УДК 621.372

А.Н. Афанасьев, Н.Н. Войт

РАЗРАБОТКА МЕТОДОВ НЕЧЕТКОЙ ПАРАМЕТРИЧЕСКОЙ АДАПТИВНОЙ ДИАГНОСТИКИ ОБУЧАЕМОГО ИНЖЕНЕРА

***Афанасьев Александр Николаевич**, кандидат технических наук, окончил факультет информационных систем и технологий Ульяновского государственного технического университета. Доцент кафедры вычислительной техники УлГТУ. Имеет монографии, статьи, учебные пособия в области создания САПР. [Тел.:(8422) 77-81-13].*

***Войт Николай Николаевич**, кандидат технических наук, окончил факультет информационных систем и технологий Ульяновского государственного технического университета. Сотрудник кафедры вычислительной техники УлГТУ. Имеет публикации в области создания автоматизированных средств обучения САПР. [Тел.: (8422) 77-81-13].*

Аннотация

Описан метод диагностики проектных характеристик (знаний, умений, навыков и компетентности) обучаемого инженера с использованием нечетких карт Кохонена. Приведен алгоритм обучения, использующий гауссову функцию принадлежности.

Ключевые слова: автоматизация проектирования, обучение инженера, нечеткие карты Кохонена, адаптивная диагностика.

Abstract

The article describes a method of diagnostics of design characteristics (knowledge, ability, skills and competence) of engineering specialist trained using CAD. It also cites a training algorithm using the Gaussian membership function.

Key words: design automatization, engineer training, Kohonen's fuzzy maps, adaptive diagnostics.

1 Метод диагностики на основе самоорганизующейся карты Кохонена

Т. Кохоненом (1982) разработана сеть, использующая алгоритм обучения без учителя, - «самоорганизующая карта признаков» (SOM), часто используемое название сети Кохонена — KCN. Применяется для отображения нелинейных взаимосвязей данных на достаточно легко интерпретируемые сетки (карты), представляющие метрические и топологические зависимости входных векторов, объединяемых в классы [1].

1.1 Способы классификации обучаемых инженеров

Разбиение на классы предусматривает наличие общего свойства обучаемых инженеров, которое представляется параметрами в векторной форме. Отнесение характеристик модели обучаемого инженера к определенному классу выполняется по правилу или критерию разбиения, представленному мерой близости (сходства). Мера вычисляется с помощью математических методов. Удобно представлять модель обучаемого инженера вектором параметров (свойств):

$X=(x_1, x_2, x_3, ..., x_n) \subset R^n$,

тогда мерой близости двух объектов X_1 и X_2 можно принять Евклидово расстояние между ними:

$$D(X_1,X_2)=\|X_1-X_2\|=\sqrt{\sum_{i=1}^{n}(\Delta X_i)^2}. \qquad (1)$$

Если нужно показать наихудшую ситуацию классификации, тогда в качестве меры сходства между объектами используют максимальную разность значений параметров векторных описаний:

$$D=max_{i=1}^{n}\left(x_{1i}-x_{2i}\right). \qquad (2)$$

Картам Кохонена, ориентированным на классификацию, свойственен принцип самоорганизации (самообучения). При обучении правильные ответы картам не сообщаются, во время процедуры обучения происходит формирование классов на основе обобщений имеющейся входной информации, сосредоточенной в обучающей выборке, руководствуясь критерием сходства.

1.2 Конкурентное обучение карт Кохонена (SOM)

Наиболее популярным способом самоорганизации является процедура конкурентной (соревновательной) кластеризации (разбиение на классы). Кластером называется группа объектов, обобщенных по определенным свойствам.

Число входов карты соответствует размерности входных векторов, а число выходов задает число классов, на которое должно быть разделено входное множество объектов. Каждый класс определяется весовым вектором, соединенным с соответствующим выходом.

