

УДК 519.816

СИНТЕЗ ОПТИМАЛЬНЫХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ С ПОМОЩЬЮ МОДИФИЦИРОВАННОГО ГЕНЕТИЧЕСКОГО АЛГОРИТМА

И.Б. Бондаренко, Ю.А. Гатчин, В.Н. Гераничев

Рассмотрены принципы синтеза нейронных сетей, показаны недостатки подходов при выборе структуры и настройке весов. Оптимизация сетей возможна с применением эволюционных алгоритмов, обладающих преимуществами, которые проявляются при работе с многосвязными и многослойными сетями. Проведенные имитационные эксперименты на конкретной тестовой функции подтверждают эффективность разработанных алгоритмов.

Ключевые слова: искусственная нейронная сеть, генетический алгоритм, кластерный анализ, нейрон, хромосома, обратное распространение ошибки, тестовая функция Бранинса.

Введение

Интеллектуальные системы на основе искусственных нейронных сетей (ИНС) позволяют с успехом решать проблемы распознавания образов, классификации, аппроксимации функций, выполнения прогнозов, оптимизации и управления. Создание ИНС сводится к выполнению следующих этапов:

- синтез структуры ИНС;
- настройка весов всех нейронов ИНС (обучение).

В настоящее время для реализации перечисленных задач не существует строгих методов решения, даются лишь рекомендации. Предлагаемые алгоритмы направлены на решение локальных задач, из-за чего структура ИНС оказывается неудовлетворительной, а время обучения – большим. В этом случае приходится создавать сеть и производить расчеты заново. Еще меньшее внимание уделяется исследователями построению многослойных несимметричных ИНС, характеризующихся сложностью и многовариантностью.

Для создания структуры и настройки весов сети (стратегия обучения «с учителем») авторами [1] предлагаются подходы, основанные на использовании эволюционных алгоритмов. Однако предлагаемые процедуры позволяют получить лишь один вариант структуры ИНС. В настоящей работе авторами предлагается разработанный модифицированный генетический алгоритм (МГА) совместно с кластерным анализом всех возможных вариантов, с целью получения полного множества оптимальных вариантов построения ИНС. Такой подход позволит расширить множество анализируемых альтернатив в системах принятия решений, основанных на применении ИНС: интеллектуальных автоматизированных систем проектирования, экспертных систем, систем управления и т.п.

Постановка задачи

Основу ИНС составляют нейроны, имеющие строение, сходное с биологическими аналогами. Каждый нейрон можно представить как микропроцессор с несколькими входами и одним выходом. Синаптические веса W_i оказывают влияние на выходной сигнал нейрона NET [2, 3]:

$$NET = \sum_i^k W_i X_i,$$

где k – количество входов нейрона; X_i – i -й вход нейронной сети. С помощью блока нелинейного преобразователя окончательно формируется выходной сигнал OUT_i . В нелинейном преобразователе в качестве преобразующей (сжимающей) использована сигмоидальная функция. При объединении нейронов между собой образуется структура – нейронная сеть. Выстроенные вертикально нейроны образуют слои: входной, внутренний/внутренние и выходной. Количество слоев определяет сложность и, в то же время, функциональность сети, которая до конца не исследована.

Для исследователей выполнение первого этапа создания сети является наиболее сложной задачей. В литературе даются следующие рекомендации [4].

1. Число нейронов скрытого слоя определяется эмпирическим путем, но в большинстве случаев используется правило

$$N_{\text{скр}} \leq N_{\text{вх}} + N_{\text{вых}},$$

где $N_{\text{скр}}$, $N_{\text{вх}}$, $N_{\text{вых}}$ – число нейронов соответственно в скрытом, входном и выходном слое.

2. Увеличение количества входов и выходов сети ведет к необходимости увеличения числа нейронов в скрытом слое.
3. Для ИНС, моделирующей многоэтапные процессы, необходим дополнительный скрытый слой, но, с другой стороны, добавление скрытых слоев может привести к перезапоминанию и к неверному решению на выходе сети.

С учетом приведенных рекомендаций количество слоев и число нейронов в скрытых слоях выбирается исследователем, исходя из его личного опыта.

