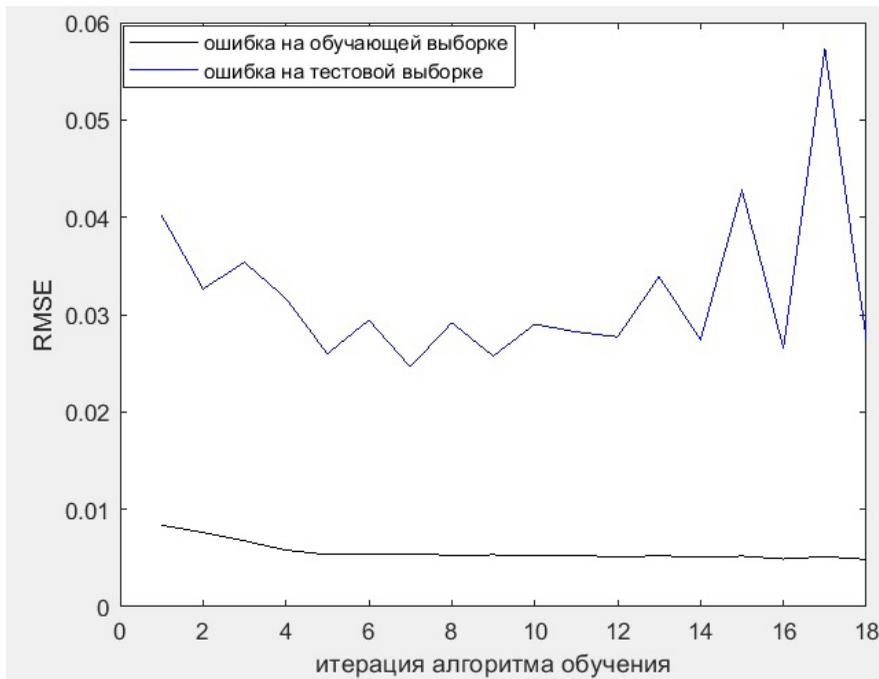


Рисунок 7 – Зависимость значений ошибок от количества итераций ANFIS-алгоритма

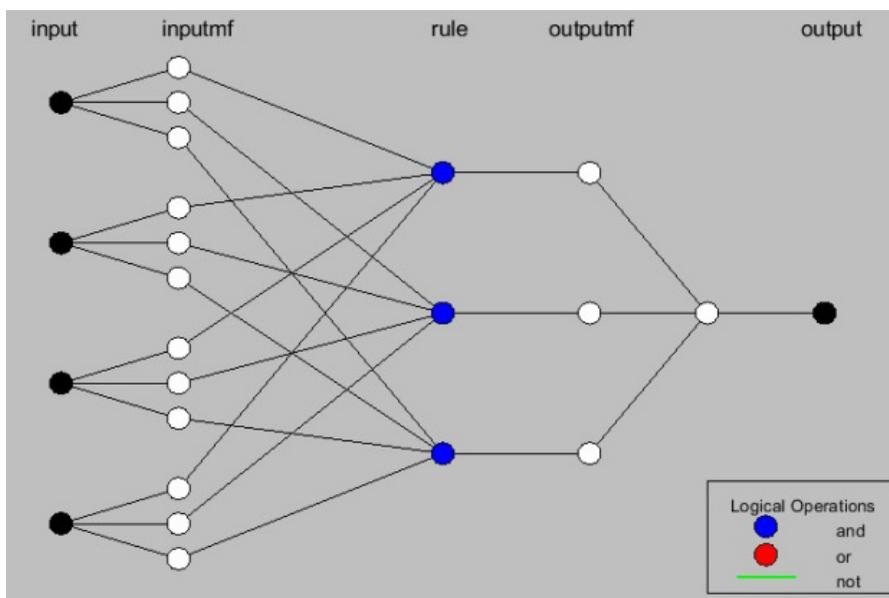


Рисунок 8 – Структура нейро-нечеткой модели

Пример вида ФП для входной переменной in4 нечетким кластерам до обучения и после ANFIS-обучения представлены на рисунке 9. Параметры всех ФП представлены в таблице 1

Полученные значения заключений правил для базы знаний Сугено представлены в таблице 2.