Входной вектор X и весовой вектор W_j предполагаются нормализованными, нормализация осуществляется следующим образом:

$$\left|X\right| = \sqrt{\sum_i x_i^2} = 1; \quad W_j = \sqrt{\sum_i w_i^2} = 1. \qquad (3)$$

Величина нейронной активности a_j выхода j определяется как Евклидово расстояние либо по формуле:

$$a_j = \sum_{i=1}^{3} x_i \cdot w_{ij}. \qquad (4)$$

Между выходами карты осуществляется конкуренция на основе (4). Нейрон–победитель определяется по наибольшему значению активности. Наибольшая нейронная активность соответствует выходу j, тогда весовые коэффициенты связей выхода j с выходами карты будут изменяться в соответствии с правилами конкуренции по следующей формуле:

$$W_j\left(t+1\right) = \frac{W_j\left(t\right) + \mu\left(X_j\left(t\right) - W_j\left(t\right)\right)}{\left\|W_j\left(t\right) + \mu\left(X_j\left(t\right) - zW_j\left(t\right)\right)\right\|}, \qquad (5)$$

где $0 \le \mu \le 1$ - скорость обучения.

В формуле (5) используются нормализованные векторы, и при предъявлении очередного входного вектора обновляется только вес нейрона-победителя, а остальные веса не изменяются. В качестве нейрона-победителя выбирается такой нейрон, весовой коэффициент которого наиболее близок к входному вектору, весовые векторы выходных нейронов на каждой итерации обучения будут изменяться в направлении кластеров входных образов.

В случае использования Евклидова расстояния в качестве меры непохожести векторов выбирается активность выходных нейронов, которая вычисляется по формуле:

$$a_j = \sqrt{\sum_{i=1}^{n}\left(x_i - w_{ij}\right)^2} = \left\|X - W_j\right\|. \qquad (6)$$

В формуле (6) определяется нейрон-победитель с номером j с минимальным Евклидовым расстоянием между входным образом и весовым вектором этого нейрона. Среднеквадратическая ошибка для j-го нейрона-победителя равна

$$E = \sum_p \left\|W_{fX_p} - X_p\right\|, \qquad (7)$$

где $f(X_p)$ - нейрон-победитель;

X_p — входной вектор;

W_{fXp} - центр класса, которому принадлежит X_p.

1.3 Архитектура карт Кохонена

Рассмотрим более подробно архитектуру карты Кохонена и правила ее обучения. Карта Кохонена имеет один слой нейронов, число входов каждого нейрона равно размерности входного вектора, число нейронов в слое равно числу различных классов, которые карта может распознать. Карта Кохонена может не только выделить основные группы образов, но и установить некоторые особенности полученных классов, при этом близким входным образам будут соответствовать близкие карты нейронной активности.

Основная цель обучения карт Кохонена состоит в выявлении структуры в n-мерных входных данных и представлении ее на карте в виде распределенных нейронных активностей. Каждый нейрон несет информацию о классе, объединяющем в группу сходные по критерию близости входные векторы, формируя для данной группы некоторый собирательный образ.

SOM способна к обобщению, активизируя нейроны подобных векторов. Конкретному кластеру может соответствовать и несколько нейронов с близким значением весовых векторов, поэтому выход из строя одного нейрона не так критичен к ошибке распознавания для функционирования этой сети, как это имеет место, например, в случае сети Хемминга. Карта Кохонена позволяет понизить размерность пространства представления многомерных входных данных, сохранив топологию связей между ними.

В большинстве случаев каждый выходной элемент карты соединен со своими соседями. Корректировка весов происходит не для всех весов сети, а только для окрестности элемента, который наилучшим образом откликается на очередной вход.

Карта Кохонена использует конкурентный алгоритм обучения. Выигрывает тот нейрон, чей весовой вектор наиболее близок (меньше расстояние) к текущему входному вектору, определяемому, например, Евклидовой метрикой (1).