Для выполнения этапа обучения ИНС разработано множество алгоритмов. Наибольшее распространение получил метод обратного распространения ошибки, который позволяет подстраивать веса в

многослойных сложных ИНС с помощью обучающих наборов. По рекомендации Е. Баума и Д. Хасслера [2] объем обучающего множества прямо пропорционален количеству всех весов ИНС и обратно пропорционален доле ошибочных решений при работе обученной сети.

Недостатки алгоритма обратного распространения, с которыми сталкиваются исследователи, несмотря на все разработанные модификации, заключаются в следующем. Процесс обучения может длиться достаточно долго, а может никогда не закончиться. Чрезмерное увеличение весов ИНС может привести к «параличу» сети, когда синаптические веса перестанут изменяться, не достигнув конца обучения. Величину шага необходимо постоянно подстраивать, сеть необходимо постоянно «доучивать», так как она постепенно «забывает» предыдущие обучающие наборы, которым она была обучена. Все эти недостатки приводят к тому, что разработка ИНС под конкретную задачу сильно усложняется. Наше исследование в рамках данной работы посвящено выполнению только второго этапа.

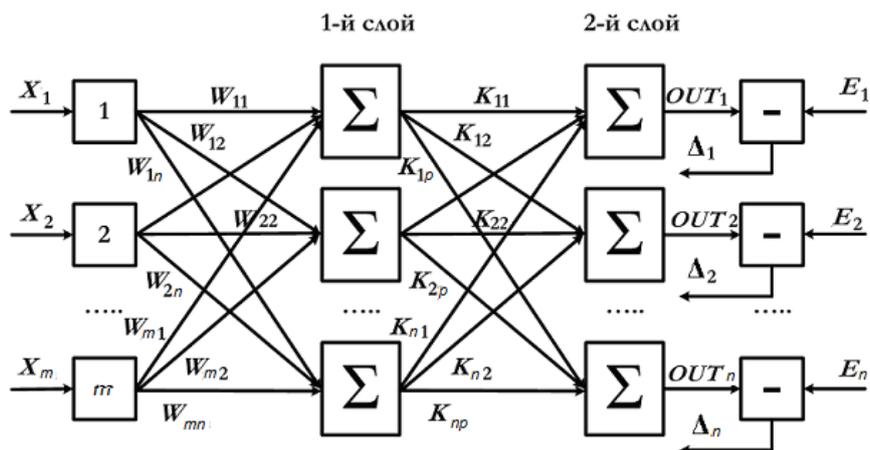


Рис. 1. Принцип обучения ИНС методом обратного распространения ошибки

Математически задача подбора весов состоит в оптимизации функции ошибки R или показателя качества обучения (рис. 1)

$$R_{opt} = \min(R), R = \frac{1}{2N_s} \sum_{i=1}^{N_s} (OUT_i - E_i)^2, \quad (1)$$

где N_s – мощность множества обучающих пар; OUT_i – полученное с помощью сети значение выходного нейрона при i -м наборе обучения; E_i – требуемое значение выходного нейрона при i -м наборе обучения.

Каждая i -я ошибка Δ_i будет определяться как расхождение между значением на i -м выходе сети и i -м обучающим значением

$$\Delta_i = OUT_i - E_i, \quad i = 1, 2, \dots, n.$$

Разработанный метод решения

Функция ошибки (1) является негладкой и многоэкстремальной, а решением задачи обучения является нахождение ее экстремума. При этом решение будет означать реализацию весов ИНС. Классические методы решения такой задачи, такие как градиентный спуск, метод Ньютона, метод многошаговой редукции размерности, стохастические методы [5], позволяют получить только одно решение задачи оптимизации.

Исследователями делаются попытки использовать для решения задачи оптимизации функции ошибки генетические алгоритмы (ГА). ГА основаны на параллельной обработке конкурирующих вариантов, а их основными достоинствами являются достаточно быстрая сходимость и нечувствительность к локальным экстремумам.

Основными процедурами ГА являются генерация вариантов, отбор, скрещивание и мутация. Разработано большое число модификаций ГА, в том числе мультиколониальные и с управляемым поиском.

