

УДК 004.94

UDC 004.94

**ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКИХ  
АЛГОРИТМОВ ДЛЯ ОБУЧЕНИЯ  
НЕЙРОННЫХ СЕТЕЙ****USING GENETIC ALGORITHMS FOR  
TRAINING NEURAL NETWORKS**Шумков Евгений Александрович  
к.т.н.Shumkov Eugene Alexandrovich  
Cand.Tech.Sci.Чистик Игорь Константинович  
*Кубанский Государственный Технологический  
Университет, Краснодар, Россия*Chistik Igor Konstantinovich  
*Kuban State Technological University, Krasnodar,  
Russia*В статье показаны способы поиска  
субоптимальных нейронных сетей с помощью  
генетических алгоритмовIn this article we have shown how to find sub-optimal  
neural networks using genetic algorithmsКлючевые слова: ГЕНЕТИЧЕСКИЕ  
АЛГОРИТМЫ, НЕЙРОННЫЕ СЕТИ, ОБУЧЕНИЕ  
НЕЙРОННОЙ СЕТИ, СТОХАСТИЧЕСКИЕ  
МЕТОДЫ ОБУЧЕНИЯKeywords: GENETIC ALGORITHMS, NEURAL  
NETWORKS, NEURAL NETWORK TRAINING,  
STOCHASTIC METHODS OF TRAINING

В настоящее время создано большое количество архитектур нейронных сетей и методов их обучения [4]. Наиболее распространенной является связка – многослойный персептрон (далее МП) и алгоритм обратного распространения ошибки. Для МП также применяют и другие методы, в частности для «грубого» обучения применяют RProp, для «точного» - квазиньютоновские методы и др. [7]. Данные методы, в основном градиентные и не лишены недостатков. Один из основных – точные значения градиента не всегда доступны и кроме того решение задачи оптимизации в условиях большой размерности требует значительных вычислительных, а следовательно и временных затрат [4,6].

Одним из крупных классов обучения нейронных сетей являются вероятностные методы, а также принцип стохастического моделирования [4]. В частности выделяют методы стохастического моделирования: алгоритм Метрополиса, метод моделирования отжига, выборки Гиббса. Популярны и стохастические машины при обучении нейронных сетей – машина Больцмана, сигмоидальные сети доверия и стохастическая машина Гельмгольца. Стохастические методы обучения нейронных сетей черпают свои идеи в основном в статистической механике [4]. Данные методы

обычно относятся, за некоторыми исключениями, к классу самоорганизующихся обучаемых систем.

Другим похожим классом обучения нейронных сетей является использование генетических алгоритмов для поиска весов синаптических связей и/или структуры сети. При использовании генетических алгоритмов для обучения МП можно действовать двумя способами. Первый – искать непосредственно комбинацию весов синаптических связей нейронной сети, а второй – искать нейронную сеть с нужными параметрами (количество слоев и нейронов в них, функция активации, входы – выходы сети и т.д.), то есть решается задача структурной оптимизации. В первом случае генерируются банки особей нейронных сетей и генетическим алгоритмом ищется комбинация синаптических весов, то есть методы обучения нейронных сетей при данном подходе не используются! А во втором случае генерируются особи с разными параметрами и с помощью выбранного метода обучения или нескольких эти нейронные сети обучаются. Здесь также важен следующий момент – можно перебирать входы и выходы нейронной сети<sup>1</sup>. Можно также использовать комбинацию вышеописанных методов, например, на первом этапе с помощью генетического алгоритма ищутся общие параметры нейронной сети, а на втором этапе, также с помощью генетического алгоритма, ищутся значения синаптических весов [1]. Отметим, что во втором случае априори используется большее количество вычислительных затрат, но при этом плюсом является то, что можно быстрее найти нейронную сеть меньшего размера решающую поставленную задачу. Отечественные разработки применения генетических алгоритмов для обучения нейронных сетей можно посмотреть в работах [1,3,6].

---

<sup>1</sup> Выходы нейросети обычно жестко задаются, в частности для прогнозирования обычно используется только один выход.

Использование генетического алгоритма для обучения нейронных сетей обладает следующими достоинствами [1,6]:

- изучение пространства поиска при помощи множества решений, а не одного. Это дает значительный эффект при поиске глобальных минимумов адаптивных рельефов;
- генетические алгоритмы малочувствительны к росту размерности множества оптимизации.

В тоже время существует значительный недостаток использования генетических алгоритмов – серьезные вычислительные, а следовательно и временные затраты. Также генетические алгоритмы содержат в себе элементы эвристики и соответственно обладают всеми недостатками присущими эвристическим методам.