В другом варианте алгоритма обучения победителем считается элемент, весовой вектор которого имеет наибольшее скалярное произведение с входным вектором. Эта величина также является некоторой мерой близости, поскольку скалярное произведение — это проекция входного вектора на весовой вектор. Очевидно, что такая проекция будет наибольшей, если векторы имеют близкие направления. При этом методе, однако, оба вектора — весовой и входной — должны быть нормированы по длине, например, быть равными единице. Напротив, Евклидово расстояние позволяет работать с векторами произвольной длины.

На текущем шаге обучения вес меняется только у нейрона-победителя m^*, а также, возможно, и у его непосредственных соседей; веса остальных нейронов при этом «замораживаются». Нейронная активность соседнего нейрона s уменьшается при увеличении расстояния между нейроном s и выигравшим нейроном m^*.

Подстройка входного вектора выигравшего нейрона и его соседей осуществляется путем их перемещения в сторону входного вектора. После обучения на достаточном количестве обучающих примеров совокупность весовых векторов приходит в соответствие со структурой входных векторов, т.е. весовые векторы в буквальном смысле моделируют распределение входных образов.

Для правильного понимания картой входного распределения нужно, чтобы каждый элемент становился победителем одинаковое число раз, следовательно, появление различных весовых векторов должно быть равновероятно. С этой целью в процедуру обучения вводится некоторый механизм «справедливости». Один из возможных способов осуществления этого механизма связан с увеличением расстояния между входным и весовым векторами на некоторую величину добавки часто выигрывающих нейронов и уменьшения этого расстояния на ту же величину для тех нейронов, которые чаще проигрывают. Шансы проигрывающих нейронов повышаются. Величина добавки меняется в процессе обучения в соответствии с изменением частоты выигрывающих нейронов.

1.4 Нечеткие карты Кохонена (FSOM)

В нечеткой карте Кохонена [1,2] нечеткость вводится на уровне выделения кластеров. Нечеткая карта имеет два слоя нейронов, первый слой представлен слоем расстояний стандартной SOM, а второй слой — слоем меры принадлежности. Для заданного входного вектора X_j узел i, $i=1, 2,\dots, k$ в слое расстояний вычисляет расстояние $dist(X_j, W_j)$ точно так же, как и для SOM по формуле (1). В FSOM введен второй слой, мера принадлежности входного образа классу представлена функцией принадлежности μ_{ij}, $i = 1, 2, \dots, k$; j - номер примера в обучающей выборке ($j = 1, 2, \dots, N$).

Суммарная степень принадлежности входного вектора ко всем классам равна 1:

$$\sum_{i=1}^{k}\mu_{ij} = 1 \quad \text{для } j = 1, 2,\dots,N. \qquad (8)$$

Второй слой нейронов - FSOM-слой функций принадлежности отображает расстояние $dist(X_j, W_j)$, где $j = 1, 2,\dots, N$ (N — количество примеров в обучающей выборке), $i = 1, 2,\dots, k$ (k — количество классов), в значениях функции принадлежности вектора некоторому классу в соответствии с уравнением:

$$\mu_{ij} = \left[\sum_{p=1}^{k}\frac{dist_{ij}}{dist_{pj}}\right]^{-1}, \qquad (9)$$

где $\mu_{ij} = 1$, если $dist_{ij} = 0$ и $\mu_{ij} = 0$, если $dist_{pj} = 0$, $p \neq i$.

Обратная связь в схеме позволяет производить обновление весовых коэффициентов на каждом шаге итерации.

Запишем алгоритм обучения карты FSOM.

1. Задать начальные значения весов W_i, $i=1, 2,\dots, k$ (случайным образом или на основе определенного алгоритма) и установить величину окрестности $r = k/2$, где k — априорно заданное число кластеров.

2. Для каждого входного вектора X_j выбирается узел i^* в слое расстояний, такой, что $1 \le i^* \le k$, где $dist_{ij}$ определяется как Евклидово расстояние. Обновить веса W_i по следующему правилу:

$$W_i(t + 1) = W_i(t) + \mu_{ij}(t)dist(X_j, W_i(t)), \qquad (10)$$

где t — номер итерации;

i — входной узел в слое расстояний и всех r соседей вокруг него.

