

УДК: 004.8:004.94

## МЕТОД ПРЕДВАРИТЕЛЬНОГО ВЫБОРА АРХИТЕКТУРЫ МНОГОСЛОЙНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ АППРОКСИМАЦИИ ПОЛИГАРМОНИЧЕСКОГО СИГНАЛА

**В. В. Фролов**, д.т.н., доцент, профессор, Харьковский национальный университет им. В. Н. Каразина, Харьков, Украина;

orcid.org/0000-0002-2770-3385, e-mail: vvicfrol@rambler.ru

**Г. Н. Жолткевич**, д.т.н., профессор, Харьковский национальный университет им. В. Н. Каразина, Харьков, Украина;

orcid.org/0000-0002-7515-2143, e-mail: g.zholtkevych@gmail.com

**О. Ю. Приходько**, к.т.н., доцент, БГТУ им. В. Г. Шухова, Белгород, Россия;

orcid.org/0000-0002-6452-0465, e-mail: prihodko.o.u@gmail.com

**Ю. В. Скурятин**, к.т.н., доцент, БГТУ им. В. Г. Шухова, Белгород, Россия;

orcid.org/0000-0001-5555-8691, e-mail: yvs-work@mail.ru

*Предлагаемый в статье метод предварительного выбора архитектуры сужающейся многослойной искусственной нейронной сети прямого распространения без обратных связей с сигмоидными функциями активации основан на определении количества слоев по форме полигармонического сигнала. Форму сигнала предлагается характеризовать количеством точек перегиба. В статье экспериментально доказано, что количество слоев многослойной сети коррелирует с количеством точек перегиба по критерию минимизации абсолютной погрешности при сравнении с универсальным аппроксиматором. Общее количество нейронов многослойной сети определяется из условия, что оно не может быть больше количества нейронов универсального аппроксиматора. Суть метода заключается в сравнительном анализе абсолютной погрешности для однослойной и многослойной сети. Полученную конфигурацию многослойной сети можно использовать как начальную для дальнейшей оптимизации структуры.*

**Ключевые слова:** искусственная нейронная сеть, функция активации, аппроксимация, генетический алгоритм, нейронная сеть прямого распространения без обратных связей, дискретная оптимизация, медиана, среднеквадратичное, абсолютная погрешность.

**DOI:** 10.21667/1995-4565-2021-78-102-111

### Введение

Использование искусственных нейронных сетей (ИНС) для решения задач аппроксимации зависимостей в современном анализе данных достаточно широко распространено. Теоретической основой применения ИНС для задач этого класса является теорема об универсальной аппроксимации [1]. В рамках этой теоремы Дж. Цибенко [2] доказал, что для аппроксимации достаточно использовать один скрытый слой нейронов. Современные подходы обеспечивают проектирование архитектуры многослойных ИНС: методами наращивания и упрощения [3], процедурной генерации архитектур модульных ИНС [4], методами алгоритмического формирования ИНС, основанными на анализе предметной области [5, 6], на основе глубоких нейронных сетей [7, 8].

Все вышеприведенное показывает, что конструирование архитектуры многослойных ИНС является сложной и многовариантной задачей. Эта задача на данный момент слабо формализована. Эффективность ее решения целиком зависит от опыта проектировщика. При этом методы её решения можно разбить на несколько групп.

Первая группа методов предполагает усложнение ИНС. Вначале принимается, что ИНС должна быть максимально простой с точки зрения конкретного метода. Затем на каждом последующем шаге сеть усложняется добавлением слоев и нейронов. Алгоритмы, которые ис-

пользуются для этого, характеризуют конкретный метод. Эти действия продолжают до тех пор, пока погрешность обучения сети не будет приемлема для проектировщика. Характерный пример такого подхода с разработкой типовых шаблонов и алгоритма эволюции сети предложен в работе [9]. Здесь процесс эволюции основан на правилах самоорганизации сети из типовых шаблонов и предполагает определённый набор шагов. Эти шаги выполняются циклически до тех пор, пока интегрированная ошибка обучения нейронной сети не будет меньше пороговой величины, при которой сеть считается обученной [9].

Вторая группа методов предполагает упрощение ИНС. На начальном этапе формируется максимально сложная ИНС. Каждый последующий шаг предполагает удаление нейронов и слоев по заданному критерию отбора. Алгоритмы и критерии отбора определены конкретными методами. В результате нужно получить максимально простую ИНС, которая должна аппроксимировать данные в рамках погрешностей, заданных проектировщиком. Так, например, в работе [10] авторы выделяют: алгоритм динамического добавления и удаления, метод «оптимального повреждения мозга».

Здесь следует отметить, что определение максимальной сложности ИНС в терминах конкретного метода – задача не простая, поскольку не существует строго формализованных подходов для ее решения.

