

ПРИМЕНЕНИЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ В ПРОЦЕССЕ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ФУНКЦИИ ПРИНАДЛЕЖНОСТИ К НЕЧЕТКОМУ МНОЖЕСТВУ

Козлов А.Ю., Адамов В.Г.

*Донецкий национальный технический университет
Кафедра автоматизированных систем управления
e-mail: Alexseyko@president.org.ua*

Рассмотрена задача выбора шаблонной формы функции принадлежности к нечетким множествам. Разработан способ применения генетических алгоритмов в поиске оптимальных параметров шаблонной функции принадлежности, применительно к задаче построения системы поддержки принятия решений в определении патологии плода.

Введение

При анализе УЗИ снимков в процессе определения патологий плода врач выполняет ряд замеров геометрических характеристик (диаметр желудочка, ширина воротникового пространства, длины конечностей и т.д.). Используя справочные данные врач может установить возможную патологию. Но часто одни и те же результаты измерений могут соответствовать различным диагнозам. Такая ситуация от части обусловлена тем, что справочные данные расходятся в значениях по некоторым замерам.

Задачу постановки диагноза можно решить при помощи компьютерных систем, спроектировав систему поддержки принятия решений (СППР). Которая, на основе: входных данных, базы знаний, справочной литературы и опыта, полученного в процессе работы, должна на выходе выдавать список предполагаемых диагнозов с вероятностями каждого из них.

При проектировании системы поддержки принятия решений возникает необходимость выбора средства построения ядра данной системы. Хорошие результаты при построении данной системы позволяет получить использование нечеткой логики.

Входные данные – численные значения проведенных замеров во время УЗИ диагностики.

Выходные данные – список предполагаемых диагнозов с вероятностью каждого из них. Эти данные служат основой врачу для принятия решения в постановке диагноза.

Технологии, основанные на нечеткой логики, наиболее подходят, потому как наверняка проверяются все правила из базы данных, а так же нет избыточности в поисках решений: сколько диагнозов, столько и правил. При построении СППР на нечеткой логике использование методики Мемдани которая дает возможность получить степень правдивости каждого правила, что в свою очередь можно трактовать как вероятность каждого из правил [1].

При построении системы поддержки принятия решений, на основе нечеткой логики, одной из основных задач, которые необходимо решить – является выбор наиболее правильной формы и расположения функции принадлежности на числовой оси. Существует два варианта определения данной функции: при помощи эксперта, и без него. Эксперт, в свою очередь, определяет где и как должны располагаться функции принадлежности к каждому множеству, и для каждой переменной. В случаях, когда такого эксперта нет, можно использовать методы, позволяющие получить наиболее

точную форму и местоположение функций. В данной работе в качестве такого метода используются генетические алгоритмы.

Постановка задачи

Нужно выбрать шаблонную форму для описания функции принадлежности к нечетким множествам у входных переменных. После этого для выбранной шаблонной формы определить способ описания через набор параметров. К данному набору описывающих параметров разработать алгоритм, позволяющий достичь наиболее оптимальной формы и расположения данного множества, используя заданные критерии соответствия входных и выходных данных.

Входные данные для алгоритма:

1. Выбранная форма искомого множества;
2. Стартовые значения параметров;
3. Последовательность входных данных полученных со снимков УЗИ с точно установленным диагнозом или из справочной литературы.

Выходные данные для данного алгоритма:

1. Наиболее правильная форма данного множества;

Выбор шаблонной формы функции принадлежности

В качестве функции принадлежности могут использоваться различные варианты уравнений: треугольной, трапециевидной, гауссовой и др. форм (рис 1).

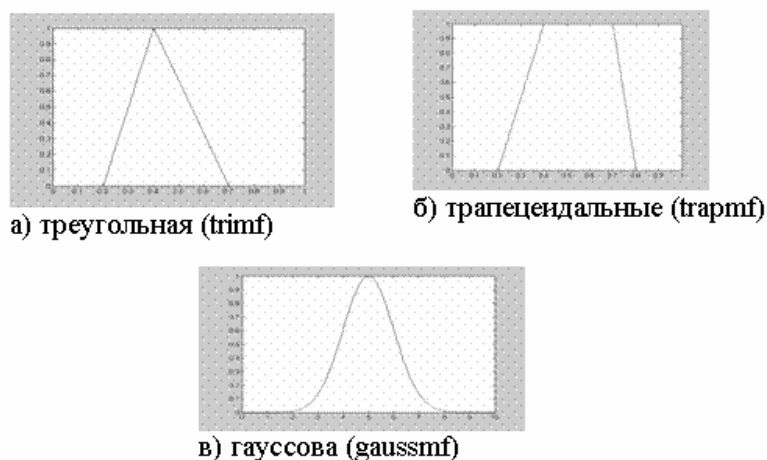


Рисунок 1 – возможные формы функций принадлежности

Исходя из универсальности и удобства в обращении наиболее подходящей формой является трапециевидная.