Исходя из таблицы 1 и 2 видно, что в результате ANFIS-обучения были уточнены как параметры ФП, так и параметры в заключениях нечетких правил.

Сравним спроектированную НБЗ Сугено

с базой знаний Мамдани. Проектирование базы знаний Мамдани реализовано с применением функции genfis3. Были выбраны следующие параметры алгоритма FCM: количество кластеров – 3; экспоненциальный вес – 2; значение улучшения целевой функции за одну итерацию – 0,00001; количество итераций – 100. В результате кластеризации получена матрица координат центров нечетких кластеров (каждая строка матрицы описывает центр одного кластера) (таблица 3) и матрица степеней принадлежности объектов кластерам.

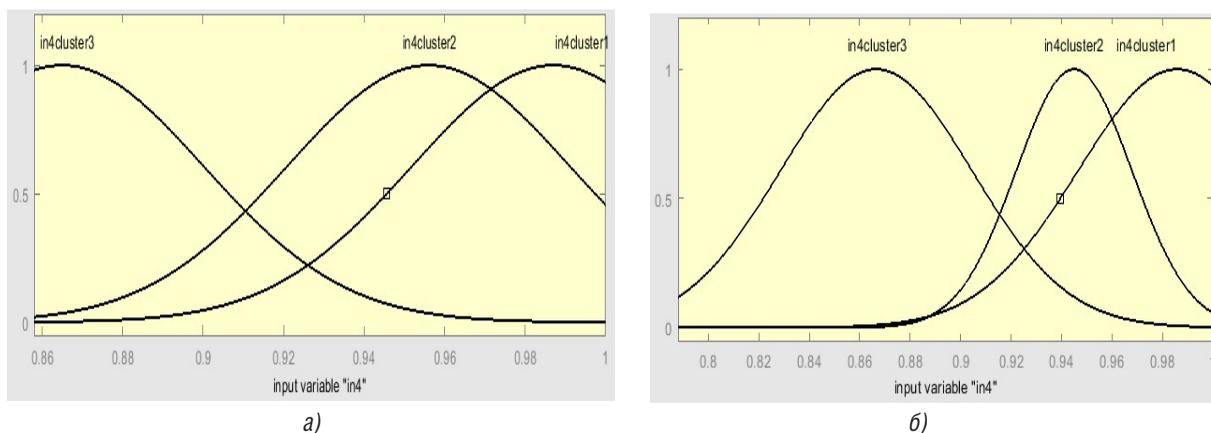


Рисунок 9 – Пример функций принадлежности а) до обучения; б) после ANFIS-обучения

Таблица 1

Параметры ФП НБЗ Сугено

Входная переменная	Кластер	до обучения		после ANFIS-обучения	
		b	c	b	c
in1	cluster1	2,668	0,441	2,671	0,443
	cluster2	3,413	0,441	3,413	0,444
	cluster3	3,573	0,441	3,572	0,441
in2	cluster1	0,1	0,074	0,077	0,066
	cluster2	0,1	0,074	0,126	0,070
	cluster3	0	0,074	0,0005	0,075
in3	cluster1	1,967	0,859	1,967	0,861
	cluster2	1,865	0,859	1,86	0,857
	cluster3	3,452	0,859	3,452	0,860
in4	cluster1	0,987	0,035	0,986	0,039
	cluster2	0,956	0,035	0,945	0,023
	cluster3	0,865	0,035	0,866	0,038

Таблица 2

Заключения НБЗ Сугено

Кластер	Выходная переменная (out1)	
	до обучения	после ANFIS-обучения
1	2	3
cluster1	0,31+0,9432in1-0,2238in2-0,01595in3+0,4891in4	1,281+0,02203in1-0,01662in2-0,00039in3-0,3849in4
cluster2	-0,005155+0,1441in1+0,1199in2+0,001565in3+0,3977in4	-0,154+0,1672in1+0,1938in2+0,003474in3+0,4531in4
cluster3	0,7259+0,003549in1-2,492in2+0,0007605in3+0,126in4	0,7096+0,00613in1+3,61in2+0,002314in3+0,1281in4

Таблица 3

Координаты центров нечетких кластеров

Кластер	Координаты			
cluster1	3,263	0,032	3,276	0,895
cluster2	3,317	0,125	2,008	0,957
cluster3	2,617	0,074	1,944	0,981

В результате из экспериментальных данных экстрагирована база знаний, также содержащая 3 нечетких правила (рисунок 10). Значения ошибок равны $trnRMSE3= 0,0304$ и $chkRMSE3= 0,0349$.