Управляемыми параметрами в генетическом алгоритме являются:

- длина хромосомы;
- наполнение хромосомы (локусы и аллели);
- параметры оператора кроссовера;
- параметры оператора мутации;
- параметры оператора инверсии;
- параметры выбора лучших особей;
- параметры генерации начальной и последующих популяций и т.д.

То есть видно, что в генетическом алгоритме можно подстраивать практически всё. При этом есть только общие рекомендации, как это делать. Конкретная реализация ложится полностью на плечи разработчика.

Относительно операции кроссовера можно управлять количеством точек разрезов и собственно позициями точек разрезов. Обычно для упрощения выбирается один из типов кроссовера – однотоочечный или многотоочечный, а также жестко закрепляются позиции разрезов. В операции мутации можно варьировать количеством изменяемых генов и

ограничениями на изменения. При проведении инверсии также можно управлять количеством переставляемых генов.

Следует отметить, так как генетические алгоритмы во многом опираются на вероятностный выбор, то при программной реализации необходимо использовать качественный генератор случайных чисел.

Будем называть *субоптимальным набором* весов синаптических связей МП, такой набор весов, при котором МП решает задачу<sup>2</sup> с ошибкой меньше заданной. Набор называется *субоптимальным*, а не *оптимальным*, потому – что всегда можно найти другой набор весов, который будет решать эту же задачу с меньшей ошибкой, но поиск «лучшего» набора, возможно займет гораздо больше времени [2].

Предложим архитектуру системы, использующую генетический алгоритм для поиска субоптимальной нейронной сети для поставленной задачи. Исходным моментом для любой нейросетевой задачи, является обучающая выборка, описывающая решаемую задачу. Здесь мы предлагаем использовать базу данных временных рядов в стандартном исполнении.

Предлагаемая топология выглядит, как показано на Рисунке 1.

---

<sup>2</sup> Мы имеем в виду «сложные» задачи, например, прогнозирование курса валют, распознавание изображения с веб – камеры и т.д.

