

М.П. Кочетков, Д.Н. Корольков, В.Ф. Петров, О.В. Петров,  
С.Б. Симонов, А.И. Терентьев

## ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА С ЭЛЕМЕНТАМИ НЕЧЕТКОЙ ЛОГИКИ ДЛЯ ОЦЕНКИ ОКРУЖАЮЩЕЙ ОБСТАНОВКИ ГРУППЫ РОБОТОТЕХНИЧЕСКИХ СРЕДСТВ

*Кочетков М.П., Корольков Д.Н., Петров В.Ф., Петров О.В., Симонов С.Б., Терентьев А.И.* Применение кластерного анализа с элементами нечеткой логики для оценки окружающей обстановки группы робототехнических средств.

**Аннотация.** Возникновение чрезвычайных ситуаций, которые угрожают жизни и здоровью людей, резко повышает требования к полноте и точности представления информации о текущей ситуации. Современные робототехнические средства оснащены датчиками, работающими на различных физических принципах. Это приводит к росту входной информации, поступающей в управляющую систему. С учетом ограниченной производительности бортовой вычислительной системы, а также высокой априорной неопределенности наземной обстановки робототехнические средства не могут быть эффективно использованы без объединения получаемой информации от группы робототехнических средств и создания единой картины наземной обстановки.

Решить задачу отождествления вектор-признаков, относящихся к одному объекту, а также оценить эффективность полученных решений можно по известным формулам теории проверки статистических гипотез и теории вероятностей только при нормальном законе распределения с известными математическим ожиданием вектор-признака и корреляционной матрицей. Однако перечисленные условия на практике, как правило, не выполняются.

Предложен новый метод решения задачи отождествления вектор-признаков, не опирающийся на статистический подход, и, следовательно, не требующий знания вида закона распределения и значений его параметров. Предлагаемый метод основан на идее сочетания кластерного анализа и нечеткой логики и отличается сравнительно простой по отношению к базовым методам многомерной непараметрической статистики.

Обсуждаются математические аспекты метода нечеткой кластеризации и возможное упрощение алгоритма нечеткого отождествления при временных ограничениях. Установлено, что применение нечеткой кластеризации объектов в сложной наземной обстановке позволяет уменьшить количество ложных распознаваний объектов по сравнению с существующим статистическим подходом, ориентированным на использование нормального закона распределения.

Показано преимущество предлагаемого метода отождествления вектор-признаков объектов, даны сравнительные значения по количеству ложных распознаваний. Даны рекомендации построения правил нечеткого вывода при создании базы знаний экспертной системы.

**Ключевые слова:** нечеткая логика, признак, отождествление целей, робототехническое средство.

**1. Введение.** Наличие множества источников информации, которые работают на различных физических принципах, обладают разными техническими характеристиками и ведут мониторинг обстановки в зонах экологических и техногенных катастроф и в районах проведения контртеррористических операций, позволяет поставить задачу со-

здания обобщенной картины наземной обстановки в области, контролируемой несколькими робототехническими средствами (РТС) [1-3].

Каждый РТС функционирует в автономном режиме, самостоятельно собирает информацию в доступном окружающем пространстве, осуществляет ее первичную обработку и выделяет признаки обнаруженных объектов. При этом распознаванию могут подлежать различные подвижные объекты, например легковые и грузовые автомобили, автобусы, военная техника, а также другие средства, состав которых определяется на этапе формирования задач для группы РТС. Для увеличения дальности обнаружения таких объектов, обеспечения всепогодности, круглосуточности и высокой помехоустойчивости отдельные источники информации должны объединяться в единую информационную систему, построенную на базе группы РТС [4, 5].

Информационные потоки в группе из трех РТС, объединенных в группу с пунктом управления (ПУ), который также взаимодействует с другими источниками информации, например беспилотным летательным аппаратом (БЛА) и, возможно, с соседним пунктом управления (ПУ), показаны на рисунке 1.

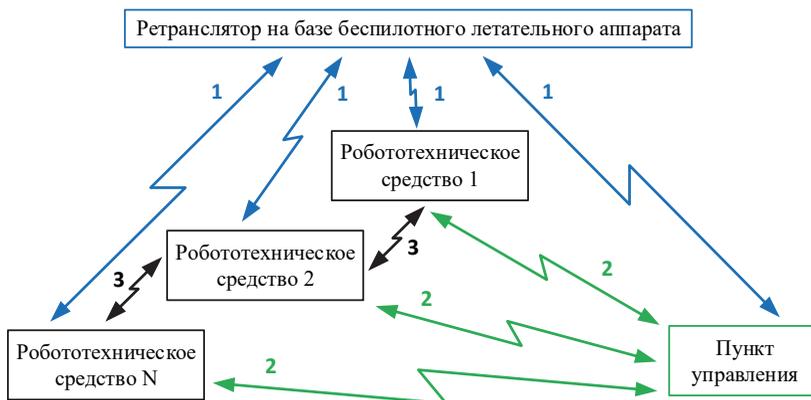


Рис. 1. Информационные потоки в группе РТС:  
1) поток ПУ-БЛА-РТС; 2) поток ПУ-РТС; 3) поток РТС-РТС

Нужно учитывать тот факт, что сведения об одном и том же объекте наблюдения (ОН) – время обнаружения, координаты и параметры движения, признаки объекта, – формируются в сообщение и могут поступать на пункт дистанционного управления (ПДУ) одновременно от разных источников, зоны наблюдений которых пересекаются. В идеальном случае такие сообщения после обработки должны

совпадать, однако на практике этого не происходит из-за погрешностей определения координат ОН бортовыми средствами РТС; временных задержек при передаче данных по каналам связи «РТС-ПДУ», погрешностей экстраполяции траекторий движения на ПДУ и ряда других факторов. В этих условиях решение задачи минимизации возможного ущерба и повышения эффективности определения окружающей обстановки невозможно без объединения (отождествления) информации, поступающей на ПДУ от различных источников. Предполагается, что отождествлению сообщений предшествует их приведение к единому времени и в единую систему координат [6].

**2. Решение задачи отождествления.** Решение задачи отождествления сообщений включает следующие этапы, показанные на рисунке 2.

На этапе 1 производится предварительное отождествление сообщений, поступивших от разных РТС, путем сравнения содержащихся в них координат объектов. Например, координаты объектов в сообщениях от первого РТС сравниваются с соответствующими координатами объектов от других РТС.

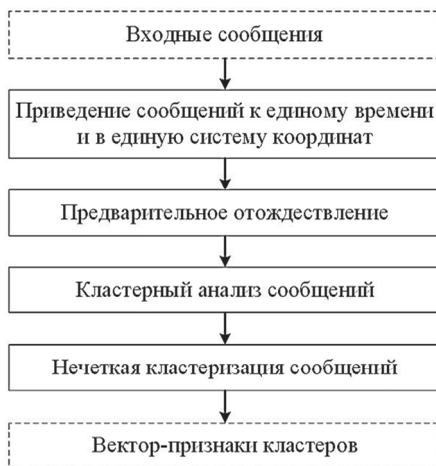


Рис. 2. Обработка сообщений при «мягком» отождествлении

Очевидно, если сообщения получены по одному и тому же ОН, то разность координат с большой вероятностью не будет превышать допустимое отклонение, определяемое ошибками обработки информации. Проведя попарное сравнение координат объектов для всех возможных вариантов, получим множество пар предварительно отождествленных сообщений.

Далее следует по степени близости координат объединять пары в группы, ассоциированные с различными объектами. Допустимые отклонения выбираются так, чтобы вероятность попадания в конкретную группу всех без исключения сообщений, относящихся к одному ОН, была близка к единице. Однако такие отклонения получаются достаточно большими, учитывая ошибки первичной обработки, а также преобразования и экстраполяции координат объектов. Поэтому некоторые группы могут включать сообщения, относящиеся к различным объектам. Следовательно, количество сформированных групп из-за ошибок отождествления может отличаться от количества реальных объектов.

На втором этапе отождествления осуществляется дополнительная обработка сообщений в группах для установления соответствий между сообщениями от разных источников. Допустим, что в некоторую группу входят три сообщения:  $s_1$  от первого РТС,  $s_2$  и  $s_3$  от второго РТС. Тогда возникают две гипотезы отождествления сообщений: первая гипотеза – отождествляются  $s_1$  с  $s_2$ , вторая –  $s_1$  с  $s_3$ .