Параметры ФП (гауссовского типа) представлены в таблице 5.

На рисунке 11 представлены примеры ФП нечетким кластерам для одной входной и выходной переменной.

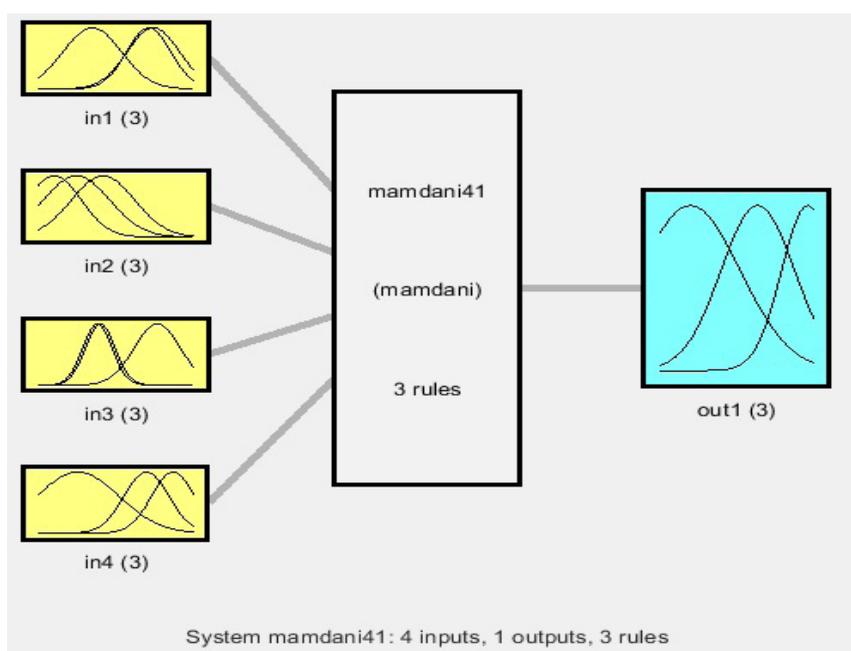
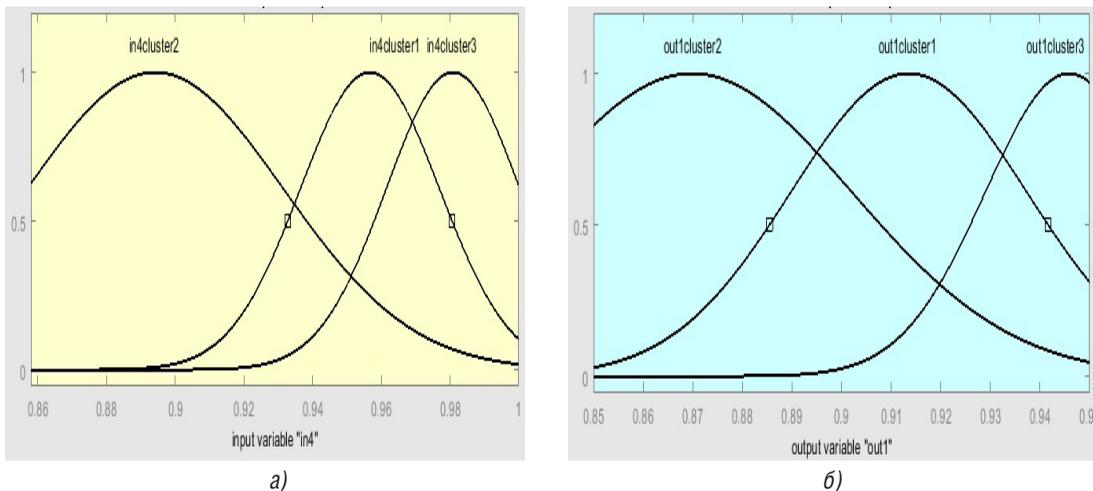