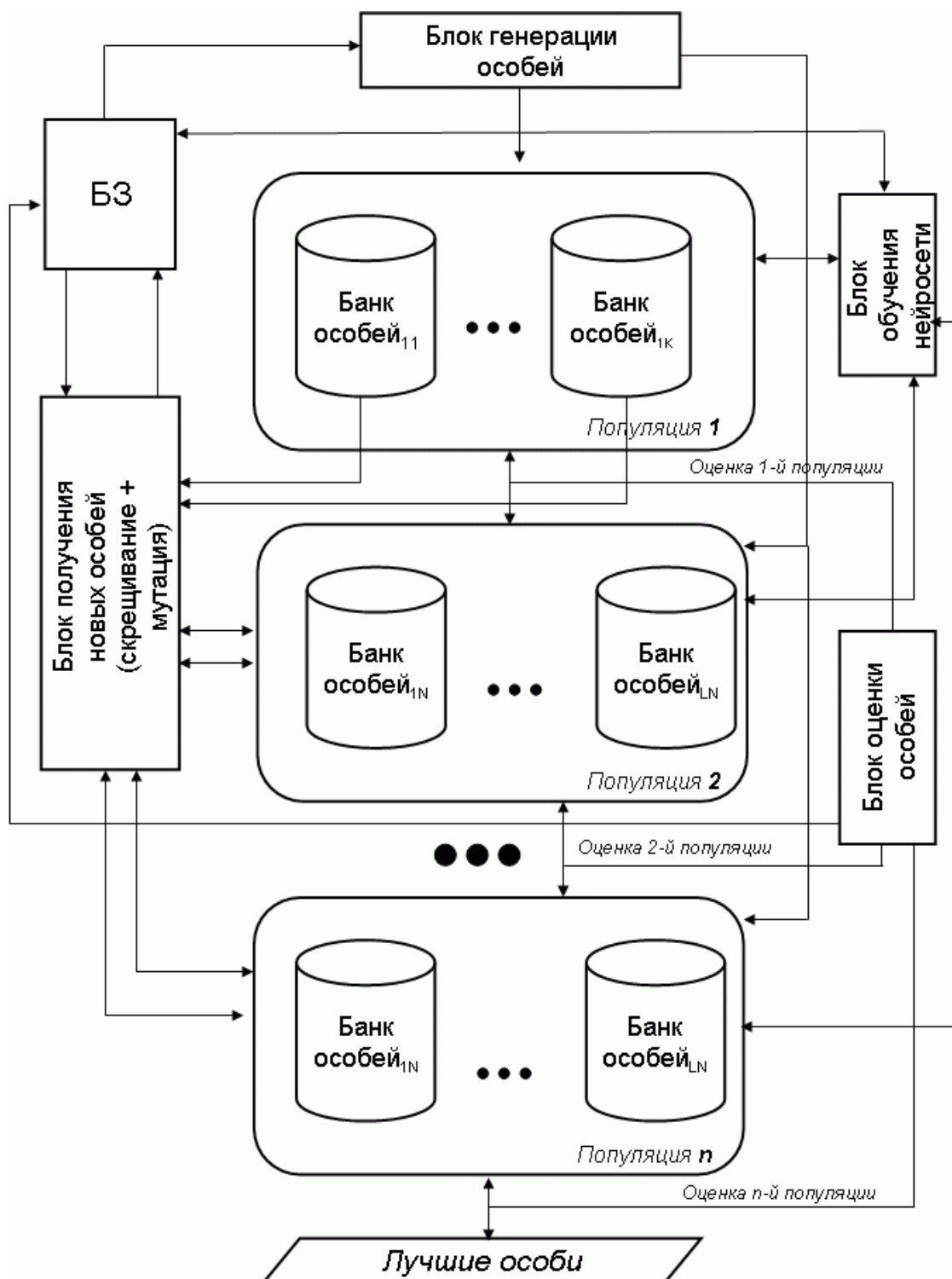


Рисунок 1. Архитектура системы поиска нейронной сети с помощью генетического алгоритма

На рисунке 1 показана общая архитектура топологии для поиска нейросети с использованием метода обучения, то есть с помощью

генетического алгоритма идет только поиск общей структуры нейронной сети.

Отметим наличие в топологии БЗ – базы знаний, основная задача которой сохранение необходимых данных о предыдущих поисках (успешных и неуспешных) и, следовательно, сокращение времени на поиск адекватной модели при новых стартах. В БЗ должны протоколироваться все запуски поиска нейронных сетей с конечными результатами, статистические параметры временных рядов решаемых задач. В топологии, кроме БЗ, ключевыми являются блоки получения новых особей, блок обучения нейросети, блок оценки особей и генератор особей. Блок оценки особей, кроме собственно оценки особей, также выбирает по определенному алгоритму особи для следующего поколения, например, по принципу «рулетки» или турнирного отбора. Блок оценки особей должен быть абсолютно независим от остальных блоков рассматриваемой топологии. Блок генерации особей работает не только при формировании первоначальных банков, но и в последующие этапы. Принцип работы БЗ не должен нивелировать стохастичность поиска генетического алгоритма, БЗ должна определять стартовую позицию генерации особей, помогающую быстрее выходить на решение задачи. Обучение нейронных сетей происходит в отдельной компоненте – «блоке обучения нейросетей». В данный блок поступает общая хромосома (см. рисунок 2) с параметрами для конструктора нейросети и рекомендации по процессу обучения (метод обучения, скорость обучения, расписание обучения и др.) от БЗ по итогам предыдущих запусков, а на выходе выдаются матрицы весов синаптических связей.

В качестве целевой функции можно выбирать из нескольких вариантов формул, оценивающих качество обучения нейронной сети, остановимся на одной из них (по примерам и по выходам):

$$GF = \sum_{j=1}^m \frac{\sum_{i=1}^k |y_i^j - \hat{y}_i^j|}{k \cdot m}, \quad (1)$$

где  $k$  - количество выходов нейронной сети;  $m$  - количество примеров в валидационной выборке;  $y$  - желательное значение выхода;  $\hat{y}$  - реальное значение выхода.

