**Лабораторная работа №1**

Используя аппарат нечеткой логики смоделировать распределение участников по возрастным группам на конкурсе.

**Лабораторная работа №2**

**Дано:**

Лингвистическая переменная «Истинность», предложенная Л. Заде:

**Задание:** проиллюстрировать эту идею гранулированных значений истинности на частном примере

**Лабораторная работа №3**

**Постановка задачи:**

Одним из базовых положений применения методов нечеткого управления на практике является задача выявления функций принадлежности для лингвистических координат технологического процесса, которая в данной постановке интерпретируется в виде задачи классификации. Последняя, в свою очередь, может быть интерпретирована как задача распознавания образа (ситуации). Распространяя эти положения на исследуемую предметную область, проблему можно представить как составную часть процедуры идентификации ситуаций и ситуационного управления.

Основные методологические подходы, принятые в настоящее время в теории распознавания, базируются на основах статистики, традиционного кластерного анализа, а также дедукционных выводов в бинарной логике. Однако эти подходы в незначительной степени отражают процесс распознавания, который свойственен человеку вообще и человеку оператору в частности. В связи с этим следует отметить, что выделение признаков – наиболее важный этап в распознавании ситуаций, он же является и исключительно сложным.

Поскольку информация даже о достаточно четком объекте может искажаться за счет шумов, для его распознавания чаще всего применяется нечеткая логика. Типичным примером является классификация элементов с помощью нечеткой кластеризации. Поскольку абсолютно идентичные элементы встречаются редко, необходима «размытая» кластеризация. Аналогичные методы применяются и при классификации образов, имеющих разброс относительно эталонного образа (распознавание рукописных знаков, речи и т.п.)

Кластеризация представляет собой метод разбиения множества разбросанных данных на несколько групп. Разбиение осуществляется так, чтобы данные в одной группе обладали похожими свойствами, а свойства в среднем между группами максимально различались.

Для решения задач классификации Л. Заде предложил использовать новое понятие – свойство размытой близости, которое означает:

1. если у объектов x и y степень сходства велика, то у них высока степень принадлежности к некоторому кластеру;
2. если у объектов x и y степени принадлежности к различным кластерам высоки, то они, то они не обладают высокой степенью сходства.

Кластеры, порожденные процедурой кластеризации, называют размытыми, если они обладают свойством размытой близости. Необходимо отметить, что свойство размытых множеств более сильно, чем то, которое подразумевается в обычных определениях. Для последних требуется, чтобы степень сходства объектов, принадлежащих одному и тому же кластеру, была больше степени сходства между объектами, принадлежащими разным кластерам.

Предлагаемый метод Fuzzy-кластеризации использует в качестве исходных данных пространство состояний *Х* моделируемого ТП

*X = { x1, x2, ..., xn},*

где *xj ∈ Rd- d-*мерный вектор состояния, элементами которого являются значения координат ТП, полученные на этапе сбора технологической информации.

Перед проведением кластеризации можно произвести выбор наиболее информативных координат, на множествах значений которых будет осуществляться кластеризация.

Задача состоит в том, чтобы разбить множество *Х* на *с* кластеров   
(*2≤ с≤n*).

Степень принадлежности xj к k-му кластеру обозначается как ukj. При жесткой кластеризации ukj принимает два значения: 0 или 1, при нечеткой - значение ukj может быть произвольным от 0 до 1:

ukj ∈ {0,1} – при жесткой кластеризации,

ukj ∈ [0,1] – при нечеткой кластеризации.

Но в любом случае должно выполняться условие:

∑j ukj> 0,

Нечеткая кластеризация допускает принадлежность данных к двум или более кластерам, но сумма степеней принадлежности составляет 1 , а u является весом принадлежности к кластеру. Определить степень принадлежности u можно, применяя нечеткую логику.

Пусть *M*fc - множество (c x n)-матриц *U* (называемых матрицами разделения), элементами которых являются степени принадлежности вектора *xj* к *k*-ому кластеру *u*kj. Кластеризация, в данном контексте, это процедура соединения множества данных *X* и матрицы разделения *U*.

В процедуре FCM для нахождения оптимального результата этого соединения за целевую функцию, определяющую качество размытого разбиения, принимается сумма квадратичных ошибок в обобщенной группе

*Jm (U, v)* =  , 1 ≤ *m* < ∞,

где *xj* - *d*-мерные измеренные данные, *vk* - *d-*мерный вектор, центр *k*-го кластера, - произвольная норма, отражающая подобие измеренных данных и центра кластера. В нашем случае целесообразно использовать в качестве нормы евклидово расстояние, как наиболее универсальное.