Авторами предложен МГА для поиска множества оптимальных решений, в котором особь представляет один из конкурирующих вариантов. В работе использовано вещественное кодирование, при котором гены в хромосоме – вещественные числа, а не биты. Процедура отбора состоит в определении приспособленности хромосомы, которая вычисляется по формуле [6]

$$p_i = \frac{f(Hr_i)}{\sum_{i=1}^{N_{hr}} f(Hr_i)}, \quad (2)$$

где $f(Hr_i)$ – пригодность i -й хромосомы; N_{hr} – количество хромосом или размер популяции.

Более приспособленные особи скрещиваются, и из потомков образуются новые варианты, которые можно оценить с помощью (2). Каждая особь содержит информацию о связях в сети между нейронами и ее весах. Таким образом, можно подстраивать как веса, так и структуру ИНС. Структурная схема алгоритма решения поставленной задачи представлена на рис. 2. С ростом числа хромосом N_{hr} , количество которых задает пользователь, прямо пропорционально возрастает и размер генофонда. В результате отбора, скрещивания и мутации постепенно формируется множество решений.

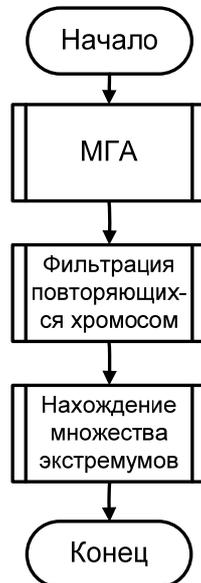


Рис. 2. Структурная схема разработанного алгоритма

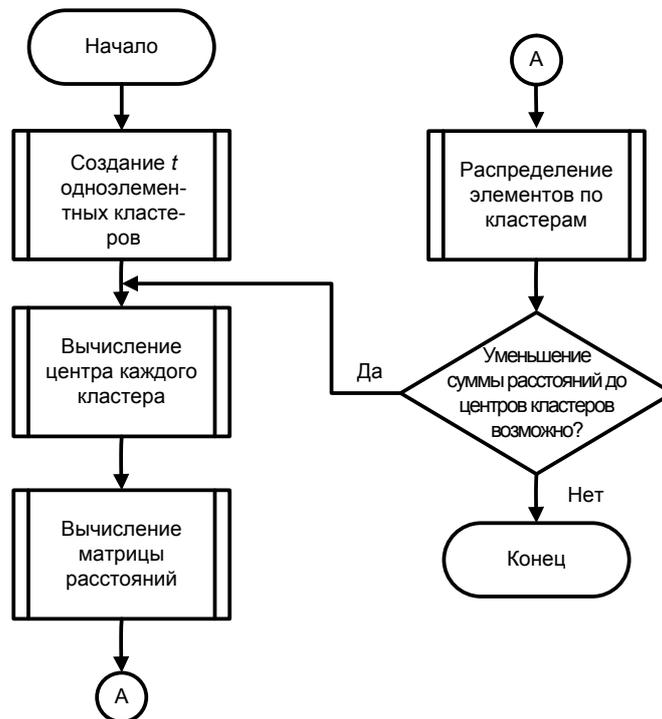


Рис. 3. Структурная схема алгоритма кластеризации данных

Для реализации последней процедуры – нахождения множества решений – использован иерархический метод теории кластерного анализа *K-средних* [7]. Использована стратегия объединения исходных данных в кластеры. Изначально каждому наблюдению соответствует свой кластер (рис. 3). Далее выполняется процедура определения центра $\mu_j, j = 1, 2, \dots, t$ каждого из t кластеров, компоненты которого вычисляются как среднее арифметическое наблюдений, входящих в этот кластер. В центре каждого кластера j вычисляется сумма квадратов расстояний от j -го центра до i -го элемента b_i

$$V = \sum_{j=1}^l \sum_{b_i \in S_j} (b_i - \mu_j(b_i))^2,$$

где S_j – полученные кластеры; $\mu_j(b_i)$ – центры масс векторов $b_i \in S_j$. Затем минимизируется суммарное квадратичное отклонение точек кластеров от центров μ_j .

Результат проверки расстояния от каждого элемента до центра его кластера показывает, следует ли повторить процедуру или закончить выполнение алгоритма.