Исследуем более подробно вопрос отождествления объектов с учетом дополнительной информации об их признаках, содержащихся в принятых сообщениях, для повышения достоверности принимаемого решения. В частности, предположим, что на ПДУ от двух РТС поступают сообщения о признаках, представляемые в виде векторов-признаков  $s_1 = \{x_{11}, x_{21}\}$ ,  $s_2 = \{x_{12}, x_{22}\}$ ,  $s_3 = \{x_{12}^*, x_{22}^*\}$ , где первый индекс соответствует номеру признака, а второй – номеру РТС. Оценим степень близости указанных векторов-признаков в признаковом пространстве, для чего используем многомерный статистический анализ. Выбор предпочтительного варианта можно реализовать путем проверки статистических гипотез. Плотность вероятности евклидова расстояния между векторами-признаками  $s_1$  с  $s_2$  для первой гипотезы вычислим по формуле:

$$w(s_1, s_2) = \frac{1}{2\pi \left[ (\sigma_{11}^2 + \sigma_{12}^2)(\sigma_{21}^2 + \sigma_{22}^2) \right]^{\frac{1}{2}}} \times \exp \left[ -\frac{1}{2} \left( \frac{(x_{11} - x_{12})^2}{\sigma_{11}^2 + \sigma_{12}^2} + \frac{(x_{21} - x_{22})^2}{\sigma_{21}^2 + \sigma_{22}^2} \right) \right],$$

где  $\sigma_{k1}^2, \sigma_{k2}^2$  ( $k = 1, 2$ ) – дисперсии  $k$ -го признака объекта от первого и второго РТС соответственно. Аналогично вычисляется  $w(s_1, s_3)$  для вто-

рой гипотезы. Проверая выполнение неравенства  $w(s_1, s_2) \geq w(s_1, s_3)$ , можно либо принять, либо опровергнуть первую гипотезу.

Недостатком указанного подхода является допущение о том, что признаки, входящие в состав сообщений, имеют нормальное распределение, которое не всегда реализуется на практике. Это обусловлено рядом нестохастических факторов, а именно: неоднородным фоном; априорной неопределенностью появления ОН под определенным ракурсом относительно средств наблюдения; разнообразием помех как искусственного, так и естественного происхождения. Задача усложняется, если на ПДУ выявлено не два, а множество сообщений, содержащих вектор-признаки,  $S = \{s_1, s_2, \dots, s_n\}$ , поступивших от  $n$  РТС (источников информации). Рассмотрим более подробно содержание третьего и четвертого этапов.

**3. Применение кластерного анализа и нечеткой логики.** При помощи методов кластерного анализа данные могут быть разделены на группы «схожих» данных, которые именуют кластерами, а также разнесены в различные сегменты или группы на основе близости их свойств [7, 8].

Возможна ситуация, когда входные сообщения могут быть сгруппированы в кластеры, внутри которых будут находиться «похожие» сообщения, сильно отличающиеся от сообщений других кластеров. В этом случае кластеризация сравнительно просто выполняется четким алгоритмом с-средних. Однако указанная ситуация является исключением и, как правило, не встречается на практике.

В процессе применения методов кластерного анализа требуется сократить объем информации и область поиска рациональных вариантов для решения задачи отождествления сообщений, тем самым уменьшив требуемое время оценки окружающей обстановки.

В большем числе алгоритмов кластеризации не используются стандартные допущения, поэтому разумно применять их в случаях, когда законы распределения данных неизвестны или известны лишь частично [9-11]. Кластеризация проводится для следующих типов объектов: с качественными, количественными или смешанными признаками.

С целью устранения отмеченного недостатка и оценки качества принимаемых решений предлагается метод «мягкого» отождествления сообщений, основанный на сочетании кластерного анализа и нечеткой логики. Пусть каждое принятое сообщение  $s_i$ , кроме координат и параметров движения, содержит  $m$  признаков объекта  $x_{1i}, x_{2i}, \dots, x_{mi}$ , которые могут интерпретироваться точкой в  $m$ -мерном признаковом пространстве  $E^m(x)$ .

Проведение кластерного анализа включает подготовительный этап и этап кластеризации. На подготовительном этапе совокупность признаков, принадлежащих разным объектам в принятых сообщениях, упорядочивается и записывается в виде следующей матрицы «объект-признак»:

$$X_{m \times n} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}.$$

Вектор-признак, содержащийся в сообщении  $s_i$ , представлен  $i$ -м столбцом этой матрицы, то есть  $x_i = \{x_{1i}, x_{2i}, \dots, x_{mi}\}^T$ , где  $T$  – знак транспонирования;  $x_{ki}$  – значение  $k$ -го признака объекта в сообщении  $s_i$ .

Полагаем, что признаки, относящиеся к одному объекту в разных сообщениях, находятся достаточно близко в пространстве  $E^m(x)$  и поэтому должны отождествляться. Для оценки степени близости вектор-признаков необходимо ввести меру расстояния (сходства) между ними. Поскольку признаки измеряются в различных шкалах, то для их однообразного описания требуется провести нормализацию. В частности, элементы матрицы «объект-признак» нормализуем следующим образом:

$$x_{ki}^* = \frac{x_{ki}}{\max_{1 \leq l \leq n} x_{kl}}, k = \overline{1, m}.$$

Расстояния между вектор-признаками  $x_p$  и  $x_q$  в пространстве признаков  $E^m(x)$  можно определить несколькими способами. Наиболее часто выбирают относительное евклидово расстояние:

$$d(x_p, x_q) = \sqrt{\frac{1}{m} \sum_{k=1}^m (x_{kp}^* - x_{kq}^*)^2},$$

для которого справедливо соотношение  $0 \leq d(x_p, x_q) \leq 1$ . Как альтернативный вариант может использоваться следующее соотношение:

$$d(x_p, x_q) = \frac{1}{m} \sum_{k=1}^m (x_{kp}^* - x_{kq}^*)^2.$$

Другим возможным вариантом является обобщенное относительное расстояние Хемминга:

$$d(x_p, x_q) = \frac{1}{m} \sum_{k=1}^m |x_{kp}^* - x_{kq}^*|,$$

однако его целесообразно использовать, когда алгоритм оптимального отождествления вектор-признаков не использует операцию дифференцирования.

Далее сформируем на основе матрицы «объект-признак»  $X_{m \times n}$  другую матрицу «объект-объект»  $D_{m \times n} = [d_{kl}]$ , элементами которой являются расстояния  $d_{kl}$  между векторами-признаками  $x_k$  и  $x_l$ , где  $k, l = 1, \dots, n$ . Введем сокращенную запись  $d(x_p, x_q) = d_{qp}$ ,  $p = \overline{1, n}$ ,  $q = \overline{1, n}$ , после чего матрицу «объект-объект» запишем в виде:

$$D_{n \times n} = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{pmatrix},$$

причем  $d_{gg} = 0$  и  $d_{pg} = d_{gp}$ .

В том случае, если от одного источника приходят несколько сообщений, то матрица  $D_{n \times n}$  модифицируется путем добавления соответствующего количества столбцов и строк, а также указанием запрета на отождествление сообщений от этого источника.

Применительно к вышерассмотренному примеру, когда сообщение  $s_1$  поступает от первого РТС, а сообщения  $s_2$  и  $s_3$  от второго РТС, матрица  $D_{n \times n}$  может иметь вид:

$$D_{3 \times 3} = \begin{pmatrix} 0 & 0,3 & 0,4 \\ 0,3 & 0 & Z \\ 0,4 & Z & 0 \end{pmatrix},$$

где  $d_{12} = d_{21} = 0,3$ ;  $d_{13} = d_{31} = 0,4$ ;  $d_{23} = d_{32} = Z$ . Здесь символ «Z» означает запрет на отождествление сообщений  $s_2$  и  $s_3$ , которые поступают от одного источника.

*Рассмотрим этап кластеризации сообщений.* Отождествление сообщений, представленных матрицей  $D_{n \times n}$ , можно свести к задаче распределения вектор-признаков по группам (кластерам). Для решения этой задачи будем использовать методы кластерного анализа. Цель кластеризации – объединить полученные вектор-признаки в кластеры, каждый из которых формируется по принципу схожих признаков в пределах одного класса. Необходимо заметить, что в большем числе алгоритмов кластеризации не используются стандартные допущения, поэтому разумно применять их в случаях, когда законы распределения данных не известны или известны лишь частично [12].