При *m=1* и *u*kj = {0, 1} процедура минимизации критерия выполняется по обычному методу *k*-средних. Чем больше *m* превышает 1, тем более нечеткой становится кластеризация. Таким образом, особенностью алгоритма FCM является возможность произвольным образом адаптироваться к нечеткостям. Т.е. чем меньше уверенность в точности данных, снимаемых датчиками с исследуемого технологического процесса, тем большее значение задается для коэффициента нечеткости.

*Jm (U, v)* выполняет роль оценки взвешенной дисперсии точек из *Х* относительно оптимального расположения центров кластеров *vi, ..., vk* .

Значения ukj и vk , при которых формула минимальна (обозначается, как ), при m>1 удовлетворяют следующим условиям:

Условие 1:  Условие 2: 

Значение ukj ,обеспечивающее минимум выражения , можно найти с помощью следующей итеративной процедуры.

Шаг 1. Выбрать значение коэффициента нечеткости *m*, число кластеров *с* и определить соответствующим образом норму в выражении . Для матрицы степеней принадлежности векторов к кластерам *U* задать начальное значение *U(0)* ∈ *M*fc (*U(0)* целесообразно выбрать случайным образом независимо от *ukj*).

Шаг 2. Вычислить вектора центров кластеров {*vk*(0)}, используя *U(0*), по формуле .

Шаг 3. Определить новые значения элементов матрицы степеней принадлежности *U(1)*, используя {*vk*(0)}, по формуле (10).

Шаг 4. Задать подходящую норму и граничное значение ε для определения момента завершения алгоритма и выполнять предыдущие шаги до тех пор, пока .

Полученные таким образом элементы ukj матрицы U характеризуют степень принадлежности xj кластеру k.

В различных литературных источниках указывается, что процедура нечеткой кластеризации FCM, схема работы которой представлена на рис. 16, всегда сходится.

Число кластеров *c* выбирается приблизительно, на основе предвари­тель­ной оценки исследуемого ТП. Изменяя произвольным образом значение *с* как в меньшую сторону, так и в большую, следует выяснить, какое из них будет наилучшим образом представлять физические процессы, протекающие в моделируемом ТП.

Для оценки качества полученного разбиения был выбран критерий на основе энтропийного, экспоненциального распределения, так называемый коэффициент матрицы разбиения:



где - полученная матрица степеней принадлежности для выбранного коэффициента нечеткости и числа кластеров.

Значения критерия вычисляются для одного значения коэффициента размытости и различного числа кластеров и лежат в интервале [0, 1]. Чем меньше размытость данных, тем большие значения принимает коэффициент. Максимальное его значение соответствует оптимальному числу кластеров. Варьируя коэффициент нечеткости и число кластеров, можно построить поверхность в трехмерном пространстве для помощи в выборе оптимального числа кластеров.

Поскольку формулы ( ) и ( ) определяют минимум функционала качества *Jm*, то, изменяя начальное значение матрицы степеней принадлежности *U*, можно изменять и результаты кластеризации. Для получения хорошего результата следует более внимательно подходить к выбору начального значения. Кроме того, в любом случае трудно распознать небольшую изолированную область, т.е. вместо глобального минимума можно достичь локальный, что приведет к получению неадекватных результатов. Поэтому рекомендуется выполнить процедуру FCM несколько раз с различными начальными значениями матрицы разделения.

Результатом работы процедуры FCM в разработанном методе является алфавит классов, на основе которого формируется нечеткая кластерная модель ТП. Для этого предусмотрено задание лингвистического соответствия полученным эталонам. Оператор, владеющий знаниями об исследуемом ТП и навыками



Рис. 16. Процедура формирования нечеткой кластерной модели ТП

в данной области, должен сформулировать лингвистические термы, соответствующие полученным нечетким множествам.

Полученная кластерная модель дает возможность осуществлять построение итоговых моделей для различных состояний и режимов функционирования ТП и объединять их. Таким образом можно получать наборы моделей ТП.

**Задание:** реализовать алгоритм нечеткой классификации и, варьируя исходные данные, проанализировать полученные результаты алгоритма.