Выбор основных параметров шаблона

У выбранного шаблона можно выделить четыре основных точки, которые обозначены на рисунке (рис. 2): A, B, C, D. Проанализировав их можно выделить четыре параметра выбранного шаблона: т. A – точка, с которой начинается множество, и три отрезка: AB, BC, CD.

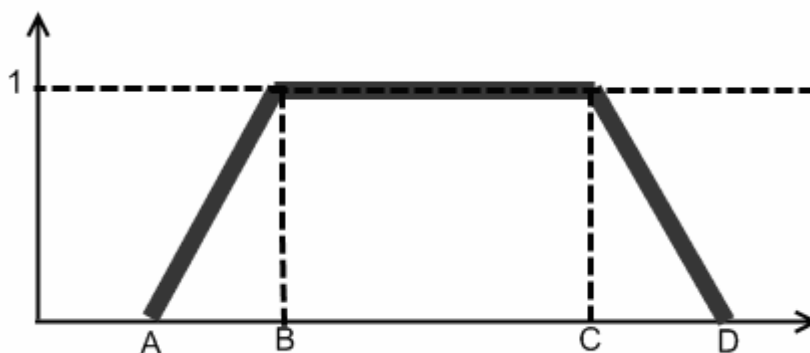


Рисунок 2 – Стандартная выбранная форма терма

При приведении длины отрезка BC к нулю возможно получить функцию принадлежности треугольной формы. Это в свою очередь существенно расширяет спектр применения такого шаблона.

В зависимости от длины отрезков AB и CD можно легко определить крутизну граней трапеции, причем данные значения могут быть не одинаковы, что дает возможность сделать форму функции принадлежности не симметричной.

В итоге использование шаблона в форме трапеции дает возможность, изменяя значения точки начала множества и отрезков AB, BC и CD получить оптимальную форму функции принадлежности, которая позволит наиболее точно формировать результирующее значение.

Разработка алгоритма определения параметров

Генетические алгоритмы применяются при разработке программного обеспечения в системах искусственного интеллекта, оптимизации, искусственных нейронных сетях и в других отраслях знаний. Следует отметить, что с их помощью решаются задачи, для которых ранее использовались только нейронные сети. Генетические алгоритмы также применяются совместно с нечеткими системами.

Генетический алгоритм представляет собой метод, отражающий естественную эволюцию методов решения проблем, и в первую очередь задач оптимизации. Генетические алгоритмы - это процедуры поиска, основанные на механизмах естественного отбора и наследования. В них используется эволюционный принцип выживания наиболее приспособленных особей [2].

Применительно к решаемой задаче особью является набор из 4 чисел: точки на числовой оси, с которой начинается восходящая часть трапеции (на рисунке это точка A), и трех значений длин отрезков соответственно AB, BC и CD.

Перед первым шагом нужно случайным образом создать начальную популяцию. Данная популяция создается на основе уже имеющейся функции принадлежности. Особи популяции создаются благодаря варьированию в пределах $\pm 10\%$ значения от каждого из четырех базовых параметров. Итогом первого шага является популяция N, состоящая из N особей.

Размножение в генетических алгоритмах обычно половое. Чтобы произвести потомка, нужны несколько родителей. Для решения данной задачи выбираем два родителя. После чего случайным образом выбранные 2 числа из одного родителя и недостающие 2 числа из другого родителя образуют новую особь.

Для создания возможности выхода из локального максимума используется мутация. Проходит она так: есть некоторая доля мутантов m , являющаяся параметром генетического алгоритма, и на шаге мутаций нужно выбрать $m(N)$ особей, а затем изменить их в соответствии с заранее определенными операциями мутации. В данном случае это увеличение или уменьшение значения одного из чисел особи не более чем на 20%.

На этапе отбора нужно из всей популяции выбрать определенную ее долю, которая останется «в живых» на этом этапе эволюции. Существуют разные способы проводить отбор. Вероятность выживания особи h должна зависеть от значения функции приспособленности $Fitness(h)$. Сама доля выживших s обычно является параметром генетического алгоритма, и ее просто задают заранее. По итогам отбора из N особей популяции N' должны остаться $s(N)$ особей, которые войдут в итоговую популяцию N' . Остальные особи погибают.