Отличие между кластеризацией и классификацией состоит в следующем. Классификация относит вектор-признак к одной из нескольких предварительно установленных групп, называемых классами. Кластеризация, напротив, разделяет вектор-признаки на группы, количество которых определяется в процессе обработки данных. Редко, но это допустимо в ряде приложений кластерного анализа, когда точно известно, сколько кластеров следует выделить. Но чаще всего этот вопрос остается открытым перед процедурой кластеризации.

Существуют общие эмпирические рекомендации для выбора количества кластеров. Например, двух или трех кластеров, как правило, недостаточно, так как кластеризация будет слишком грубой, приводящей к потере полезной информации. Напротив, если кластеров больше 10, то эксперту трудно умозрительно обосновать их количество. Таким образом, недостатком традиционных методов кластеризации при решении задачи отождествления сообщений, одновременно поступающих от разных РТС, является требование задания количества кластеров. В общем случае число таких кластеров неизвестно.

Методы кластерного анализа можно условно разделить на иерархические и итерационные. Применение иерархических методов предполагает четкое указание количества кластеров. Среди итеративных методов существуют методы, которые не требуют первоначального задания числа кластеров, однако нужно анализировать алгоритмы, меняя значение двух переменных – порог близости для объединения вектор-признаков в кластеры или число образуемых кластеров. После чего появляется возможность выбрать наилучшее разбиение по выбранному критерию качества.

Сравнительный анализ различных методов кластер-анализа свидетельствует, что для группирования сообщений, поступающих от не-

скольких РТС, преимуществом обладает модифицированный метод горной кластеризации [13]. Выбор указанного метода обусловлен отсутствием необходимости иметь информацию о количестве классов. Для этого алгоритма требуется, чтобы число потенциальных центров кластеров было конечным. В рассматриваемом случае это условие выполняется, так как число центров кластеров не может превышать общего количества сообщений, поступивших на ПДУ от группы РТС. Следовательно, данный алгоритм отличается тем, что количество кластеров определяется самим алгоритмом, центры кластеров здесь выбираются из множества экспериментальных данных.

Рассмотрим кратко суть выбранного метода [14]. *На первом шаге* определяются точки в пространстве признаков (центры кластеров). Если же определить их нельзя, то потенциальными центрами могут быть точки, определяемые всеми вектор-признаками входных сообщений  $s_i$ ,  $i = \overline{1, n}$ .

*Второй шаг* предназначен для вычисления потенциала всех точек. При этом преследуется цель оценить целесообразность создания новых кластеров с центрами, обладающими высоким значением потенциала. Считается, что потенциал точки зависит от количества соседних точек и их удаления. По мере удаления соседних точек потенциал центральной точки ослабляется. Вычисление потенциала в точке  $z_i$  производится по формуле:

$$P(z_i) = \sum_{k=1}^n \exp[-ad(z_i, x_k)],$$

где  $d(z_i, x_k)$  – расстояние между точками  $z_i$  и  $x_k$  в пространстве признаков;  $a$  – положительная константа, которая равна единице, деленной на среднее расстояние между точками.

Название метода обусловлено ассоциациями с горным рельефом местности, возникающим при рассмотрении графика потенциала для точек, кластеризация которых производится по двум признакам.

*На третьем шаге* из множества точек с высоким потенциалом путем нескольких итераций определяются центры кластеров. Процедура последовательного уточнения потенциалов и селекции центров кластеров прекращается, когда уровень потенциалов оставшихся точек становится ниже предварительно заданного порога. В нашем случае он выбирался на уровне 0.55 от потенциала первой максимальной вершины.

Стоит отметить, что 0.55 – это рекомендуемое значение порога для конкретных условий рассматриваемой задачи. Здесь есть некото-

рая аналогия с выбором порога при распознавании зашумленного сигнала, когда приходится искать компромисс между ошибками первого и второго родов. В нашем случае малый заданный уровень – много ложных кластеров, большой порог – потеря реальных кластеров. Подробное описание горного алгоритма приведено в работах [15-17].

Таким образом, результатом проведенного кластерного анализа является выделение вектор-признаков, являющихся центрами кластеров, после чего появляется возможность распределения остальных вектор-признаков по образованным кластерам (рис. 3).

Будем полагать, что вектор-признак  $x_i$  считается принадлежащим кластеру  $V_k$ , который образован вектор-признаком  $x_k$ , если расстояние  $d(x_i, x_k)$  меньше, чем расстояние от  $x_i$  до других кластеров. В этом случае все входные вектор-признаки будут принадлежать каким-либо кластерам.

Если же допустить, что  $x_i$  считается принадлежащим кластеру  $V_k$ , когда расстояние до него  $d(x_i, x_k)$  меньше, чем некоторое пороговое значение, то могут появиться вектор-признаки, не принадлежащие ни одному из найденных кластеров.

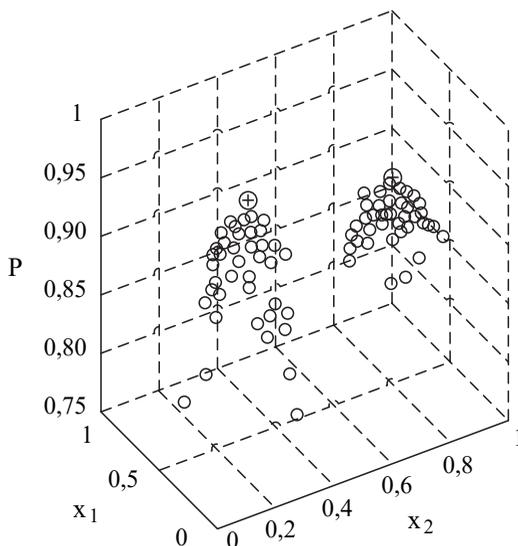


Рис. 3. Результат кластерного анализа – выделение центров кластеров

К достоинствам алгоритма горной кластеризации можно отнести его простоту, которая, однако, компенсируется недостатками –

невысокой наглядностью представления и низкой точностью [18, 19]. Кластеризация по горному алгоритму является четкой, при которой вектор-признак попадает только в один кластер. Такое ограничение не всегда приемлемо в системах специального назначения. Для частично видимых объектов, чьи признаки расположены на границе кластеров, более «естественной» будет именно нечеткая кластеризация.

**4. Использование методов нечеткой логики.** Сформулируем задачу нечеткой кластеризации. Известно множество вектор-признаков  $X = \{x_1, x_2, \dots, x_n\}$ . Необходимо произвести разбиение  $X$  на нечеткие кластеры, которыми являются нечеткие подмножества  $A_i, \left( i = \overline{1, c} \right)$ .

Введем обозначение  $\mu_{ik} = \mu_{A_i}(x_k)$  – степень принадлежности вектор-признака  $x_k \in X$  к подмножеству  $A_i$ . Матрицу степеней принадлежности представим в виде  $M = [\mu_{ik}]$ , при этом будем полагать:

$$\mu_{ik} \in [0, 1], \quad i = \overline{1, c}, \quad k = \overline{1, n},$$

$$\sum_{i=1}^c \mu_{ik} = 1, \quad k = \overline{1, n}.$$

Последнее условие в нечеткой кластеризации обусловлено тем обстоятельством, что вектор-признак может относиться к разным кластерам с определенными значениями функции принадлежности. В силу чего возникает необходимость нормализации суммы степеней принадлежности.

В настоящее время широкое распространение в различных приложениях получил FCM-алгоритм нечеткой кластеризации [20-22]. Применение этого алгоритма предполагает определение матрицы степеней принадлежности  $M$  и центров кластеров  $V = \{v_1, v_2, \dots, v_c\}$ , которые удовлетворяют определенному критерию. Наиболее часто требуется минимизировать критерий:

$$\sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m d^2(x_k, v_i) \rightarrow \min,$$

где  $d(x_k, v_i)$  – евклидово расстояние между двумя векторами-признаками  $x_k$  и  $v_i$  или соответствующими точками в многомерном

пространстве признаков;  $m$  – весовой коэффициент фаззификации. Коэффициент  $m$  усиливает влияние вектор-признаков с высокими значениями степеней принадлежности и ослабляет иные. FCM-алгоритм является итерационным алгоритмом. При его выполнении требуется многократно повторить определенную последовательность действий с изменяющимся количеством кластеров для различных исходных нечетких разбиений множества вектор-признаков на нечеткие подмножества. FCM-алгоритм подразумевает выполнение ряда шагов [23]:

Шаг 1. Активация исходного нечеткого разделения вектор-признаков.