3. Проверить равенство $r = 0$. Если $r = 0$, алгоритм завершен, в противном случае - $r: = r-1$ и переход к шагу 2.

4. Остановка.

При выполнении шага 2 (при обновлении весов) так же, как и в алгоритме обучения SOM, выигрыш во времени при вычислениях достигается за счет обновления только весов соседей. Количество обновлений весов меньше, чем в SOM, за счет введения нечеткости — функции принадлежности входных векторов кластерам. Завершение процедуры обучения автоматизировано, когда r становится равным нулю.

Большое влияние на скорость сходимости алгоритма обучения FSOM оказывает закон изменения скорости обучения и окрестности соседних нейронов.

1.5 Анализ видов функций принадлежности, применяемых в классификации обучаемых инженеров

Для формирования множеств входных показателей (в виде нечетких множеств) и построения их логико-лингвистических шкал целесообразно использовать типовые $L-R$ - функции, например, колокообразного вида.

Построение μ_{ci} может проходить по следующим схемам:

1. По принципу симметричной треугольной функции:

$$\mu_{ci}(x)=\begin{cases}1-\dfrac{|x-c|}{d} & \text{для } x\in[c-d,c+d];\\ 0 & \text{для остальных.}\end{cases} \qquad (11)$$

2. При помощи обобщенной гауссовой функции:

$$\mu_{ci}(x) = exp\left[-\left(\frac{x-c}{\sigma}\right)^{2b}\right]. \quad (12)$$

3. При помощи трапециевидных функций:

$$\mu_{ci}(x) = \begin{cases} 1 - \frac{|x-c|}{d} & \text{для } x \in [c_1 - d, c_2 + d], \\ & \text{причем } \; c_1 \neq c_2\,; \\ 0 & \text{для остальных,} \end{cases} \quad (13)$$

где c — центр кластера;

d — ширина треугольной функции;

σ, b — параметры формы для гауссовой кривой.

Производная гауссовой функции равна экспоненте, скорость изменения функции имеет экспоненциальный характер. Функция медленно меняется в окрестностях ее максимума и минимума, т.е. выделяются две окрестности функции (5) в глобальных экстремумах, которые явно делят входное множество на две части, что улучшает качество вычисления принадлежности.

Производная треугольной функции равна константе, т.е. скорость изменения функции постоянная, функция одинаково меняется на всех участках, следовательно, не возникает окрестностей, в которых бы наблюдалось сосредоточение значений функции.

Следовательно, наиболее эффективной функцией для классификации является колокообразная гауссова функция (12).

Способы оценок погрешности карт Кохонена.

1. Формула оценки максимума правдоподобия Кульбака-Лейблера [1] имеет следующий вид:

$$E = \sum_{i=1}^{M} x_{целевое_i} \log \frac{x_{целевое_i}}{x_{фактическое_i}} + \left(1 - x_{целевое_i}\right) \log \frac{1 - x_{целевое_i}}{1 - x_{фактическое_i}}, \quad (14)$$

где M — число выходов сети;

$x_{фактическое}$ — оценочная фактическая величина;

$x_{целевое}$ — эталонное значение величины.

2. Метод интервальных оценок показывает наихудший вариант оценки. Формальная запись следующая:

$$\Delta x = x_{целевое} - x_{фактическое}. \quad (15)$$

3. Метод среднеквадратического отклонения показывает лучший вариант оценки, формальная запись имеет вид:

$$\Delta x = \sqrt{\left(x_{целевое} - x_{фактическое}\right)^2}. \quad (16)$$

2 Разработка нечеткой классификации обучаемых инженеров на основе FSOM

2.1 Структурно-параметрический анализ проектного решения

Входные оценочные значения баллов за проектные задачи, выполненные инженером, получаются в результате структурно-параметрической оценки учебного объекта проектирования. Анализ проектного решения на структурном уровне в обучающих целях использует дерево построения, что обеспечивает контроль процесса и последовательности проектирования, структуры решения и повышает качество обучения. Сравнительный анализ разработанного проектного решения с эталоном позволяет выявить наличие лишних элементов или отсутствие элементов, неправильно совершенные операции над объектом. Поэтапное сравнение иерархического дерева построения с эталонным вариантом может проводиться в глубину и в ширину дерева.