Очень важным понятием в генетических алгоритмах считается функция приспособленности (fitness function), иначе называемая функцией оценки. Она представляет меру приспособленности данной особи в популяции. Эта функция играет важнейшую роль, поскольку позволяет оценить степень приспособленности конкретных особей в популяции, и выбрать из них наиболее приспособленные (т.е. имеющие наибольшие значения функции приспособленности) в соответствии с эволюционным принципом выживания «сильнейших» (лучше всего приспособившихся). В данном случае в качестве значения фитнес функции выступает количество точек из обучающего набора. То есть набора состоящего из множества значений входных переменных и множества соответствующих значений выходных переменных.

Данный алгоритм рассчитан на то, что с течением времени будет накапливаться опыт в виде дополнения обучающего набора, что в свою очередь будет требовать периодического пересчета параметров функций принадлежности. На этом принципе основаны значения погрешности при формировании начальной популяции.

Разработанный генетический алгоритм состоит из следующих шагов:

1. Выбор исходной популяции хромосом путем внесения погрешности;
2. Оценка приспособленности хромосом в популяции;
3. Проверка условия остановки алгоритма;
4. Селекция хромосом;
5. Применение генетических операторов;
6. Формирование новой популяции;
7. Выбор «наилучшей» хромосомы.

Блок-схема генетического алгоритма изображена на рис. 3.

В итоге получаем наиболее подходящее сочетание четырех параметров для данной переменной. Которые используем для описания функции принадлежности в системе принятия решений.

Данный алгоритм реализован в рамках системы поддержки принятия решений в акушерстве и гинекологии. Программное обеспечение реализовано на базе языка программирования высокого уровня C#.

Проверка функционирования алгоритма на тестовых данных показала вполне удовлетворительные результаты.

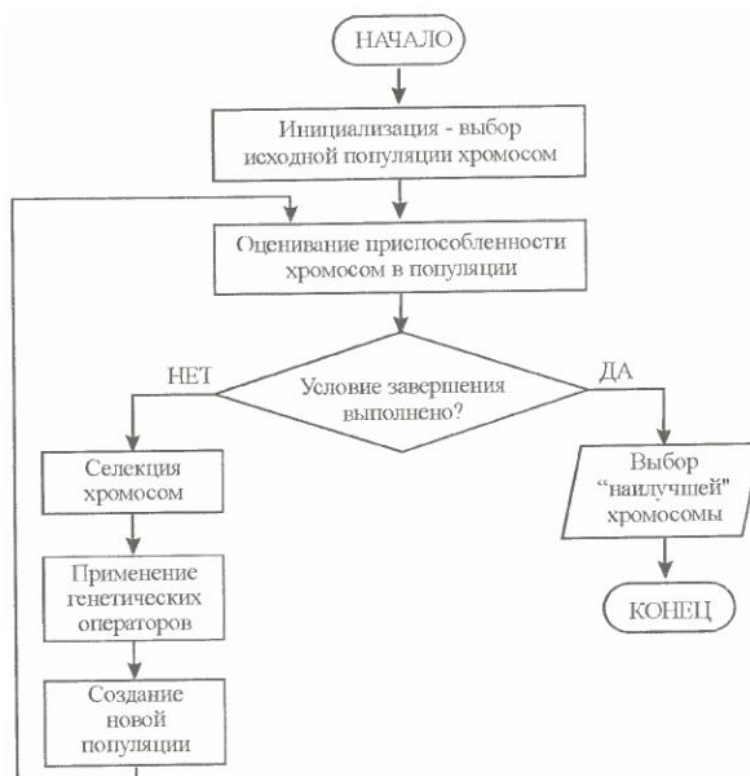


Рисунок 3 – Блок-схема генетического алгоритма

Вывод

В результате проведенных исследований в области определения нечетких множеств удалось получить алгоритм, который можно использовать в системе поддержки принятия решений для периодической оптимизации функции принадлежности к различным переменным в системе с нечеткими множествами.

По мере работы системы поддержки принятия решений планируется накопление эмпирической информации по каждой из возможных переменных. Что в свою очередь позволяет, периодически запуская генетический алгоритм поиска оптимального значения функции принадлежности, получать все более и более оптимальное описание нечетких множеств. Что в свою очередь способно сделать системы намного более точной.

Литература

- [1] В. Дьяконов, В. Круглов. Математические пакеты расширения MATLAB. Специальный справочник. -Санкт-Петербург: Питер, 2001 - с. 307-309
- [2] Д. Руткоская, М. Пилинский, Л. Рутковский Нейронные сети, генетические алгоритмы и нечеткие системы – М.: «Горячая линия-Телеком», 2006, 303с.