

УДК 007.52

ВЫЧИСЛИТЕЛЬНЫЕ РЕГУЛЯТОРНЫЕ СЕТИ И САМОАДАПТИВНЫЙ НЕЙРОЭВОЛЮЦИОННЫЙ АЛГОРИТМ*

Ю.Р. Цой¹

В докладе описывается концепция вычислительных регуляторных сетей (ВРС), объединяющая искусственные нейронные сети, эволюционные вычисления и генные регуляторные сети. Рассматриваются основные отличия ВРС от нейронных сетей. Описывается алгоритм эволюции ВРС с самоадаптацией, и представлены результаты первых исследований на примере задач классификации, показывающие его работоспособность.

Введение

Одной из центральных проблем нейроинформатики является разработка методов и алгоритмов генерации искусственных нейронных сетей (ИНС) большого размера (с количеством нейронов порядка 10^3 - 10^5 и более), способных решать сложные многоплановые задачи распознавания образов, адаптивного управления и поведения в динамических средах и др. Данная проблема особенно актуальна для понимания того, что «могут» ИНС, где границы их применимости.

Одной из проблем использования нейроэволюционных (НЭ) алгоритмов [Yao, 1999, Цой, 2007] для поиска ИНС большого размера является проблема кодирования информации. Неоднократно отмечалось, что использование прямого кодирования, например, списка межнейронных связей, либо матрицы смежности, малоэффективно для поиска ИНС со сложной структурой, содержащей тысячи нейронов [Stanley, Miikkulainen, 2003]. При этом, очевидно, что в ДНК информация о структуре нейронных связей в мозгу человека кодируется непрямым, косвенным образом, т.к. количество нейронов в мозгу, равно 10^{10} - 10^{11} (по разным оценкам), значительно превышает количество активных генов, составляющее по различным оценкам от 20000 до 30000 [Zigmond et al., 1999]. Поиск ИНС большого размера осложняется также тем, что эффект от

* Работа выполнена при финансовой поддержке РФФИ (проект № 06-08-00840).

¹ *Томский политехнический университет, 634050, Томск, пр. Ленина, 30
qai@mail.ru*

выполнения элементарных операций модификации ИНС (добавление/удаление нейрона или связи, изменение веса связи и т.д.), использующихся в НЭ алгоритмах, в большинстве случаев будет иметь очень малое воздействие на результат работы ИНС. Это значительно увеличивает время поиска сложной ИНС.

Реальные нейросетевые структуры мозга представляют собой сети с достаточно большим количеством узлов. В рамках активно развивающегося со второй половины 20 века направления по исследованию сложных сетей были выявлены свойства, общие для многих реальных сложных сетей (Интернет, социальные сети и др.), которые не могут быть смоделированы в рамках теории случайных графов и теории перколяций [Albert, Barabasi, 2002]. Этими свойствами являются ярко выраженная кластерная структура реальных сложных сетей, а также степенной закон распределения количества связей каждого узла. При этом наличие данных свойств не зависит от размеров сети, что послужило основанием для того, чтобы называть сети, обладающие такими свойствами, безмасштабными (scale-free) [Albert, Barabasi, 2002].

В докладе предлагается способ неявного кодирования информации об ИНС, основанный на идее генных регуляторных сетей (ГРС), а также алгоритм поиска ИНС, использующий предлагаемый способ и учитывающий свойство безмасштабности.

1. Генные регуляторные сети

Набор взаимодействующих сегментов (генов) ДНК в клетке образует генную регуляторную сеть (gene regulatory network). Взаимодействие между различными генами осуществляется неявным образом посредством белков и ферментов, производимых в результате экспрессии генов, а также с использованием имеющихся в клетке веществ. Динамика ГРС существенно зависит от начальной концентрации химических соединений. Схематичный пример генной регуляторной сети приведен на рис. 1, для упрощения будем считать, что гены производят только белки.

