

Інфокомунікаційні системи

УДК 519.71

А.А. Бессонов

Харьковский национальный университет радиоэлектроники, Харьков

ОБУЧЕНИЕ РАДИАЛЬНО-БАЗИСНЫХ СЕТЕЙ С ПОМОЩЬЮ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ С АДАПТИВНОЙ МУТАЦИЕЙ

Рассматривается эффективность применения адаптивных мутаций в генетических алгоритмах, используемых для обучения радиально-базисной сети и оптимизации ее структуры. Исследуются следующие типы адаптивных мутаций: уменьшающаяся мутация Коши, адаптивные мутации Гаусса и Лапласа. Проводится их сравнительный анализ с помощью имитационного моделирования.

Ключевые слова: нейронная сеть, генетический алгоритм, адаптивная мутация, базисная функция, селекция, фитнес-функция, хромосома, ген.

Введение

Радиально-базисные сети (РБС) являются весьма популярным инструментом при моделировании различных нелинейных систем. В связи с этим возникла необходимость разработки новых эффективных алгоритмов одновременной настройки как параметров сети, так и ее структуры. В работе [1] для решения этой задачи было предложено использовать генетические алгоритмы (ГА), которые с помощью операторов скрещивания и мутации позволили получить оптимальную для решаемой задачи структуру РБС и произвели предварительную настройку ее параметров. Однако особенностью РБС, в отличие, например, от многослойного персептрона, является неоднородность диапазонов значений ее параметров. Очевидно, что значения таких параметров сети, как вес связи между нейронами скрытого и выходного слоев, центр и дисперсия базисной функции, лежат в различных диапазонах. Это приводит к тому, что использование традиционной равномерной мутации значительно затягивает процесс обучения. В данной работе для настройки параметров РБС предлагается использовать адаптивные мутации различных типов: уменьшающаяся мутация Коши, адаптивные мутации Гаусса и Лапласа. С помощью имитационного моделирования проводился сравнительный анализ данных типов мутаций и исследовалось их влияние на скорость обучения РБС.

Структура РБС

РБС являются универсальными аппроксиматорами и при этом обладают простой структурой и отличаются высокой скоростью обучения. Они осуществляют аппроксимацию функции $f(\mathbf{x})$ некоторой системой базисных функций (БФ) – нелинейных функций $\varphi(\mathbf{x}, \mathbf{t})$, зависящих от расстояния (радиального) $r = \|\mathbf{x} - \mathbf{t}\|$, где \mathbf{t} – вектор центров БФ.

Представление нелинейной модели радиально-базисной сетью имеет вид

$$\hat{y}(k) = a_0 + \sum_{i=1}^N w_i \Phi_i(\mathbf{x}(k), \boldsymbol{\mu}, \boldsymbol{\sigma}), \quad (1)$$

где a_0 – смещение нейрона выходного слоя; w_i – вес связи i -го нейрона скрытого слоя с нейроном выходного слоя; N – число нейронов в скрытом слое; Φ_i – базисная функция (БФ) i -го нейрона; $\boldsymbol{\mu}, \boldsymbol{\sigma}$ – параметры (центры и радиусы соответственно) БФ.

Если в качестве базисной функции $\Phi_i(x)$ выбрать

$$\Phi_i(x) = \exp\left\{-\frac{\|\mathbf{x} - \boldsymbol{\mu}_i\|^2}{\sigma_i^2}\right\}, \quad (2)$$

где $\boldsymbol{\mu}_i, \sigma_i$ – центры и радиусы базисных функций соответственно; $\|\cdot\|$ – евклидова норма, то в процессе обучения определению подлежит вектор

$$\mathbf{w}(k) = (a_0(k), w_1(k), \boldsymbol{\mu}_1^T(k), \sigma_1(k), \dots, w_N(k), \boldsymbol{\mu}_N^T(k), \sigma_N(k))^T.$$

В настоящее время разработано и хорошо изучено достаточно большое количество алгоритмов настройки параметров сети, однако вопросы выбора ее оптимальной структуры остаются открытыми.