Анализ в ширину дерева позволяет контролировать все одноуровневые элементы проектного решения. Такой анализ используется для вычисления количества элементов, входящих в проектное решение.

Метод сравнения основан на обходе дерева по горизонтали путем сравнения элементов дерева проектного решения с эталонным вариантом (анализ «слева-направо»). Одинаковые элементы в проектном решении и в эталонном отмечают совпадение. В проектном решении элементы могут отсутствовать или отличаться от эталонных элементов.

Анализ в глубину дерева позволяет детально контролировать любой элемент проектного решения на всех иерархических уровнях. Анализ в глубину сравнивает подэлементы элемента дерева проектного решения (анализ «сверху-вниз»).

Параметрический анализ учитывает параметры элементов объекта и совершенные операции, сравнивает их в проектном решении и в эталонном варианте. Параметрический анализ применяется вместе со структурным при обходе дерева проектного решения.

Метод анализа может быть интегрирован в САПР с помощью разработки дополнительных сервисных программ, что дает возможность использовать в обучении примеры, взятые с производства, и контролировать правильность проектных решений в САПР.

2.2 Классификация проектных характеристик обучаемого инженера

Параметры модели обучаемого инженера меняются событийно в контрольных точках K_i сценария. Оценочные входные векторы поступают на вход нечеткой карты, которая классифицирует полученные данные и формирует нечеткие характеристики уровня подготовленности инженера САПР. Число входов нейрона класса знаний равно числу элементов входного вектора баллов за ответы на вопросы. Число входов нейрона класса умений равно числу элементов входного вектора баллов за решенные проектные задачи. Число входов нейрона класса навыков равно числу элементов входного вектора затраченного времени на выполнение проектных действий. Число входов нейрона класса компетентности равно трем (знания, умения, навыки).

Классификация выполняется с помощью нечеткой карты [3] и разработанной оценочной шкалы в процентном, интервальном и лингвистическом видах. Функция принадлежности имеет вид:

$$\mu_{ij} = exp\left[-\frac{1}{2}\left(dist_{ij}\right)^2\right],\ \ i=\overline{1,C},\ \ j=\overline{1,4}, \tag{17}$$

где C — число классов;

j имеет значение от 1 до 4, потому что есть четыре класса: знания, умения, навыки, компетентность.

Активность нейронов слоя расстояния вычисляется как Евклидово расстояние (отклонение) от требуемой активности:

$$dist_{ij} = \sqrt{\sum_{e=1}^{T}\left(x_e - w_{ije}\right)^2}, \tag{18}$$

где T — число критериальных параметров входного оценочного вектора X;

x_e — элемент вектора X;

w_{ije} — вес дуги, связывающей элемент x_e и нейрон ij класса.

Минимальное значение активности нейрона класса j слоя расстояния имеет вид $min_j = min_j\ (dist_{ij})$, максимальная принадлежность входного вектора к классу j имеет вид $max_j = max_j\ (\mu_{ij})$.

При многократном применении нечетких карт обратная связь позволяет обновлять весовые коэффициенты на каждом шаге t:

$$W_{ij}(t+1) = W_{ij}(t) + \mu_{ij}(t)\ dist_{ij}(X, W_{ij}\ (t)), \tag{19}$$

что позволяет повысить точность оценки.