Шаг 2. Определение центров всех кластеров. Для вычисления центра  $i$ -го кластера используется формула:

$$v_i = \frac{\sum_{k=1}^n \left( (\mu_{ik})^m x_k \right)}{\sum_{k=1}^n (\mu_{ik})^m},$$

где  $\mu_{ik}$ , по сути, является весовым коэффициентом, который определяет вес вектор-признаков  $x_k$ ,  $k = \overline{1, n}$ , зависящий от величины параметра  $m$ , при вычислении значений  $v_i$ ,  $i = \overline{1, c}$ .

Шаг 3. Вычисление новых функций принадлежности:

$$\mu_{ik} = \left[ \sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1},$$

где  $d_{ik} = d(x_k, v_i)$ ;  $d_{jk} = d(x_k, v_j)$ .

Следует учитывать тот факт, что с ростом коэффициента  $m$ , конечная матрица, сформированная FCM-алгоритмом, «размывается», а при  $m \rightarrow \infty$  все вектор-признаки уже принадлежат всем кластерам с одинаковой степенью принадлежности.

Шаги 2 и 3 будут циклически повторяться. Завершение цикла возможно при выполнении одного из двух условий: либо исполнится количество циклов, определенное заранее, либо будет достигнута требуемая точность вычислений.

Для сокращения вычислительных затрат предлагается число кластеров и значения их центров, полученные алгоритмом горной кластеризации, использовать как приближительные начальные значения в FCM-алгоритме.

Распределив входные вектор-признаки по группам FCM-алгоритмом, получим несколько кластеров. Кластеры описываются своими центрами, то есть вектор-признаками того же пространства признаков, которому принадлежит исходное множество вектор-признаков. После чего можно воспользоваться построенной базой знаний, позволяющей путем логического вывода оценить степень принадлежности вектор-признаков центров кластеров к каждому классу распознаваемых объектов.

**5. Построение и использование нечеткой базы знаний.** Проблема отождествления вектор-признаков в условиях высокой априорной неопределенности окружающей обстановки не может быть решена на основе аппарата математической статистики. Однако ее решение возможно в рамках теории нечеткого логического вывода, которая оперирует качественными понятиями, учитывает характер мышления опытного специалиста и ход его рассуждений при выработке решений.

Большой сложностью при создании базы знаний экспертной системы на нечеткой логике является необходимость явно сформулировать правила принятия решений в форме продукций.

Важным понятием нечеткой логики является лингвистическая переменная (ЛП), которая в общем случае определяется несколькими термами. Например, признак «длина контура» ОН может быть представлен термами: «маленькая», «средняя», «большая». При построении нечеткой базы знаний требуется по обучающей выборке определить нечеткие множества, соответствующие различным термам ЛП [24-26]. Формализуем поставленную задачу.

Пусть определены наименования термов для всех признаков, кроме того указаны области их определения. Имеется обучающая выборка (ОВ) «множество значений признаков – номер класса». Обозначим вектор обучающей выборки:  $z = \{z_1, z_2, \dots, z_m\}$ . Необходимо определить нечеткие множества, представляющие термы каждого признака  $z_j \cdot (1 \leq j \leq m)$  для распознавания ОН  $h$  классов  $q_1, q_2, \dots, q_h$ .

Разобьем множество векторов ОВ на  $M$  групп по количеству классов. В каждую группу включим векторы, относящиеся к одному классу. Таким образом, будут образованы группы:  $L_1, L_2, \dots, L_M$ .

Допустим, что признак  $z_1$  включает  $V$  термов  $T_1 = \{T_{11}, T_{12}, \dots, T_{1V}\}$ , определенные на множестве  $U = \{u\}$ , и необходимо построить нечеткое множество:

$$\tilde{T}_1 = \left( \frac{\mu_{T_{11}}(u)}{u}, \frac{\mu_{T_{12}}(u)}{u}, \dots, \frac{\mu_{T_{1V}}(u)}{u} \right),$$

где  $\mu_{T_{1j}}(u)$ ,  $j = \overline{1, V}$  – функция принадлежности (ФП) значения  $u$  признака  $z_1$  к терму  $T_{1j}$ ,  $\left( j = \overline{1, V} \right)$  при распознавании объектов класса  $q_1$ .

Рассмотрим последовательность действий для вычисления функций принадлежности термов лингвистической переменной.

1) Определим диапазон  $\Delta z$  изменения признака  $z_1$  при распознавании ОН классов  $q_1, q_2, \dots, q_h$ .

2) Разделим  $\Delta z$  на непересекающиеся интервалы. Учитывая фактор неравномерного распределения значений признака  $z_1$ , получим  $N$  интервалов.

3) Строим таблицу, имеющую  $N$  столбцов и  $V$  строк для термов признака  $z_1$ .

4) Эксперт анализирует значение признака  $z_1$  из примера и определяет, к какому из  $V$  термов необходимо его отнести.

5) Выбирается строка, соответствующая этому терму, и добавляется единица в ячейку, принадлежащую одному из  $N$  столбцов. Номер столбца зависит от номера интервала, в который попало значение этого признака. Для определенности полагаем, что номер столбца равен  $k$ .

6) Нормируем числа в каждом столбце таблицы относительно суммы чисел, стоящих в этом столбце. В результате получим оценку ФП элементов нечеткого множества  $\tilde{T}_1$ . Например, степень принадлежности к терму  $T_{1j}$  значения  $u$ , попавшего в  $k$ -ый интервал, вычисляется следующим образом:

$$\tilde{\mu}_{T_{1j}}(u) = \frac{n_{jk}}{n_k}, \quad j = \overline{1, V},$$

где  $n_{jk}$  – количество примеров ОВ для распознавания класса  $q_1$ , подсчитанное в ячейке на пересечении  $j$ -ой строки и столбца с номером  $k$ ;  $n_k$  – сумма чисел в столбце таблицы с номером  $k$ .

Указанная последовательность действий повторяется для построения ФП термов признака  $z_2$  по множеству  $L_2$ , ФП термов признака  $z_3$  по  $L_3$  и так далее.

После того как термы различных признаков для распознавания объектов будут представлены в виде нечетких множеств, переходим к проектированию базы знаний [27, 28]. Введем обозначения:  $\beta$  – лингвистическая переменная,  $\alpha$  – ее значение. Для классификации применим логические высказывания вида: « $\beta$  ЕСТЬ  $\alpha$ », полагая, что  $\alpha$  определено на базовом терм-множестве указанной лингвистической переменной.

Используем нечеткий логический вывод для упрощения модели композиционного правила вывода Заде. Упрощение основывается на предположении, что установленные входные лингвистические переменные имеют известные числовые значения, а также наличие только одной выходной лингвистической переменной, а именно: «класс объекта».

Классификация ОН  $h$  классов  $q_1, q_2, \dots, q_h$  по  $m$  признакам  $z_1, z_2, \dots, z_m$  соответствует отображению вида:

$$z = (z_1, z_2, \dots, z_m) \rightarrow q \in \{q_1, q_2, \dots, q_h\}$$

и может производиться на основании следующих правил логического вывода:

ЕСЛИ ( $z_1$  ЕСТЬ  $T_{11}$ ) И ( $z_2$  ЕСТЬ  $T_{21}$ ) И ...И ( $z_m$  ЕСТЬ  $T_{m1}$ )

ТО ( $q$  ЕСТЬ  $q_1$ )

ЕСЛИ ( $z_1$  ЕСТЬ  $T_{12}$ ) И ( $z_2$  ЕСТЬ  $T_{22}$ ) И ...И ( $z_m$  ЕСТЬ  $T_{m2}$ )

ТО ( $q$  ЕСТЬ  $q_2$ )

...

ЕСЛИ ( $z_1$  ЕСТЬ  $T_{1V}$ ) И ( $z_2$  ЕСТЬ  $T_{2V}$ ) И ...И ( $z_m$  ЕСТЬ  $T_{mV}$ )

ТО ( $q$  ЕСТЬ  $q_V$ ),

где  $T_{ij}$  – лингвистический терм (нечеткая переменная, представленная нечетким множеством) для оценки признака  $z_i$  при распознавании объектов класса  $q_j$ .

Условие ( $z_i \in T_{ij}$ ) реализуется функцией фаззификации, которая представляется обобщенной гауссовой функцией отдельно для каждой переменной  $z_i$  и соответствующего термина  $T_{ij}$ , где  $j = \overline{1, V}$ .