Применение же нейроэволюционного алгоритма (НА) для обучения РБС позволяет одновременно решать задачи настройки параметров сети и определения ее оптимальной структуры. Самым распространенным видом эволюционных алгоритмов являются генетические (ГА).

Обучение РБС с помощью ГА

В ГА каждая особь кодируется сходным с ДНК методом – в виде строки (хромосомы), содержащей определенный набор генов. Длина хромосомы постоянна, а популяция, состоящая из некоторого количества особей, подвергается процессу эволюции с использованием операций скрещивания и мутаций. Классический ГА содержит следующие шаги:

1. Создание начальной популяции.
 - 1.1. Инициализация хромосомы каждой особи.
 - 1.2. Оценивание начальной популяции.
2. Этап эволюции – построение нового поколения.
 - 2.1. Отбор кандидатов на скрещивание (селекция).
 - 2.2 Скрещивание, т.е. порождение каждой парой отобранных кандидатов новых индивидов.
 - 2.3. Мутация.
 - 2.4. Оценивание новой популяции.

В начале работы алгоритма случайным образом инициализируется популяция P_0 , состоящая из S особей (РБС сетей): $P_0 = \{H_1, H_2, \dots, H_S\}$. Каждая особь в популяции при этом получает свое уникальное описание, закодированное в хромосоме $H_j = \{h_{1j}, h_{2j}, \dots, h_{Lj}\}$, которая состоит из L генов, где $h_{ij} \in [h_{\min}, h_{\max}]$ – значение i -го гена j -й хромосомы (h_{\min} – минимальное, и h_{\max} – максимальное допустимые значения). Формат хромосомы и соответствие между генами и параметрами РБС представлены на рис. 1.

Как видно из рисунка, каждая хромосома состоит из генов, в которых хранится информация о соответствующих параметрах сети. В начале хромосомы находятся гены, которые хранят информацию о параметрах помехи, затем идет ген, кодирующий параметр a_0 . Параметры нейронов скрытого слоя закодированы в последующих блоках генов. Так первый ген каждого такого блока (1/0) определяет, присутствует ли соответствующий нейрон в структуре сети, т.е. участвует он или нет в вычислении выходной реакции сети на поступивший входной сигнал. Ген BF определяет, какая базисная функция из заданного набора БФ будет использоваться для вычисления реакции нейрона.

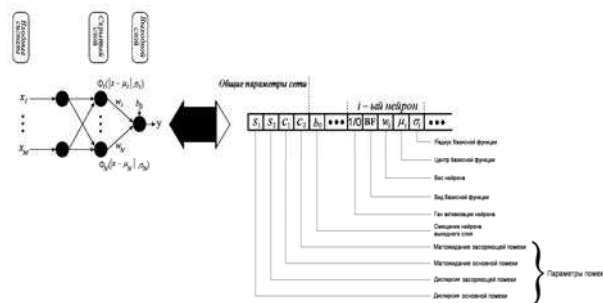


Рис. 1. Формат хромосомы и соответствие между генами и параметрами РБС

Оценивание популяции. После формирования начальной популяции оценивается приспособленность каждой особи, входящей в нее, на основе анализа функции приспособленности (фитнесс-функции)

$$f_i(x_j) = \frac{1}{M} \sum_{j=1}^M \left(y_j^*(x_j) - \hat{y}_j(x_j) \right)^2, \quad (3)$$

где $y^*(k)$ – желаемая реакция сети; $\hat{y}(k)$ – реальный выходной сигнал; M – размер выборки.

Скрещивание. После того как родительские особи были отобраны методом селекции, осуществляется их скрещивание (кроссовер). Кроссовер применяется с целью воспроизведения потомства, и заключается в обмене генетической информацией между родительскими особями. Пусть родительские особи описываются выражениями

$$H^{(1)} = \{h_1^{(1)}, \dots, h_i^{(1)}, \dots, h_L^{(1)}\};$$

$$H^{(2)} = \{h_1^{(2)}, \dots, h_i^{(2)}, \dots, h_L^{(2)}\},$$

тогда хромосомы их потомков принимают вид

$$Y^{(1)} = \{h_1^{(1)}, \dots, h_i^{(1)}, h_{i+1}^{(2)}, \dots, h_L^{(2)}\},$$

$$Y^{(2)} = \{h_1^{(2)}, \dots, h_i^{(2)}, h_{i+1}^{(1)}, \dots, h_L^{(1)}\},$$

где i – случайная величина из интервала $[1, L]$. Пример операции скрещивания двух РБС показан на рис. 2. Из рисунка видно, что при скрещивании происходит не только обмен информацией о параметрах сети, но и образуются сети с новой для популяции структурой.