Четыре выходных кортежа нечеткой карты имеют вид:

$$<dist_j\ ,\ \mu_j>,\quad j=\overline{1,\ 4},$$

где $dist_j$ — значение активности нейрона j класса;

μ_j — значение функции принадлежности входного вектора к конкретному j классу.

Проведена оценка эффективности использования FSOM по отношению к конкурентным методам, основанным на скрытых Марковских переменных (СМП). Погрешность классификации проектных характеристик обучаемого инженера с помощью FSOM равна 15%, а с помощью методов, использующих СМП, — 21%.

2.3 Алгоритм обучения FSOM, использующий гауссову функцию принадлежности

Заранее обычно неизвестно, насколько полезны (информативны) те или иные входные параметры для предсказания значений выходов, у разработчиков возникает соблазн увеличить число входных переменных в надежде на то, что сеть сама определит, какие у них наиболее значимые входные параметры. Но сложность обучения нечетких карт при этом быстро возрастает, а также, что достаточно важно, с увеличением числа входов ухудшается точность предсказаний, и, как следствие, снижается предсказательная способность карты, она просто заучивает данные вместе с шумом (бесполезной информацией) [4].

Данную проблему можно решить, пользуясь большой выборкой [5] при обучении либо жестко лимитируя число входов, и тогда набор наиболее информативных переменных представляется важным этапом подготовки данных для обучения нечетких карт.

Алгоритм обучения FSOM имеет вид.

1. Задать равновероятные начальные значения весов W_{ij}, $i=\overline{1,C};\ \ j=\overline{1,\ 4}$.

2. Задать погрешность $e_{ij}^{standart}$.

3. Начало цикла:

а) Вычислить функцию принадлежности $\mu_{ij} = exp\left[-\frac{1}{2}\left(dist_{ij}\right)^2\right]$.

б) Обновить весовые векторы $W_{ij}(t+1) = W_{ij}(t) + \mu_{ij}(t)\ dist_{ij}(X, W_{ij}\ (t))$.

в) Вычислить $e_{ij} = dist_{ij}(W_{ij}(t+1), W_{ij}(t))$.

г) $W_{ij}(t) = W_{ij}(t+1)$.

д) Пока $e_{ij} > e_{ij}^{standart}$ выполняется цикл.

4. Остановка.

Выводы

Разработан метод нечеткой параметрической адаптивной диагностики обучаемого с помощью критериальных параметров модели обучаемого, таких, как знание, умение, навык и компетентность, на основе карт Кохонена. Входной вектор карт имеет параметры: балл, время, сложность, число правильных решений, число повторов учебного материала. Первый слой содержит нейроны, которые вычисляют активность как сумму Евклидовых расстояний (не требуется нормализация). Второй слой имеет нейроны, которые вычисляют значение принадлежности активности к классу. Обратная связь карт позволяет корректировать и самообучать карту автоматически. Алгоритмом обучения сети является «Обратное распространение ошибки» с функцией принадлежности Гаусса. При многократном использовании карт достигается повышение точности метода классификации. Выходным вектором карт является пара значений активности и принадлежности.

СПИСОК ЛИТЕРАТУРЫ

1. Комарцева Л.Г., Максимов А.В. Нейрокомпьютеры. — М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. — 400 с.

2. Осовский С. Нейронные сети. — М.: Финансы и статистика, 2002. — 344 с.

3. Афанасьев А.Н., Войт Н.Н. Разработка автоматизированной обучающей системы САПР с использованием нечетких нейронных сетей // Труды международной конференции по системам искусственного интеллекта AIS/CAD' 08. — Краснодар, 2008. — С. 43 - 46.

4. Безруков Н.С., Еремин Е.Л. Выделение информативных признаков для системы поддержки принятия решения на основе нейро-нечеткой сети // Нейрокомпьютеры. — 2007. - № 1-2. — С. 22 — 25.

5. Горбань А.Н., Дунин-Барковский В.Л., Кирдин А.Н. Нейроинформатика. — Новосибирск: Наука, 1998. — 296 с.