Пусть требуется классифицировать объект по совокупности  $m$  конкретных значений признаков  $u_1, u_2, \dots, u_m$ . Тогда для каждого  $u_i, i = \overline{1, m}$  следует определить значение функции принадлежности к нечетким множествам, представленными терминами  $T_{ik}, \left( k = \overline{1, h} \right)$ . Степень истинности нечеткого высказывания « $u_i$  ЕСТЬ  $T_{ik}$ » определяется значением функции принадлежности  $\mu_{T_{ik}}(u_i)$  для конкретного значения  $u_i$ , как показано на рисунке 4.

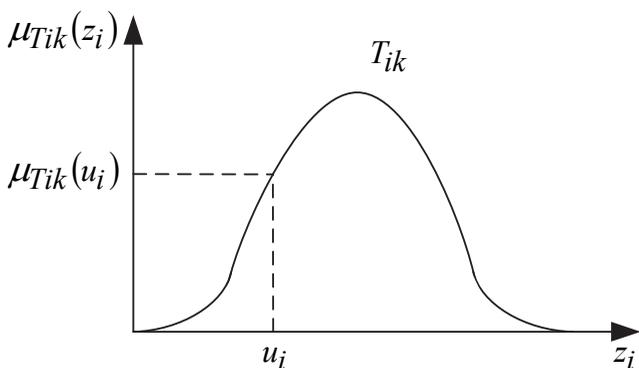


Рис. 4. Пример введения нечеткости с использованием четкого значения  $u_i$  признака

После определения степени принадлежности конкретных значений признаков к терминам, указанным в  $j$ -ом правиле, вводим для краткости обозначение  $t_{ij} = \mu_{T_{ij}}(u_i), i = \overline{1, m}$ . Логическую связку «И» будем интерпретировать как Т-норму нечетких множеств или операцию их пересечения. Тогда степень истинности предпосылок  $j$ -го правила можно вычислить следующим образом:

$$t_j = t_{1j} \cap t_{2j} \cap \dots \cap t_{mj}.$$

Полученное значение  $t_j$  применяется для оценки  $u_j(q)$  – истинности заключения этого правила или того, что входной вектор  $(z_1, z_2, \dots, z_m)$  соответствует классу  $q_j$ . Считается, что  $u_j(q) = t_j$ . Аналогичным образом определяются степени истинности для всех других правил, представленных в базе знаний. После чего предпочтение отдается тому классу, которому соответствует правило с максимальным значением истинности заключения. Номер искомого класса определяется по формуле:

$$k = \arg \max_{1 \leq p \leq h} [t_p].$$

В условиях априорной неопределенности для построения правил нечеткого логического вывода может непосредственно использоваться обучающая выборка прецедентов [29-31]. Рассмотрим подход к решению задачи распознавания ОН, заданного вектором информативных признаков  $z = (z_1, z_2, \dots, z_m)$ , к одному из  $h$  классов  $\{q_1, q_2, \dots, q_h\}$  на основании информации, содержащейся в обучающей выборке.

Для классификации требуется создать нечеткую базу знаний, содержащую  $h$  правил, каждое из которых включает информацию о прецедентах в обучающей выборке по соответствующему классу объектов. Пусть классификация объектов класса  $q_j$  производится по экспертной информации из  $k_j$  примеров. Тогда правило распознавания будет состоять из системы нечетких логических высказываний, представленных  $k_j$  следующими строками логических связей «И»:

ЕСЛИ  $(z_1$  ЕСТЬ  $T_{1j,1}$ ) И  $(z_2$  ЕСТЬ  $T_{2j,1}$ ) И ... И  $(z_{mj,1}$  ЕСТЬ  $T_{nj,1}$ )  
с весом  $w_{j,1}$  ИЛИ

ЕСЛИ  $(z_1$  ЕСТЬ  $T_{1j,2}$ ) И  $(z_2$  ЕСТЬ  $T_{2j,2}$ ) И ... И  $(z_{mj,2}$  ЕСТЬ  $T_{nj,2}$ )  
с весом  $w_{j,2}$  ИЛИ

...

ЕСЛИ  $(z_1$  ЕСТЬ  $T_{1j,k_j}$ ) И  $(z_2$  ЕСТЬ  $T_{2j,k_j}$ ) И ... И  $(z_{mj,k_j}$  ЕСТЬ  $T_{nj,j}$ )  
с весом  $w_{j,k_j}$

ТО  $q = q_j, j = 1, 2, \dots, h,$

где  $T_{ij,p}$  – лингвистический терм, которым оценивает признак  $z_i$  при распознавании объектов класса  $q_j$  в  $p$ -ой строке правила;  $p = \overline{1, k_j}$ ;  $k_j$  – количество правил (прецедентов), использованных для описания класса  $q_j$ ,  $w_{j,p}$  – весовой коэффициент, ( $0 < w_{j,p} \leq 1$ ) означающий экспертную оценку достоверности или уверенность в степени истинности логического условия в  $p$ -ой строке правила. Другими словами, коэффициент  $w_{j,p}$  определяет значимость условия, по умолчанию  $w_{j,p} = 1$ .

Выбирая в качестве  $T$ -нормы пересечение по Заде и учитывая весовой коэффициент, получим взвешенную степень истинности условия о принадлежности входного вектора признаков классу  $q_j$  по условиям  $p$ -ой строки:

$$\mu_{j,p}(Z^*) = w_{j,p} \times \min_{i=1,m} [\mu_{jp}(Z_i^*)].$$

Взвешенная степень истинности принадлежности вектора  $Z^* = (z_1^*, z_2^*, \dots, z_n^*)$  классу  $q_j$  по условиям всех  $k_i$  строк:

$$\mu_{qi}(Z^*) = \max_{p=\overline{1,k_i}} \left\{ w_{jp} \times \min_{i=1,n} [\mu_{jp}(Z_i^*)] \right\},$$

где  $\mu_{jp}(Z_i^*)$  – функция принадлежности  $Z_i^*$  нечеткому терму  $T_{i,jp}$ .

В качестве решения задачи распознавания выбираем класс объектов с максимальной взвешенной степенью принадлежности:

$$d = \arg \max_{q_j \in \{q_1, q_2, \dots, q_h\}} [\mu_{qj}(Z^*)].$$

Представленный подход предполагает использование набора четких значений информативных признаков и может применяться в системах автоматического распознавания при наличии соответствующих датчиков и средств обработки первичной информации.

Рассмотренный подход к созданию единой картины наземной обстановки в зоне действия группы РТС и последующей клас-

сификации ОН позволяет путем автоматического сбора и объединения информации [32], поступающей от различных источников, повысить дальность обнаружения, а также достоверность отождествления и, следовательно, распознавания ОН в недетерминированных условиях.

**6. Моделирование предложенных методов.** Для оценки эффективности нового метода решения задачи отождествления проводился компьютерный эксперимент с двумя системами идентификации. В ходе эксперимента предлагаемый метод сравнивался с классическим методом статистической теории распознавания образов на основе гипотезы нормального закона распределения. Эффективность оценивалась количеством ложных распознаваний. При моделировании использовались различные законы распределения поступающих вектор-признаков, причем предполагалось, что вид закона распределения системе идентификации неизвестен.

Численные оценки эффективности зависят не только от метода решения задачи идентификации, а также от многих других факторов (количества РТС в группе, взаимного расположения их на местности, секторов обзора систем технического зрения, дальности обнаружения объектов, их общего количества и т.д.). Поэтому было принято решение проводить сравнение методов при одних и тех же условиях функционирования системы идентификации.

Как и ожидалось, метод статистической теории распознавания оказался эффективнее предлагаемого, когда вектор-признаки имели нормальное распределение. Наиболее близким по виду функциональной зависимости к нормальному закону среди других рассматриваемых законов является релейский закон. Экспериментальным путем получено, что эффективность предлагаемого метода при релейском законе относительно классического метода приблизительно на 18% выше. При других видах формульных зависимостей законов распределения эффективность обоих исследуемых методов снижалась, однако предлагаемый метод имел эффективность на 18% выше классического подхода.

Предложенные методы промоделированы в среде Matlab, имеющей два способа для работы с кластеризацией: при помощи командной строки или же графического интерфейса [33, 34].