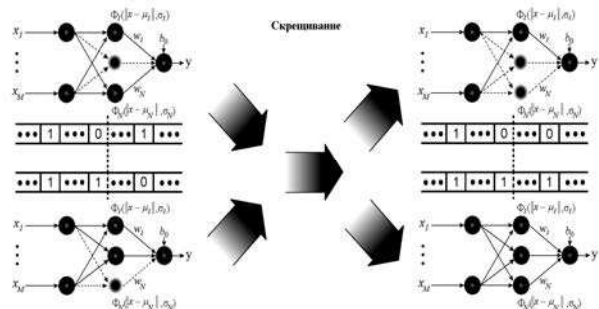


Рис. 2. Пример операции скрещивания двух РБС

Оператор мутации. Мутация представляет собой генетический оператор, который изменяет одно или несколько значений генов в хромосоме. Следует отметить, что в ГА механизм мутаций является единственным способом внесения новой информации в хромосому особи. Это может привести к совершенно новым значениям генов, которые впоследствии могут быть добавлены в генофонд популяции. С помощью этих новых значений генов ГА получает возможность найти лучшие решения. Мутация является важной частью генетического поиска и помогает предотвратить застревание популяции в локальных минимумах. Она происходит в хромосомах особей, полученных на этапе скрещивания с некоторой обычно заранее заданной вероятностью. Вероятность мутации также может быть привязана к значению фитнесс-функции, т.е. чем хуже значение фитнесс-функции, тем выше вероятность мутации.

Наиболее часто в ГА используются следующие два способа мутации генов. В первом случае значение случайно выбранного гена хромосомы потомка заменяется на новое значение в диапазоне его допустимых значений

$$h_{ij} = \text{rand}[h_{\min}, h_{\max}], \quad (4)$$

где $\text{rand}[x, y]$ – случайное число в интервале $[x, y]$, распределенное по равномерному закону.

Во втором же случае к существующему значению гена прибавляется некоторое случайное смещение

$$h'_{ij} = h_{ij} + \Delta h_{ij}, \quad (5)$$

где h_{ij} – ген до мутации; h'_{ij} – ген после мутации.

Однако при настройке РБС с помощью данных типов мутаций процесс обучения затягивается, так как настраиваемые параметры (вес связи, центр и дисперсия базисной функции), как уже отмечалось, разнородны и их значения лежат в различных диапазонах. В связи с этим возникает необходимость применения адаптивных мутаций, которые позволяют определить параметры мутации для каждого гена отдельно. При этом параметры мутаций хранятся в дополнительном векторе v .

Рассмотрим основные типы существующих адаптивных мутаций.

Уменьшающаяся мутация Коши. При использовании данного подхода изменение шага мутации происходит следующим образом:

$$v_j(k) = \gamma v_j(k-1), \quad (6)$$

где $\gamma \in [0, 1]$ – некоторая константа (в работе [2] предлагается использовать $\gamma = 0.95$).

Изменение же значения гена происходит по следующему правилу:

$$h'_{ij} = h_{ij} + v_j(k)C(k, \tau), \quad (7)$$

где $C(k, t)$ – случайная величина, распределенная по закону Коши с параметром τ . Плотность вероятности данной случайной величины описывается следующим выражением:

$$f(x) = \frac{1}{\pi} \left(\frac{\tau}{x^2 + \tau^2} \right). \quad (8)$$

Адаптивная мутация Гаусса. Данный тип мутации был предложен в работе [3]. Здесь также вначале происходит коррекция шага мутации, а затем производится непосредственно изменение значения мутирующего в хромосоме гена. Эту процедуру можно описать следующим образом:

$$v_j(k) = v_j(k-1) \exp[\text{LN}(0, 1) + I'N(0, 1)]; \quad (9)$$

$$h'_{ij} = h_{ij} + v_j(k)N(0, 1), \quad (10)$$

где $N(0, 1)$ – случайная величина, распределенная по нормальному закону с нулевым математическим ожиданием и единичной дисперсией; 1 и I' – некоторые параметры, для которых в работе [4] были