Рисунок 10 – Схема синтезированной нечеткой базы знаний Мамдани

Таблица 5

Параметры функций принадлежности базы знаний Мамдани

Переменная	Кластер	b	c
1	2	3	4
in1	cluster1	3,317	0,299
	cluster2	3,264	0,254
	cluster3	2,617	0,335
in2	cluster1	0,125	0,061
	cluster2	0,052	0,031
	cluster3	0,064	0,074
in3	cluster1	2,008	0,329
	cluster2	3,279	0,516
	cluster3	1,945	0,326
in4	cluster1	0,957	0,020
	cluster2	0,894	0,037
	cluster3	0,981	0,019
out1	cluster1	0,914	0,024
	cluster2	0,0869	0,032
	cluster3	0,946	0,017



а) б)
Рисунок 11 – Функции принадлежности нечетким кластерам

а) ФП нечетким кластерам переменной in4; б) ФП нечетким кластерам переменной out1

Для оценки качества разбиения исходных данных на 3 нечетких кластера рассчитаем индекс Хей-Бени (Хеi-Beni, ХВ) по формуле (7):

$$\chi = \frac{\sum_{i=1, c=1, M} (\mu_{ki})^m \|X_k - V_i\|^2}{M \min_{i \neq j} (\|X_k - V_i\|^2)} \quad (7)$$

Полученное значение индекса ХВ согласно фор-

муле (7) составило 0,5348. Хороший результат разбиения на нечеткие кластеры соответствует критерию ХВ<1. Таким образом, можно сделать вывод, что получено оптимальное разбиение на кластеры.

Согласно таблицы 6 видно, что НБЗ Сугено на основе субтрактивной кластеризации и ANFIS-алгоритма лучше НБЗ Мамдани аппроксимирует исследуемую зависимость.

Таблица 6

Результаты сравнения полученных моделей

Модель	RMSE на обучающей выборке	RMSE на тестовой выборке
Модель Сугено (без ANFIS-обучения)	0,0084	0,0445
Модель Сугено (с ANFIS-обучением)	0,0049	0,0275
Модель Мамдани	0,0304	0,0349

Таким образом, в данной работе предложен и апробирован новый подход к оценке уровня компетентности операторов досмотра, заключающийся в применении аппаратных методов мониторинга психофизиологического состояния обучающихся с последующим проектированием НБЗ.

Список литературы

1. Волков Ан.К., Ионов В.В. Совершенствование организации профессиональной подготовки операторов досмотра с применением системы регистрации движения глаз и методов кластерного и дискриминантного анализа // Научный вестник МГТУ ГА. – 2018. – Т. 21 (3). – С. 45-36.
2. Штовба С.Д. Проектирование нечетких систем средствами MATLAB. – М.: Горячая линия – Телеком, 2007. – 288 с.
3. Штовба С.Д., Мазуренко В.В., Тылец Р.О. Информационная технология нечеткой идентифи-

кации для синтеза точных, компактных и интерпретабельных баз знаний // Computer Science and Telecommunications. – 2016. – № 1 (47). – С. 8-22.

4. Глебов А.А. Модель краткосрочного прогнозирования электропотребления с помощью нейро-нечетких систем // Южно-Российский вестник геологии, географии и глобальной энергии. – 2006. – № 7 (20). – С. 142-146.
5. Полковникова Н.А., Курейчик В.М. Нейросетевые технологии, нечеткая кластеризация и генетические алгоритмы в экспертной системе // Известия ЮФУ. Технические науки – 2014. – №7 (156). – С. 7-15.
6. Гашев С.Н., Бетляева Ф. Х., Лупинос М.Ю. Математические методы в биологии: Анализ биологических данных в системе Statistica: учебное пособие. Тюмень: Издательство Тюменского государственного университета, 2014. – 208 с.

Статья поступила в редакцию 14 января 2019 г.
Принята к публикации 21 марта 2019 г.