Центры кластеров определяются при помощи функции «fcm». Для работы в функцию необходимо передать аргументы:

- 1) data – данные, которые должны быть кластеризованы; точка в многомерном пространстве признаков задается одной строкой;
- 2) NC — количество кластеров;

3) options — дополнительные опции кластеризации.

Результаты работы функции отражаются в следующих переменных:

1) centers – матрица центров кластеров, каждая строка которой содержит координаты центра отдельного кластера;

2) U – результирующая матрица функциональных признаков;

3) obj\_Func – значение целевой функции на каждой итерации.

В работе использовался следующий программный код для нечеткой кластеризации:

```
// загрузка данных
load fd.dat;
// определение центра кластеризации
[centers, U, objFunc] = fcm(fd, 2);
// определение максимальной степени принадлежности
max_U = max(U);
// распределение строк матрицы данных между кластерами
ind1 = find(U(1, :) == max_U);
ind2 = find(U(2, :) == max_U);
// построение данных
plot(fcmdata(ind1, 1), fcmdata(ind1, 2), 'ko', 'markersize', 5, 'LineWidth', 1);
hold on
// построение данных
plot(fcmdata(ind2, 1), fcmdata(ind2, 2), 'kx', 'markersize', 5, 'LineWidth', 1);
// построение кластерных центров
plot(center(1, 1), center(1, 2), 'ko', 'markersize', 15, 'LineWidth', 2)
plot(center(2, 1), center(2, 2), 'kx', 'markersize', 15, 'LineWidth', 2)
```

Можно отметить, что для расширения возможностей анализа получаемых результатов функцию кластеризации можно вызывать с дополнительными параметрами. Эти параметры управляют процессом кластеризации:

– opt\_1 – показатель степени для матрицы U (по умолчанию – 2.0);

– opt\_2 – максимальное количество итераций (по умолчанию – 100);

– opt\_3 – предельное изменение значений целевой функции (по умолчанию –  $1e^{-5}$ );

– opt\_4 – отображение информации на каждом шаге (по умолчанию – 1).

Для визуализации получаемых результатов можно использовать встроенный в Matlab инструмент «findcluster». Графический

интерфейс пользователя позволяет загрузить исходные данные для кластеризации в следующем формате: одна строка – одна точка в многомерном пространстве характеристик. Количество строк равно количеству элементов данных. Оператор также определяет тип алгоритма кластеризации.

Для запуска алгоритма кластеризации необходимо указать параметры:

- количество кластеров (Cluster Num);
- максимальное количество итераций (Max Iteration);
- минимальное значение целевой функции (Min. Improvement);
- показатель степени при матрице ФП ( Exponent).

Для координат найденных центров матрицы доступна функция сохранения в виде матрицы, что позволяет пользоваться ими в дальнейшем. Строка матрицы в файле – это набор координат кластера. Количество строк соответствует количеству кластеров.

**7. Заключение.** Исследован вопрос отождествления наземных объектов, с повышением достоверности принимаемого решения, с учетом дополнительной информации об их признаках, содержащихся в принятых сообщениях. Показано, что недостатком статистического подхода является допущение о том, что признаки, входящие в состав сообщений, имеют нормальное распределение, которое не всегда реализуется на практике.

Предложен новый метод «мягкого» отождествления вектор-признаков, поступающих от группы РТС на пункт управления, который основан на сочетании кластерного анализа и нечеткой логики и который может эффективно использоваться в условиях высокой априорной неопределенности относительно законов распределения входных данных для оценки окружающей обстановки в зоне действия группы РТС.

Рассмотрено создание базы знаний экспертной системы на основе нечёткой логики в форме продукции. Для построения правил нечеткого вывода рекомендуется воспользоваться обучающей выборкой прецедентов, заданных векторами информативных признаков.

Проведенное в среде Matlab моделирование информационных процессов, протекающих в системе «группа РТС-ПДУ» показало, что предлагаемый метод отождествления вектор-признаков объектов на базе нечеткой логики и кластерного анализа уменьшает количество ложных распознаваний на 18% по сравнению с существующим статистическим подходом, основанным на предположении, что вектор-признак подчиняется нормальному закону распределения.

**Литература**

1. Zhou J., Bai X., Caelli T. Computer Vision and Pattern Recognition in Environmental Informatics // IGI Global. 2016. 437 p.
2. Kim J. et al. Robot Intelligence Technology and Applications // Springer. 2018. 579 p.
3. Shih F.Y. Image processing and pattern recognition Fundamentals and Techniques // Wiley & Sons. 2010. 552 p.
4. Ермолов И.Л., Хрипунов С.П., Хрипунов С.С. Проблемы группового применения робототехнических комплексов военного назначения и пути их решения // Материалы XII мультиконференции по проблемам управления (МКПУ-2019). 2019. С. 136–138.
5. Абросимов Э.А., Куличенко А.Д., Можжев А.Н., Смирнова Е.Ю. Перспективные технологии организации группового взаимодействия и группового управления для гибридных групп мобильных роботов // Материалы XII мультиконференции по проблемам управления (МКПУ-2019). 2019. С. 116–118.
6. Панасюк Ю.Н., Пудовкин А.П. Обработка радиолокационной информации в радиотехнических системах // ФГБОУ ВПО «ГТТУ». 2016. 84 с.
7. Härdle W.K., Simar L. Applied Multivariate Statistical Analysis. 5th edition // New York: Springer. 2019. 550 p.
8. Morlini I., Minerva T., Vichi M. Advances in Statistical Models for Data Analysis // Springer. 2015. 268 p.
9. Wierchoń S.T., Kłopotek M.A. Modern Algorithms of Cluster Analysis // Springer. 2018. 433 p.
10. Bouveyron C. et al. Model-Based Clustering and Classification for Data Science: With Applications in R // Cambridge University Press. 2019. 447 p.
11. Seetha H., Murty M.N., Tripathy B.K. Modern Technologies for Big Data Classification and Clustering // IGI Global. 2018. 382 p.
12. Егоров А.И., Курьянова Н.И. Особенности методов кластеризации данных // Известия ЮФУ. Технические науки. 2011. №11. С. 174–177.
13. Yager R., Filev D. Essentials of Fuzzy Modeling and Control // John Wiley & Sons. 1984. 387 p.
14. Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику. URL: [http://matlab.exponenta.ru/fuzzy\\_logic/book1/index.php](http://matlab.exponenta.ru/fuzzy_logic/book1/index.php) (дата обращения: 20.08.2019).
15. Пташко Е.А., Ухоботов В.И. Автоматическая генерация нечетких правил для управления мобильным роботом с гусеничным шасси на основе числовых данных // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2017. Т. 6. № 3. С. 60–72.
16. Прудковский Н.С., Беспечная Д.Э. Кластеризация данных лувенским методом и методом к-средних // Молодой исследователь: вызовы и перспективы сб. ст. по материалам LXXV междунар. науч.-практ. конф. 2018. С. 159–166.
17. Аль-Раммахи А.А.Х.Х., Сари Ф.А.А., Минин Ю.В. Модификация метода нечеткой кластеризации с-средних с использованием метода роя частиц для обработки больших данных // Материалы 1-ой Всероссийской (национальной) научно-практической конференции «Современная наука: теория, методология, практика». 2019. С. 231–233.
18. Янчуковский В.Н. Использование параллельных вычислений в кластерном анализе для формирования комплексных деталей // Вестник Иркутского государственного технического университета. 2012. № 6(65). С. 25–30.
19. Подколзина Л.А. Системный анализ данных с применением иерархического алгоритма кластеризации в совокупности с методами аппарата нечеткой логи-