получены следующие оптимальные значения:

$$1 = (2N)^{-0.5} \text{ и } I' = (2\sqrt{N})^{-0.5} \text{ соответственно.}$$

Адаптивная мутация Лапласа. Данный тип мутации в последнее время все чаще используется в качестве альтернативы гауссовской мутации. Особенно популярна мутация Лапласа в задачах минимизации функции со множеством локальных минимумов. Настройка шага мутации осуществляется по правилу (9), т.е. как и при использовании адаптивной мутации Гаусса, мутация же гена осуществляется с помощью уравнения

$$h'_{ij} = h_{ij} + v_j(k)L(\alpha), \quad (11)$$

где L – случайная величина, распределенная по закону Лапласа с параметром α , плотность вероятности которой имеет вид:

$$f(x) = \begin{cases} \frac{1}{2} e^{-\alpha x}, & \text{если } x \leq 0; \\ 1 - \frac{1}{2} e^{-\alpha x}, & \text{если } x > 0. \end{cases} \quad (12)$$

На рис. 3 представлены плотности вероятности распределения Лапласа с $\alpha = 2$ (пунктирная линия), Гаусса с единичной дисперсией (сплошная линия), Коши с $\tau = 1$ (линия с кружками).



Рис. 3. Плотности вероятности распределения

Моделирование

Решалась задача аппроксимации функции

$$f[x_1, x_2] = \frac{x_1 x_2 [x_1 + 2.5]}{1 + x_1^2 + x_2^2} \quad (13)$$

с помощью эволюционной РБС с использованием различных типов мутаций. Популяция состояла из 64 особей (РБС), максимально допустимое количество нейронов в каждой сети было ограничено 15-ю. На рис. 4 показана поверхность, описываемая уравнением (13). Результаты аппроксимации функции (13) с помощью ГА с адаптивной мутацией Лапласа приведены на рис. 5. На рис. 6 показано количество нейронов в скрытом слое сети-победителя (сети с наименьшим значением фитнес-функции) на каждой итерации ГА. Изменение же значения фитнес-функции (3) в процессе эволюции приведено на рис. 7. Здесь линией с кружками показаны результаты работы ГА при использовании мутации (4), при использовании

уменьшающейся мутация Коши (6) – (7) – пунктирной линией, а при использовании адаптивной мутации Лапласа (11) – сплошной жирной. Как видно из рисунка, наиболее точные результаты дает применение адаптивной мутации Лапласа, однако на начальных этапах работы алгоритма мутация Коши показывает лучшие результаты. Равномерная же мутация (4) показывает худшие результаты на протяжении всего времени работы ГА.

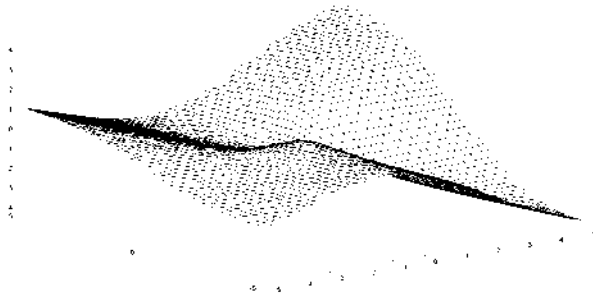


Рис. 4. Поверхность, описываемая уравнением (13)