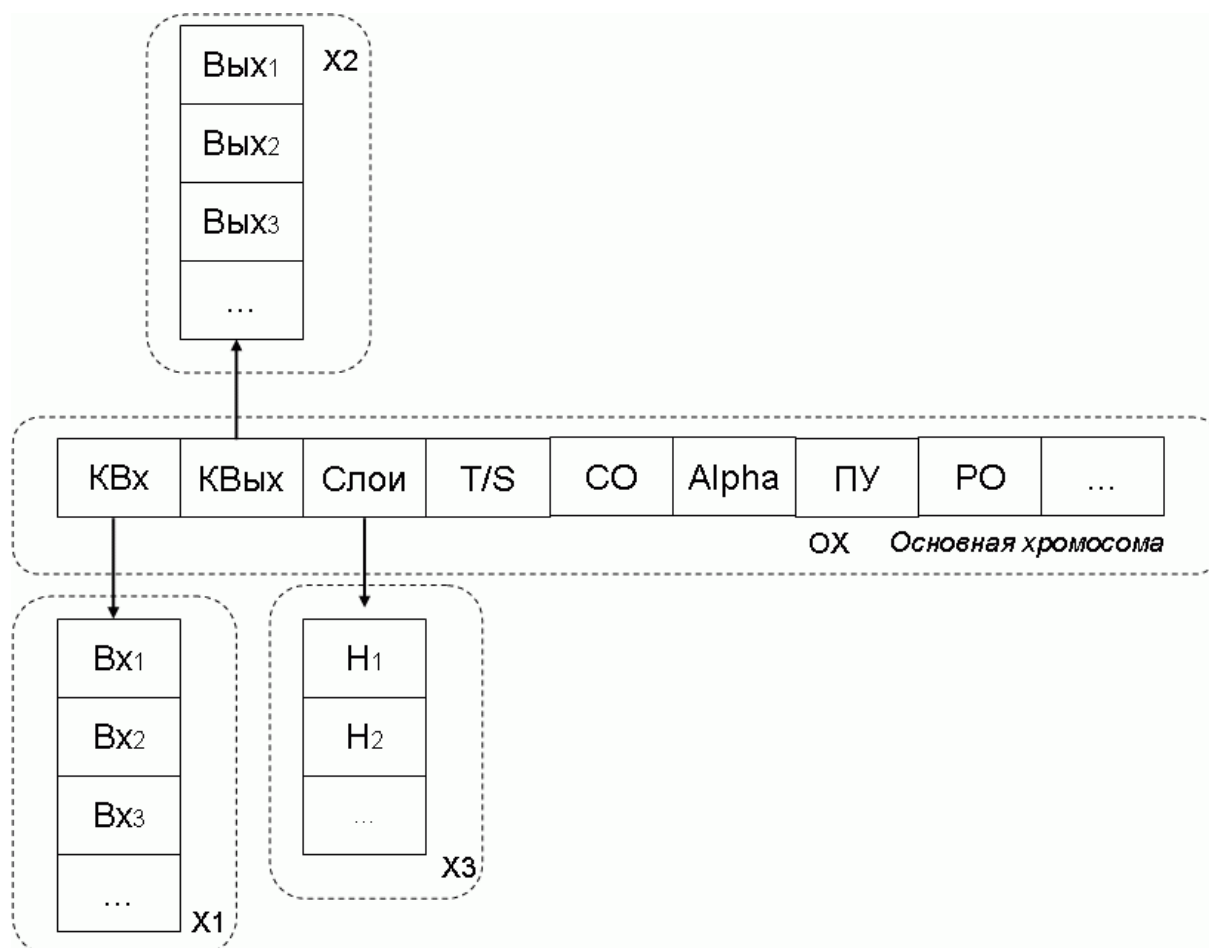


Рисунок 2. Хромосомы, описывающие нейронную сеть

Особь предлагается делать составной из нескольких хромосом, т. к. вариаций по количеству входов, слоев и нейронов в них может быть большое количество. Пусть основная хромосома состоит из следующих генов:  $KVx$  – количество входов нейросети;  $KВых$  – количество выходов нейросети; Слои – количество слоев;  $T/S$  – тип функции активации (сигмоида/ тангенсоида/ ..); Alpha – коэффициент функции активации;  $CO$  – тип расчета скорости обучения; ПУ – пороговый уровень нейронов; PO –

тип расчета ошибки на выходе нейросети и т.д. К генам  $KV_{\text{вх}}$ ,  $KV_{\text{вых}}$  и  $\text{Слои}$  прикреплены дочерние хромосомы, определяющие соответственно непосредственные входы, выходы и количество нейронов в слоях.

Принцип работы операторов генетического алгоритма - мутации, кроссовера и инверсии для модифицированной хромосомы следующий. Оператор кроссовер работает только с основной хромосомой. Оператор инверсии работает с основной хромосомой и дочерними, отвечающими за входы и выходы. Оператор мутации работает с основной хромосомой и дочерней, отвечающей за количество нейронов. То есть за входы и выходы отвечает только генератор новых особей и БЗ. Суммарный оператор на каждой итерации генетического алгоритма при работе с модифицированной хромосомой имеет следующий вид:

$$\{ \text{мутация}_{\text{OX}}(g_1); \text{мутация}_{\text{X3}}(g_2); \text{кроссовер}_{\text{OX}}(g_3); \text{инверсия}_{\text{OX}}(g_4); \text{инверсия}_{\text{X1}}(g_5); \text{инверсия}_{\text{X2}}(g_6) \},$$

где  $g_i$  - вероятность выполнения определенного оператора.