Экспериментальное исследование алгоритма

Для проведения имитационного эксперимента в качестве функции ошибки (1) была использована тестовая функция Бранинса (Branin's cosine function) [6]:

$$Q(a_1, a_2) = (a_2 - \frac{5,1}{4\pi^2} a_1^2 + \frac{5}{\pi} a_1 - 6)^2 + 10 \cdot (1 - \frac{1}{8\pi}) \cdot \cos(a_1) + 10.$$

Аргументы a_1 и a_2 , изменялись на интервале $[-10; 20]$ для достижения минимума функции $Q(a_1, a_2)$. Сложность поиска состоит в том, что функция Бранинса имеет четыре глобальных экстремума – минимума, расположенных вблизи друг от друга.

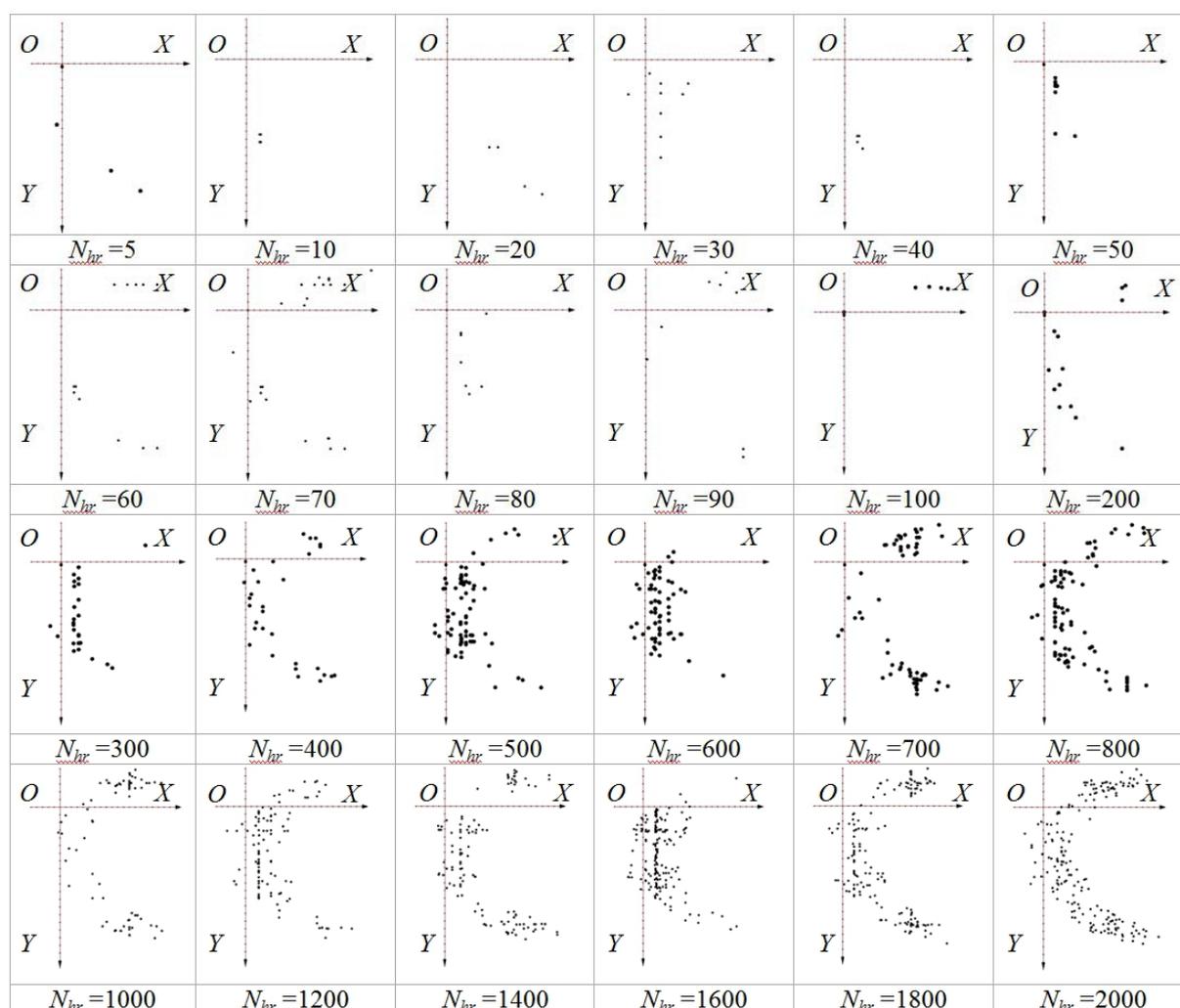


Рис. 4. Рост числа приближений с ростом числа хромосом (проекция на плоскость XOY)