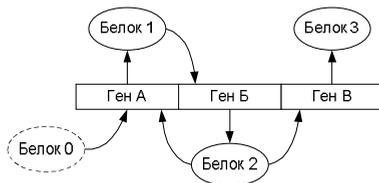


Рис. 1. Схематичный пример ГРС

В примере на рис. 1 считаем, что экспрессия любого гена начинается, если в среде присутствует хотя бы один белок, который используется

данным геном. Показанный пример соответствует случаю, когда наличие белка 0 приводит к производству белка 3 путем последовательной экспрессии генов А, Б и В. При этом исчезновение белка 0 не приведет к деактивации ГРС в силу появления обратной связи от гена Б к гену А.

Интересным свойством ГРС является то, что в ГРС отсутствует понятие «связи», как самостоятельной единицы. Если пренебречь конечной скоростью распределения вещества в пространстве, то для связывания двух генов необходимо и достаточно, чтобы один ген производил некоторый белок, а другой ген этот белок использовал.

Отметим, что идея использования ГРС в качестве основы для представления (кодирования) вычислительных сетей (таких как искусственные нейронные сети) в эволюционных алгоритмах не нова [Banzhaf, 2003, Eggenberger, 1997, Mattiussi, Floreano, 2007].

2. Вычислительная регуляторная сеть

Описанный выше принцип «связывания» генов можно использовать и для описания произвольной сети.

Обозначим $\mathbf{V} = \{V_i\}, i = 1, 2, \dots, N$ – набор узлов сети. Каждому i -му узлу поставим в соответствие множество идентификаторов входных сигналов $\Psi_i = \{\psi_{i,k}\}, k = 1, 2, \dots, N_{\psi_i}$, и множество идентификаторов выходных сигналов $\mathbf{j}_i = \{j_{i,k}\}, k = 1, 2, \dots, N_{j_i}$. Будем называть два узла i_1 и i_2 связанными, если выполняется хотя бы одно из следующих условий:

1. $\exists j, k : j_{i_1, j} = y_{i_2, k}$;
2. $\exists j, k : y_{i_1, j} = j_{i_2, k}$.

Полученную в результате связывания описанных узлов сеть будем называть *регуляторной сетью* (РС). Пример отдельного узла, а также сети из нескольких связанных узлов показан на рис. 2.

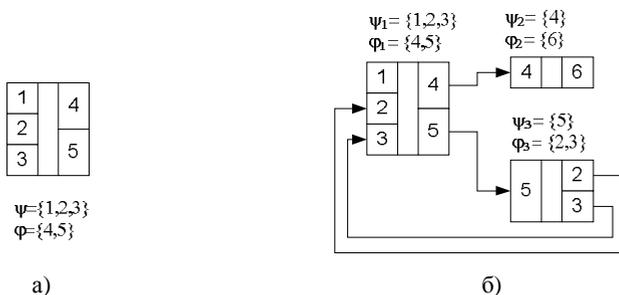


Рис. 2. Пример отдельного узла (а), а также сети из нескольких связанных узлов (б). Идентификаторы в левой части узла соответствуют входным сигналам, а в правой части – выходным. Связи между узлами показаны стрелками

Сопоставим входным и выходным сигналам i -го узла соответственно

входные и выходные веса $\mathbf{w}_i = \{w_{i,k}\}$, $k = 1, 2, \dots, N_{\psi_i}$ и $\mathbf{v}_i = \{v_{i,k}\}$, $k = 1, 2, \dots, N_{j_i}$ таким образом, что входному сигналу с идентификатором $\psi_{i,k}$ соответствует вес $w_{i,k}$, а выходному сигналу $j_{i,k}$ – вес $v_{i,k}$.

Полагая, что РС находится в некоторой среде, обозначим мощность сигнала с идентификатором a в этой среде в момент времени t как x_a , а вектор, соответствующий значениям мощности сигналов, поступающих на вход i -го узла, как $\mathbf{x}_i(t)$.