Можно модифицировать данную топологию под задачу поиска нейронной сети с варьируемыми топологией и методом обучения. То есть ищется та связка топология + метод обучения, которая справится с решением поставленной задачи.

Другой модификацией является наличие на конечном этапе принципа конгломерата нейронных сетей. То есть, с помощью генетического алгоритма ищется  $n$  субоптимальных нейронных сетей и далее используется агрегированный ответ. Данный принцип неплохо себя зарекомендовал в задачах прогнозирования с помощью нейронных сетей [7]. Также нередко используют принцип «конкуренции и кооперации» в генетических нейро-нечетких системах [8].



Отметим еще одно совместное применение генетических алгоритмов и нейронных сетей – это настройка функций принадлежности нечетких и нейро-нечетких систем [8].

В настоящее время нейронные сети и генетические алгоритмы перешли из перспективных методов решения задач в доступные и широко применяемые методы. Основное направление при разработке, как новых методов, так и при решении практических задач, связано с поиском путей уменьшения временных затрат и поиску шаблонов быстрого прототипирования. Один из подходов рассмотрен в данной статье – поиск параметров нейронной сети с помощью генетического алгоритма. Применение данного подхода позволяет упростить с позиции инженера-разработчика поиск структуры нейронной сети, в частности поиск входов, количества слоев и нейронов в них, метод обучения и т.д.

#### **Литература:**

1. Божич В.И., Лебедев О.Б., Шницер Ю.Л. Разработка генетического алгоритма обучения нейронных сетей // Известия ТРТУ. Тематический выпуск. 2001. Т. 22. №4. С. 170-174.
2. Зуева В.Н. Математическое обеспечение информационной системы учета и прогнозирования продовольственных ресурсов. Дисс. ... канд. техн. наук. Краснодар. КубГТУ. 2007. 141 с.
3. Курейчик В.М. Генетические алгоритмы: Монография. Таганрог: Изд-во ТРТУ, 1998. 242 с.
4. Хайкин С. Нейронные сети: полный курс, 2-е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2006. 1104 с.
5. Частикова В.А. Исследование основных параметров генетического алгоритма метода генетических схем в интеллектуальных системах, основанных на знаниях // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2011. №69. С. 151-163.
6. Чипига А.Ф., Воронкин Р.А. Обучение искусственных нейронных сетей путем совместного использования методов локальной оптимизации и генетических алгоритмов // Известия ТРТУ. Т. 33. №4. С. 172-174.
7. Шумков Е.А. Система поддержки принятия решений предприятия на основе нейросетевых технологий. Дисс. ... канд. техн. наук. Краснодар: КубГТУ, 2004. 158 с.
8. Ярушкина Н.Г. Основы нечетких и гибридных систем: Учеб. Пособие. – М.: Финансы и статистика. 2004. 320 с.

**Literatura:**

1. Bozhich V.I., Lebedev O.B., Shnicer Ju.L. Razrabotka geneticheskogo algoritma obucheniya nejronnyh setej // Izvestija TRTU. Tematicheskij vypusk. 2001. T. 22. №4. S. 170-174.
2. Zueva V.N. Matematicheskoe obespechenie informacionnoj sistemy ucheta i prognozirovaniya prodovol'stvennyh resursov. Diss. ... kand. tehn. nauk. Krasnodar. KubGTU. 2007. 141 s.
3. Kurejchik V.M. Geneticheskie algoritmy: Monografija. Taganrog: Izd-vo TRTU, 1998. 242 s.
4. Hajkin S. Nejronnye seti: polnyj kurs, 2-e izdanie.: Per. s angl. – M.: Izdatel'skij dom «Vil'jams», 2006. 1104 s.
5. Chastikova V.A. Issledovanie osnovnyh parametrov geneticheskogo algoritma metoda geneticheskikh shem v intellektual'nyh sistemah, osnovannyh na znanijah // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2011. №69. S. 151-163.
6. Chipiga A.F., Voronkin R.A. Obuchenie iskusstvennyh nejronnyh setej putem sovmestnogo ispol'zovaniya metodov lokal'noj optimizacii i geneticheskikh algoritmov // Izvestija TRTU. T. 33. №4. S. 172-174.
7. Shumkov E.A. Sistema podderzhki prinjatija reshenij predpriyatija na osnove nejrosetevykh tehnologij. Diss. ... kand. tehn. nauk. Krasnodar: KubGTU, 2004. 158 s.
8. Jarushkina N.G. Osnovy nechetkih i gibridnyh sistem: Ucheb. Posobie. – M.:Finansy i statistika. 2004. 320 s.