При малом числе хромосом N_{hr} генофонд содержал одно оптимальное решение и несколько приближений к нему (рис. 4). С ростом параметра N_{hr} , возрастает и число попаданий в область всех четырех глобальных минимумов, а также приближений к ним. Исследования показали, что при превышении порога в 900 хромосом дальнейшее увеличение этого параметра приводит к росту числа приближений к истинному значению экстремумов функции $Q(a_1, a_2)$. Результаты получены для значения точности поиска 0,01.

Исследована процедура кластеризации. Исходным являлось множество, полученное в результате работы МГА (генофонд). За достаточное число хромосом принята величина 900. Мощность множества вариантов после фильтрации одинаковых решений (рис. 2) для функции Бранинса равна 83. Результаты работы алгоритма, представленного на рис. 3, приведены в таблице.

Номер найденного кластера	Координаты центра кластера (a_1, a_2)	Значение функции $Q(a_1, a_2)$ в центре кластера	Соответствующие координаты экстремума функции	Значение функции $Q(a_1, a_2)$ в точке экстремума
1	(8,9443;2,3523)	1,5515	(9,4248;2,475)	0,3979
2	(3,3686;1,6596)	0,8423	(3,141;2,275)	
3	(-3,1132;12,2195)	0,4019	(-3,141;12,275)	
4	(15,8464;13,3937)	0,5203	(15,714;12,900)	

Таблица. Результаты работы алгоритма кластеризации

Заключение

Предложена методика для синтеза оптимальной структуры искусственной нейронной сети, основанная на применении двух алгоритмов. Разработанный модифицированный генетический алгоритм, в отличие от классического, позволяет находить множество оптимальных вариантов. Предложенный метод кластеризации данных позволяет разделить области расположения оптимумов на области, в которых найденные центры кластеров максимально приближены к реальным экстремумам. По сравнению с методом обратного распространения ошибки разработанная процедура выгодно отличается тем, что подбор весов осуществляется вне зависимости от формы и количества локальных экстремумов. Это дает возможность при настройке весов нейронных сетей, а впоследствии и ее структуры, отыскать множество оптимальных вариантов, что предоставляет исследователю больше вариантов реализации сложных систем в процессе принятия технических решений. Для развития предложенного подхода планируется разработать алгоритмы настройки структуры искусственной нейронной сети.

Литература

1. Курейчик В.В. Эволюционные, синергетические и гомеостатические методы принятия решений: Монография. – Таганрог: Изд-во ТРТУ, 2001. – 99 с.
2. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир, 1992. – 161 с.
3. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. – 2-е изд. – М.: Физматлит, 2006. – 320 с.
4. Назаров А.В., Лоскутов А.И. Нейросетевые алгоритмы прогнозирования и оптимизации систем. – СПб: Наука и техника, 2003. – 380 с.
5. Бондаренко И.Б. Методы оптимизации проектных решений и технология искусственного интеллекта в интегрированных САПР // Научно-технический вестник СПбГУ ИТМО. – 2005. – № 20. – С. 167–171.
6. Бондаренко И.Б., Гатчин И.Ю., Коробейников А.Г. Разработка модификации генетического алгоритма для поиска множества оптимальных решений тестовой функции. Информационные технологии в профессиональной деятельности и научной работе // Сборник материалов Всероссийской научно-практической конференции с международным участием. Ч. 1. – Йошкар-Ола: Мар. ГТУ, 2011. – С. 111–115.
7. Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д. Анализ данных и процессов. – СПб: БХВ-Петербург, 2009. – 331 с.

Бондаренко Игорь Борисович

– Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, igorlitmo@rambler.ru

Гатчин Юрий Арменакович

– Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gatchin@mail.ifmo.ru

Гераничев Владимир Николаевич

– ФГУП Санкт-Петербургское ОКБ "Электроавтоматика", начальник отдела, geranichev@mail.ru