Также поставим в соответствие каждому узлу сети величину спонтанной активности f_i и вычисляемую этим узлом функцию $f_i(t)$. Будем считать, что мощности выходных сигналов узла вычисляются как:

$$x_{a,j}(t) = f_i(\mathbf{x}_i(t), \mathbf{w}_i, f_i, v_{i,k}), \quad (1)$$

где k – номер идентификатора выходного сигнала i -го узла такого, что $j_{i,k} = a$. Если «подходящего» k не существует, то $v_{i,k} = 0$.

Зададим правило обновления мощности сигналов:

$$x_a(t+1) = g(x_a) + \sum_i x_{a,j}(t), \quad (2)$$

где $g(\cdot)$ – функция, соответствующая активности среды и/или других (неизвестных, либо не учитываемых) РС, производящих сигнал a . При отсутствии этих факторов $g = 0$.

Будем называть РС, функционирующую согласно (1) в некоторой среде, обновляемой по правилу (2), *вычислительной регуляторной сетью* (ВРС).

Отличительными особенностями РС и ВРС являются: 1) неявная связанность узлов за счет сигналов с одинаковыми идентификаторами; 2) возможность существования множественных связей между двумя узлами; 3) функционирование в некоторой среде, в которой могут присутствовать сигналы, отличные от входных и выходных для данной РС.

Преимуществом использования РС для задания структуры ИНС является возможность объединения нескольких узлов в один, если рассматриваемые узлы имеют одинаковые наборы входных сигналов (рис. 3).

3. Операторы модификации ВРС

Рассмотрим следующие операторы модификации (мутации) ВРС:

1. Оператор изменения веса входного сигнала узла.
2. Оператор изменения веса выходного сигнала узла.
3. Оператор изменения величины спонтанной активности узла.
4. Оператор изменения типа узла.

5. Оператор изменения идентификатора входного сигнала узла.
6. Оператор изменения идентификатора выходного сигнала узла.
7. Оператор включения/выключения узла.
8. Оператор добавления/удаления входного сигнала узла.
9. Оператор добавления/удаления выходного сигнала узла.
10. Оператор разделения узла.
11. Оператор добавления/удаления узла.

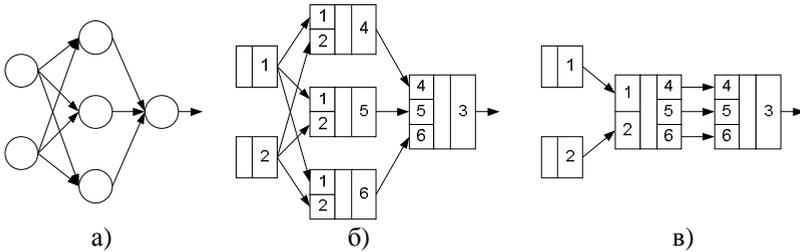


Рис. 3. Пример использования РС для компактной записи информации о структуре ИНС: а) структура ИНС; б) соответствующая РС; в) РС с объединенными по входам узлами

Вероятность выбора различных операторов характеризуется вектором параметров мутации $\mathbf{m} = \{m_k\}$, $k = 1, 2, \dots, 11$, $m_k \in \mathfrak{X}$, причем вероятность выбора k -го оператора пропорциональна величине:

$$\frac{|m_k|}{\sum_k |m_k|}.$$

Вероятность мутации определяется по формуле:

$$P_m = \begin{cases} 1, & \bar{m} > 1 \\ \bar{m}, & \bar{m} \leq 1 \end{cases}, \quad \bar{m} = \frac{b}{11} \sum_k |m_k|,$$

где b – константа.

В результате мутации вектор параметров мутации также подвергается изменениям: $m_k(t+1) = \left(\frac{1}{11} + \frac{c_k}{\sum_i c_i}\right) x_G$. где x_G – случайная величина,

распределенная по закону $N(0;1)$; c_k – количество использований k -го оператора мутации для данной особи с момента ее инициализации, другими словами, вектор $\chi = \{c_k\}$, $k = 1, 2, \dots, 11$, $c_k \in \mathbf{N}$, представляет вектор истории мутаций данной особи.