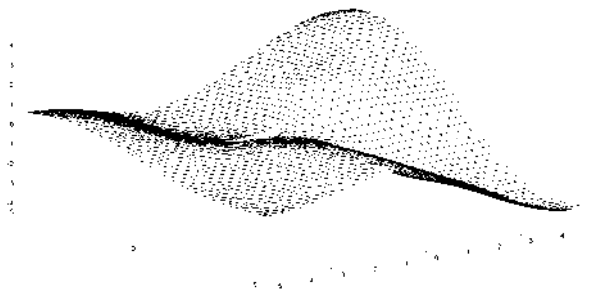


Рис. 5. Результаты аппроксимации функции (13) с помощью ГА с адаптивной мутацией Лапласа



Рис. 6. Количество нейронов в скрытом слое сети-победителя на каждой итерации ГА

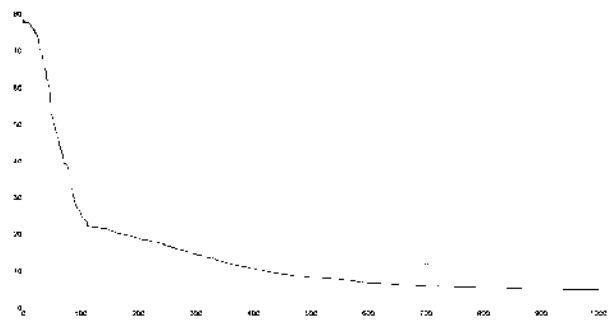


Рис. 7. Изменение значения фитнес-функции в процессе эволюции

Выводы

Выбор структуры РБС является достаточно сложной задачей, которую можно эффективно решить с помощью ГА. Основным способом настройки параметров сети и уточнения ее структуры при этом является применение операции мутации. Применение равномерной мутации (4) затягивает процесс обучения, в связи с чем возникает необходимость в адаптивных процедурах определения шага мутации для каждого гена индивидуально. Как показывают результаты моделирования, мутации Коши и Лапласа позволяют получить несколько лучшие результаты.

Список литературы

1. Maillard E.P. RBF neural network, basis functions and genetic algorithm / E.P. Maillard, D. Gueriot // Proc. Int. Conf. Neural Networks, Houston, TX. – 1997. – V.4. – P. 2187-2192.
2. Yang J.-M. A genetic algorithm with adaptive mutations and family competition for training neural networks / J.-M. Yang, J.-T. Horng, C.-Y. Kao // International Journal of Neural Systems. – 2000. – 10(5). – P. 333-352.
3. Schwefel H.-P. Numerical Optimization of Computer Models / H.-P. Schwefel. – John Wiley & Sons, Chichester, UK, 1981.
4. Bäck T. An overview of evolution algorithms for parameter optimization / T. Bäck, H.-P. Schwefel // Evolutionary Computation. – 1993. – 1(1). – P. 1-23.

Поступила в редколлегию 23.02.2012

Рецензент: д-р техн. наук, проф. О.Г. Руденко, Харьковский национальный университет радиоэлектроники, Харьков.

НАВЧАННЯ РАДІАЛЬНО-БАЗИСНИХ МЕРЕЖ ЗА ДОПОМОГОЮ ГЕНЕТИЧНИХ АЛГОРИТМІВ З АДАПТИВНОЮ МУТАЦІЄЮ

О.О. Безсонов

Розглядається ефективність застосування адаптивних мутацій в генетичних алгоритмах, що використовуються для навчання радіально-базисної мережі та оптимізації її структури. Досліджуються такі типи адаптивних мутацій: мутація Коші, що зменшується, адаптивні мутації Гаусса і Лапласа. Проводиться їх порівняльний аналіз за допомогою імітаційного моделювання.

Ключові слова: нейронна мережа, генетичний алгоритм, адаптивна мутація, базисна функція, селекція, фитнес-функція, хромосома, ген.

TRAINING OF RADIAL BASIS NETWORKS BY GENETIC ALGORITHMS WITH ADAPTIVE MUTATIONS

O.O. Bezsonov

In this paper the efficiency of adaptive mutations in the genetic algorithms used for training radial basis network and optimisation of its structure is considered. We investigate the following types of adaptive mutations: decreasing Cauchy mutation, adaptive Gauss and Laplace mutations. A comparative analysis using simulation is conducted.

Keywords: neural network, genetic algorithm, adaptive mutation, basis function, selection, fitness function, chromosome, gene.