- ки // Системный анализ, управление и обработка информации материалы VI международного семинара. 2015. С. 110–113.
20. *Ярушикина Н.Г.* Нечеткие нейронные сети в когнитивном моделировании и традиционных задачах искусственного интеллекта // М.: МИФИ. 2005. 214 с.
  21. *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.
  22. *Петров В.Ф. и др.* Сетевая структура обработки информации в распределенных системах управления наземными робототехническими комплексами // Изв. вузов. Электроника. 2018. Т. 23, № 4. С. 389–398.
  23. *Демидова Л.А., Кираковский В.В., Пылькин А.Н.* Принятие решений в условиях неопределенности // Горячая линия – Телеком. 2015. 283 с.
  24. *Mendel J.M.* Uncertain Rule-based Fuzzy Systems // Springer. 2017. 696 p.
  25. *Fidanova S.* Recent Advances in Computational Optimization // Springer. 2019. 239 p.
  26. *Hooda D.S., Raich V.* Fuzzy Logic Models and Fuzzy Control: An Introduction // Alpha Science International Ltd. 2017. 409 p.
  27. *Melin P., Castillo O., Kacprzyk J., Reformat M., Melek W.* Fuzzy Logic in Intelligent System Design // Springer. 2018. 415 p.
  28. *Nguyen H.T., Walker C.L., Walker E.A.* A First Course in Fuzzy Logic. 4th ed. // CRC Press. 2018. 458 p.
  29. *Tanveer M. et al.* Machine Intelligence and Signal Analysis // Springer. 2018. 757 p.
  30. *Woyczyński W.A.* A First Course in Statistics for Signal Analysis. 3rd ed. // Birkhäuser. 2019. 337 p.
  31. *Ley C., Verdebout T.* Applied Directional Statistics Modern Methods and Case Studies // CRC Press. 2019. 317 p.
  32. *Серегин М.Ю.* Интеллектуальные информационные системы : учебное пособие // ФГБОУ ВПО «ГГТУ». 2012. 205 с.
  33. *Kochetkov M., Terentev A., Chernovolenko A.* A hybrid model of the 3D-objects recognition system // 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIconRus). 2018. pp. 1521–1524.
  34. *Kochetkov M., Terentev A., Chernovolenko A.* Fuzzy Model of Recognition of Environmental Objects for a Mobile Robotic Complex // Proceedings of the 2019 IEEE Conference of Russia Young Researchers in Electrical and Electronic Engineering, EIconRus 2019.

**Кочетков Михаил Петрович** — канд. техн. наук, доцент, доцент, Институт микроприборов и систем управления имени Л.Н. Преснухина, Национальный исследовательский университет «МИЭТ» (НИУ МИЭТ). Область научных интересов: нечеткая логика, кластерный анализ, методы идентификации. Число научных публикаций — 110. kmp@miec.ru; пл. Шокина, 1, 124498, Москва, Зеленоград, Россия; р.т.: +7(499)720-8751.

**Корольков Дмитрий Николаевич** — ведущий инженер-электроник, Институт микроприборов и систем управления имени Л.Н. Преснухина, Национальный исследовательский университет «МИЭТ» (НИУ МИЭТ). Область научных интересов: электронные вычислительные устройства. Число научных публикаций — 29. jimmy@olvs.miec.ru; пл. Шокина, 1, 124498, Москва, Зеленоград, Россия; р.т.: +7(499)720-8751.

**Петров Владимир Федорович** — канд. техн. наук, заместитель директора, Институт микроприборов и систем управления имени Л.Н. Преснухина, Национальный исследова-

тельский университет «МИЭТ» (НИУ МИЭТ). Область научных интересов: вооружение и военная техника, системы управления, робототехнические комплексы. Число научных публикаций — 83. [pvf@olvs.miee.ru](mailto:pvf@olvs.miee.ru); пл. Шокина, 1, 124498, Москва, Зеленоград, Россия; р.т.: +7(499)720-8751.

**Петров Олег Владимирович** — канд. техн. наук, начальник отдела, Институт микроприборов и систем управления имени Л.Н. Преснухина, Национальный исследовательский университет «МИЭТ» (НИУ МИЭТ). Область научных интересов: схемотехника устройств вычислительной техники и систем управления; моделирование; измерительные устройства. Число научных публикаций — 21. [rovmiel@yandex.ru](mailto:rovmiel@yandex.ru); пл. Шокина, 1, 124498, Москва, Зеленоград, Россия; р.т.: +7(499)720-8751.

**Терентьев Алексей Игоревич** — канд. техн. наук, доцент, Институт микроприборов и систем управления имени Л.Н. Преснухина, Национальный исследовательский университет «МИЭТ» (НИУ МИЭТ). Область научных интересов: моделирование; информационно-управляющие системы, нечеткая логика, элементы искусственного интеллекта. Число научных публикаций — 43. [terentev@olvs.miee.ru](mailto:terentev@olvs.miee.ru); пл. Шокина, 1, 124498, Москва, Зеленоград, Россия; р.т.: +7(499)720-8751.

**Симонов Сергей Борисович** — канд. техн. наук, начальник отдела, Институт микроприборов и систем управления имени Л.Н. Преснухина, Национальный исследовательский университет «МИЭТ» (НИУ МИЭТ). Область научных интересов: программное обеспечение систем управления, робототехнические системы. Число научных публикаций — 12. [sb.simonov@gmail.com](mailto:sb.simonov@gmail.com); пл. Шокина, 1, 124498, Москва, Зеленоград, Россия; р.т.: +7(499)720-8751.

M. KOCHETKOV, D. KOROLKOV, V. PETROV, O. PETROV, A. TERENCEV,  
S. SIMONOV

**APPLICATION OF CLUSTER ANALYSIS WITH FUZZY  
LOGIC ELEMENTS FOR GROUND ENVIRONMENT  
ASSESSMENT OF ROBOTIC GROUP**

---

*Kochetkov M., Korolkov D., Petrov V., Petrov O., Terentev A., Simonov S. Application of Cluster Analysis with Fuzzy Logic Elements for Ground Environment Assessment of Robotic Group.*

**Abstract.** Emergency situations, that cause risks for human life and health, dictate elevated requirements to completeness and accuracy of the presentation of information about current ground environment. Modern robotic systems include sensors, that operate on different physical principles. This causes incrementation of information entering control system. Computing resources and technical capabilities of robotic systems are limited in range and detection probabilities of appearing objects. In case of insufficient performance of the on-board computer system and high uncertainties of ground environment, robotic systems are not able to perform without combining information from robotic group and producing a single view of ground environment. Complex information from a group of robotic systems occurs in real time and a non-deterministic environment.

To solve the problem of identifying attribute vectors related to a single object, as well as to evaluate the effectiveness of obtained solutions, is possible using known formulas of the theory of statistical hypothesis testing and probability theory only under the normal distribution law with the known mathematical expectation of an attribute vector and a correlation matrix. However, these conditions are usually not met in practice. Problems also arise when methods of nonparametric statistics are used with an unknown law of probability distribution.

The new method of identifying attribute vectors is proposed, that does not rely on a statistical approach and, therefore, does not require knowledge of the type of distribution law and the values of its parameters. Proposed method is based on the idea of combining cluster analysis and fuzzy logic, and is relatively simple to the basic methods of multidimensional nonparametric statistics.

The results of modeling information processes are presented. The advantages of proposed method are shown. The comparative values for the number of false recognitions are given. The recommendations are given for constructing fuzzy inference rules when creating an expert system knowledge base.

**Key words:** Fuzzy Logic, Attribute, Target Authentication, Robotic System.

---

**Kochetkov Mihail** — Ph.D., Associate Professor, Associate Professor, Institute of Microdevices and Control Systems, National Research University of Electronic Technology (MIET). Research interests: fuzzy logic, cluster analysis, identification methods. The number of publications — 110. [kmp@miec.ru](mailto:kmp@miec.ru); 1, Shokin sq., 124498, Moscow, Zelenograd, Russia; office phone: +7(499)720-8751.

**Korolkov Dmitry** — Leading Electronics Engineer, Institute of Microdevices and Control Systems, National Research University of Electronic Technology (MIET). Research interests: electronic computing devices. The number of publications — 29. [jimmy@olvs.miec.ru](mailto:jimmy@olvs.miec.ru); 1, Shokin sq., 124498, Moscow, Zelenograd, Russia; office phone: +7(499)720-8751.

**Petrov Vladimir** — Ph.D., Deputy Director, Institute of Microdevices and Control Systems, National Research University of Electronic Technology (MIET). Research interests: military technical equipment, control system, robotic systems. The number of publications — 83. [pvf@olvs.micee.ru](mailto:pvf@olvs.micee.ru); 1, Shokin sq., 124498, Moscow, Zelenograd, Russia; office phone: +7(499)720-8751.

**Petrov Oleg** — Ph.D., Head of Department, Institute of Microdevices and Control Systems, National Research University of Electronic Technology (MIET). Research interests: circuit design of computer equipment and control systems; modeling; measuring devices. The number of publications — 21. [povmiet@yandex.ru](mailto:povmiet@yandex.ru); 1, Shokin sq., 124498, Moscow, Zelenograd, Russia; office phone: +7(499)720-8751.