Для обеспечения свойств, присущих безмасштабным сетям, будем выбирать узлы, подвергающиеся мутации пропорционально количеству

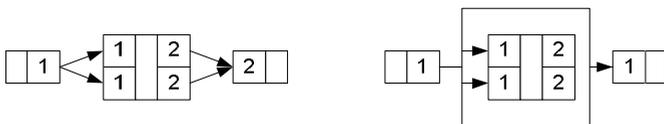
связей (степеней) d_i узлов. Выбор входных и выходных сигналов, характеристики которых подвергаются изменениям в результате мутации, также зависит от того, насколько часто эти сигналы используются в РС.

К примеру, выбор i -го узла, который будет разделен оператором разделения, осуществляется пропорционально величине:

$$P(i) = \begin{cases} \frac{d_i + g}{\sum_k (d_k + g)}, & m > 0 \\ \frac{d_{\max} - d_i + g}{\sum_k (d_{\max} - d_k + g)}, & m \leq 1 \end{cases}$$

где d_{\max} – максимальная степень узла в сети, g – константа, m – значение параметра мутации, соответствующего оператору разделения узла.

Отметим, что, несмотря на элементарность действий, осуществляемых многими из вышеперечисленных операторов, эффект может носить значительный характер для всей ВРС, поскольку эти операторы влияют на характеристики сигналов, распространяемых по ВРС. Так, например, изменение веса выходного сигнала в одном из узлов окажет влияние на мощность этого сигнала во всей среде и, тем самым, повлияет на все узлы, в которых этот сигнал является входным. Также изменение идентификатора входного сигнала может существенно изменить структуру ВРС. На рис. 4 изображен иллюстративный пример такого изменения.



а) до изменения

б) после изменения

Рис. 4. Пример изменения структуры ВРС в результате изменения идентификатора входного сигнала в крайнем правом узле

4. Результаты экспериментов

Для экспериментов выбраны следующие параметры: размер популяции $n = 200$ особей, используется турнирная селекция с размером турнира равным 8, длительность эволюционного поиска $g = 2000$ поколений. Все особи в начальной популяции представляют ВРС с одним узлом, входные сигналы которого соответствуют входным сигналам ИНС, а выходные – выходным сигналам ИНС.

Рассматривается задача реализации с помощью ИНС логической функции «Исключающее ИЛИ». Результаты решения задачи для

различных b , усредненные по 100 запускам, приведены в табл. 1.

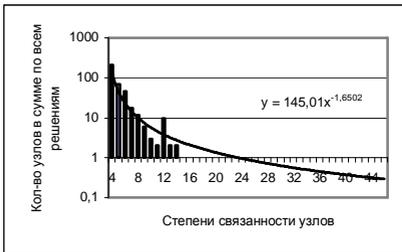
Из приведенных данных видно, что увеличение вероятности мутации способствует повышению эффективности работы алгоритма и уменьшению размеров полученных ВРС. Можно предположить, что эффективность поиска зависит от того, насколько существенно изменяется в результате мутации ВРС, поскольку множественные модификации ВРС позволяют лучше «проверить» содержимое вектора \mathbf{m} параметров мутации и тем самым способствуют отбору тех особей, изменения в которых имеют более направленный характер.

Табл. 1. Результаты экспериментов для различных значений b

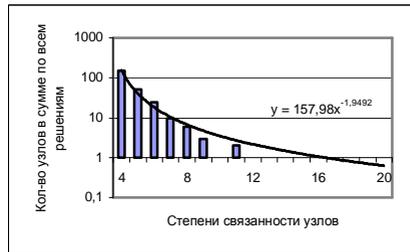
	$b = 1$	$b = 10$	$b = 100$
Количество вычислений целевой функции, N_{FE}	52023,4 $\pm 49830,61$	8355,97 $\pm 3262,22$	4612,53 $\pm 2594,07$
Время поиска, поколения	259,65 $\pm 197,12$	41,4 $\pm 36,58$	22,62 $\pm 13,02$
Время поиска, сек.	3,05 \pm 4,20	0,50 \pm 0,58	0,32 \pm 0,34
Среднее количество узлов	10,03 \pm 4,95	8,54 \pm 3,60	6,89 \pm 2,31
Успешных запусков, %	100%	100%	100%

Хвосты общего распределения степеней d_i для найденных решений со степенной линией тренда представлены на рис. 5, по оси абсцисс отложены степени узлов, а по оси ординат – общее количество узлов с соответствующими степенями в полученных решениях (логарифмическая шкала). Также приведены уравнения аппроксимирующей кривой.

Из приведенных на рис. 5 графиков видно, что распределение степеней подчиняется показательному закону, что характерно для безмасштабных сетей. Отметим тот важный факт, что показатель экспоненты, характеризующий распределение степеней узлов, зависит от величины b .



а) $b = 1$



б) $b = 100$

Рис. 5. Вид хвоста распределения степеней d_i узлов в зависимости от b

Тестирование предлагаемого алгоритма на задачах классификации показало его применимость по сравнению с ИНС, настроенными

«вручную», однако из-за недостатка места, результаты тестирования не приводятся.

Заключение

Предложенный способ кодирования и механизм использования операторов модификации ВРС позволяет реализовать более компактное представление информации о структуре ИНС с возможностью существенных воздействий на свойства этой ИНС за счет действия на сигналы, распространяемые в ВРС. Предложенный алгоритм поиска ВРС позволяет получать безмасштабные сети, соответствующие взвешенным ориентированным графам, что представляет интерес с точки зрения теории сложных сетей, поскольку большинство существующих в настоящее время моделей не рассматривают направленность связей и оперируют бинарными весами связей [Albert, Barabasi, 2002].

Возможности и свойства ВРС подлежат дальнейшему изучению. Представляют интерес исследования возможности применения ВРС для решения различных задач классификации, аппроксимации, моделирования, оптимизации и др.

Программная реализация ВРС и операторов осуществлена в инструментальной библиотеке классов ECWorkshop (<http://qai.narod.ru/ecw/>) и находится в открытом доступе.

Список литературы

- [Yao, 1999] Yao X. Evolving artificial neural networks // Proceedings of the IEEE, 1999, vol. 87, no. 9, pp. 1423-1447. WEB: <http://www.cs.bham.ac.uk/~xin/>
- [Цой, 2007] Цой Ю.Р. Введение в нейроэволюционный подход: основные концепции и приложения // IX Всерос. науч.-техн. Конф. "Нейроинформатика-2007": Лекции по нейроинформатике. Часть 2. – М.:МИФИ, 2007. – С. 43-76.
- [Stanley, Miikkulainen, 2003] Stanley K.O., Miikkulainen R. A taxonomy for artificial embryogeny // Artificial Life, 2003, vol. 9, no. 2, pp. 93–130.
- [Zigmond et al., 1999] Fundamental neuroscience / Eds.: Zigmond M.J., Bloom F.E., Landis S.C., Roberts J.L., Squire L.R. – London: Academic Press, 1999.
- [Albert, Barabasi, 2002] Albert R., Barabasi A.-L. Statistical mechanics of complex networks // Reviews of Modern Physics, vol. 74, no. 1, 2002, pp. 47-97.
- [Banzhaf, 2003] Banzhaf W. On the dynamics of an artificial regulatory networks // In W. Banzhaf, T. Christaller, P. Dittrich, J. Kim, J. Ziegler (Eds): Advances in Artificial Life, 7th European Conference ECAL-2003, Berlin: Springer, 2003
- [Eggenberger, 1997] Eggenberger P. Creation of neural networks based on developmental and evolutionary principles // W. Gerstner, A. Germond, M. Hasler, J. Nicoud (Eds.): Proc. Int. Conf. Artif. Neural Netw., Lausanne, 1997, vol. 1327, pp. 337–342. <http://citeseer.ist.psu.edu/eggenberger97creation.html>
- [Mattiussi, Floreano, 2007] Mattiussi C., Floreano D. Analog Genetic Encoding for the Evolution of Circuits and Networks // IEEE Transactions on Evolutionary Computation, 2007, Vol. 11, No. 5. P. 596-607.