**Terentev Alexey** — Ph.D., Associate Professor, Institute of Microdevices and Control Systems, National Research University of Electronic Technology (MIET). Research interests: modeling; information and control systems, fuzzy logic, elements of artificial intelligence. The number of publications — 43. [terentev@olvs.micee.ru](mailto:terentev@olvs.micee.ru); 1, Shokin sq., 124498, Moscow, Zelenograd, Russia; office phone: +7(499)720-8751.

**Simonov Sergey** — Ph.D., Head of Department, Institute of Microdevices and Control Systems, National Research University of Electronic Technology (MIET). Research interests: control system's software; robotic systems. The number of publications — 12. [sb.simonov@gmail.com](mailto:sb.simonov@gmail.com); 1, Shokin sq., 124498, Moscow, Zelenograd, Russia; office phone: +7(499)720-8751.

## References

1. Zhou J., Bai X., Caelli T. Computer Vision and Pattern Recognition in Environmental Informatics. IGI Global. 2016. 437 p.
2. Kim J. et al. Robot Intelligence Technology and Applications. Springer. 2018. 579 p.
3. Shih F.Y. Image processing and pattern recognition Fundamentals and Techniques. Wiley & Sons. 2010. 552 p.
4. Ermolov I.L., Xripunov S.P., Xripunov S.S. [Problems of group application of military-purpose robotic systems and ways to solve them]. *Materialy` XII mul'tikonferencii po problemam upravleniya (MKPU-2019)* [XII Multiconference on Management Problems (MKPU-2019)]. 2019. pp. 136–138. (In Russ.).
5. Abrosimov E.A., Kulichenko A.D., Mozhaev A.N., Smirnova E.Yu. [Advanced technologies for organizing group interaction and group control for hybrid groups of mobile robots]. *Materialy` XII mul'tikonferencii po problemam upravleniya (MKPU-2019)* [XII Multiconference on Management Problems (MKPU-2019)]. 2019. pp. 116–118. (In Russ.).
6. Panasyuk Yu. N., Pudovkin A.P. *Obrabotka radiolokacionnoj informacii v radio-texnicheskix sistemax* [Processing of radar information in radio engineering systems]. FGBOU VPO "TGTU". 2016. 84 p. (In Russ.).
7. Härdle W.K., Simar L. Applied Multivariate Statistical Analysis. 5th edition. New York: Springer. 2019. 550 p.
8. Morlini L., Minerva T., Vichi M. Advances in Statistical Models for Data Analysis. Springer. 2015. 268 p.
9. Wierzchoń S.T., Kłopotek M.A. Modern Algorithms of Cluster Analysis. Springer. 2018. 433 p.
10. Bouveyron C. et al. Model-Based Clustering and Classification for Data Science: With Applications in R. Cambridge University Press. 2019. 447 p.
11. Seetha H., Murty M.N., Tripathy B.K. Modern Technologies for Big Data Classification and Clustering. IGI Global. 2018. 382 p.

12. Egorov A.I., Kupriyanova N.I. [Features of data clustering methods]. *Izvestiya YuFU. Tekhnicheskie nauki – Izvestiya Sfedu. Engineering Sciences*. 2011. vol. 11. pp. 174–177.
13. Yager R., Filev D. *Essentials of Fuzzy Modeling and Control*. John Wiley & Sons. 1984. 387 p.
14. Shtovba S.D. *Vvedenie v teoriyu nechetkix mnozhestv i nechetkuyu logiku*. [An introduction to fuzzy set theory and fuzzy logic.]. Available at: [http://matlab.exponenta.ru/fuzzy\\_logic/book1/index.php](http://matlab.exponenta.ru/fuzzy_logic/book1/index.php) (accessed: 20.08.2019). (In Russ.)
15. Ptashko E.A. Uxobotov V.I. [Automatic generation of fuzzy rules for the control of a mobile crawler robot based on numerical data]. *Vestnik YuUrGU. Seriya: Vy chislitel'naya matematika i informatika – Bulletin of the South Ural State University. Series "Computational Mathematics and Software Engineering"*. 2017. Issue 6. vol. 3. pp. 60–72. (In Russ.)
16. Prudkovskij N.S., Bespechnaya D.E. [Data clustering by Louvain and k-means]. *Molodoj issledovatel': vy'zovy' i perspektivy` sb. st. po materialam LXXV mezhdunar. nauch.-prakt. konf.* [Young Researcher: Challenges and Prospects Sat. Art. based on materials from LXXV int. scientific-practical conf.]. 2018. pp. 159–166. (In Russ.)
17. Al-Rammaxi A.A.X.X., Sari F.A.A., Minin Yu.V. [Modification of the c-means fuzzy clustering method using the particle swarm method for processing big data]. *Materialy 1-oy Vserossijskoj (nacional'noj) nauchno-prakticheskoy konferencii "Sovremennaya nauka: teoriya, metodologiya, praktika"* [Materials of the 1st All-Russian (national) scientific-practical conference "Modern science: theory, methodology, practice"]. 2019. pp. 231–233. (In Russ.)
18. Yanchukovskij V.N. [Using parallel computing in cluster analysis to generate complex parts]. *Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta – Proceedings of Irkutsk State Technical University*. 2012. vol. 6(65). pp. 25–30. (In Russ.)
19. Podkolzina L.A. [System analysis of data using a hierarchical clustering algorithm in conjunction with the methods of the apparatus of fuzzy logic]. *Sistemnyj analiz, upravlenie i obrabotka informacii materialy VI mezhdunarodnogo seminara* [System analysis, management and information processing materials of the VI international seminar]. 2015. pp. 110–113. (In Russ.)
20. Yarushkina N.G. *Nechetkie neyronnye seti v kognitivnom modelirovanii i tradicionnyx zadachax iskusstvennogo intellekta* [Fuzzy neural networks in cognitive modeling and traditional tasks of artificial intelligence]. M.: MIFI. 2005. 214 p. (In Russ.)
21. Polkovnikova N.A., Kurejchik V.M. [Neural network technologies, fuzzy clustering and genetic algorithms in the expert system]. *Izvestiya YuFU. Tekhnicheskie nauki – Izvestiya Sfedu. Engineering Sciences*. 2014. vol. 7(156). pp. 7–15. (In Russ.)
22. Petrov V.F. et al. [Network structure of information processing in distributed control systems of ground-based robotic complexes]. *Izv. vuzov. Elektronika – Proceedings of Universities. Electronics*. 2018. Issue 23. no. 4. pp. 389–398. (In Russ.)
23. Demidova L.A., Kirakovskij V.V., Pylkin A.N. *Prinyatie reshenij v usloviyax neopredelennosti* [Decision making under uncertainty]. Goryachaya liniya – Telekom. 2015. 283 p. (In Russ.)
24. Mendel J.M. *Uncertain Rule-based Fuzzy Systems*. Springer. 2017. 696 p.
25. Fidanova S. *Recent Advances in Computational Optimization*. Springer. 2019. 239 p.
26. Hooda D.S., Raich V. *Fuzzy Logic Models and Fuzzy Control: An Introduction*. Alpha Science International Ltd. 2017. 409 p.
27. Melin P., Castillo O., Kacprzyk J., Reformat M., Melek W. *Fuzzy Logic in Intelligent System Design*. Springer. 2018. 415 p.
28. Nguyen H.T., Walker C.L., Walker E.A. *A First Course in Fuzzy Logic*. 4th ed. CRC Press. 2018. 458 p.

29. Tanveer M. et al. Machine Intelligence and Signal Analysis. Springer. 2018. 757 p.
30. Woyczyński W.A. A First Course in Statistics for Signal Analysis. 3rd ed. Birkhäuser. 2019. 337 p.
31. Ley C., Verdebout T. Applied Directional Statistics Modern Methods and Case Studies. CRC Press. 2019. 317 p.
32. Seregin, M.Yu. *Intellektualnye informacionnye sistemy: uchebnoe posobie* [Intelligent information systems: a tutorial]. FGBOU VPO “TGTU”. 2012. 205 p. (In Russ.)
33. Kochetkov M., Terentev A., Chernovolenko A. A hybrid model of the 3D-objects recognition system. 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). 2018. pp. 1521–1524.
34. Kochetkov M., Terentev A., Chernovolenko A. Fuzzy Model of Recognition of Environmental Objects for a Mobile Robotic Complex. Proceedings of the 2019 IEEE Conference of Russia Young Researchers in Electrical and Electronic Engineering, EIConRus 2019.