Третья группа методов предполагает перебор возможных архитектур ИНС до тех пор, пока не будет найдена архитектура, описывающая заданный набор данных с точностью, достаточной для решения практических задач. К этой группе можно отнести методы, использующие генетические алгоритмы для формирования топологии сети, представленные в работах [11-15]. Анализ работ, реализующих третью группу методов, показывает, что обычно с помощью генетических алгоритмов решается либо задача выбора структуры сети, либо задача обучения сети. При выборе топологии сети реализуются два подхода – прямой и косвенный. Первый подход предполагает перебор возможных межнейронных связей, а второй перебор правил формирующих эти связи.

### Постановка задачи

Данные направления объединяет один аспект – наличие архитектуры ИНС, относительно которой выполняются дальнейшие шаги по ее трансформации для поиска оптимальной конфигурации с точки зрения того или иного метода. Поэтому актуальным для этих методов будет обоснование задания такой начальной конфигурации сети для конкретных условий. При этом дальнейшие действия будут определяться экономической целесообразностью применения того или иного метода структурно-параметрической оптимизации архитектуры ИНС.

В связи с этим цель данной статьи – разработка метода выбора архитектуры сужающейся многослойной ИНС прямого распространения без обратных связей с сигмоидальными функциями активации скрытых слоев для задач аппроксимации полигармонических сигналов электронной аппаратуры.

Актуальность выбора архитектуры многослойной нейронной сети обусловлена развитием аппаратных средств и программного обеспечения, которые позволяют решать задачи структурной оптимизации по выбору приемлемой архитектуры нейронной сети для решения конкретных задач за время, удовлетворяющее минимуму соотношения затрат на оптимизацию и эксплуатацию обученных нейронных сетей. В противном случае нет смысла решать задачу, поскольку эти затраты не будут компенсироваться эффективностью использования полученных нейронных сетей.

Задачи аппроксимации сигналов с использованием ИНС обычно решаются с помощью либо радиально-базисных нейронных сетей, либо «универсального аппроксиматора» (УА), который представляет нейронную сеть с одним скрытым слоем и одним выходным слоем. Выходной слой представлен одним нейроном с линейной функцией активации, который обеспечивает суперпозицию выходов нейронов скрытого слоя. В скрытом слое обычно используются нейроны с нелинейными функциями активации, в частности, может быть использо-

вана смещенная сигмоида. В этом случае, возможность решения задачи аппроксимации, может быть теоретически обоснована теоремой Колмогорова о суперпозиции функций.

По данным работы [16], использование многослойных ИНС без обратных связей с сигмоидными функциями активации скрытых слоев для аппроксимации сигналов позволяет повысить точность аппроксимации по сравнению с УА при одном и том же количестве нейронов. Но при этом теоретическое обоснование возможности аппроксимации на основе теоремы Колмогорова не совсем корректно, поскольку мы имеем не суперпозицию функций, а композицию функций и предсказать, какая композиция будет лучшей для конкретного сигнала довольно проблематично.

Поскольку корректное теоретическое обоснование выбора архитектуры многослойных нейронных сетей в современных условиях отсутствует, предложим решение этой задачи на основе эмпирических данных по результатам серии численных экспериментов.

### Экспериментальные исследования

В экспериментальной постановке задача выбора архитектуры ИНС может быть сформулирована так: определить параметры сигнала, оказывающие доминирующее влияние на количество скрытых слоев сужающейся ИНС прямого распространения без обратных связей, инвариантные к форме полигармонического сигнала.

Ограничимся рассмотрением полигармонических сигналов с количеством гармоник от 2 до 5. Начнем с рассмотрения сигнала (см. рисунок 1), имеющего 3 гармоники: одну несущую и две наложенные. Диапазон, в котором будем рассматривать форму сигнала, от 0 до 1. На рисунке 1 видно, что сигнал периодический с периодом, равным 0,4. Форму сигнала здесь можно однозначно характеризовать точками перегиба. По второй производной (см. рисунок 2) сигнала определяем, что точек перегиба в рамках периода будет 4. Проведем серию численных экспериментов, чтобы определить связь количества точек перегиба и количества слоев ИНС для аппроксимации сигнала. Поскольку ИНС сужающаяся, необходимо задать закон убывания количества нейронов в скрытых слоях. Учитывая положения теоремы Колмогорова, определяем, что в слое, который будет связан с выходным слоем, число нейронов должно быть не меньше числа гармоник полигармонического сигнала. Это вытекает из того, что выходной слой состоит из одного нейрона, который, по сути, обеспечивает суперпозицию функций последнего скрытого слоя, и каждый нейрон этого слоя тогда должен отвечать за одну гармонику. Общее количество нейронов многослойной ИНС будет зависеть от количества нейронов УА. Последнее вытекает из того, что при одном и том же количестве нейронов многослойная ИНС обеспечивает меньшую погрешность аппроксимации.

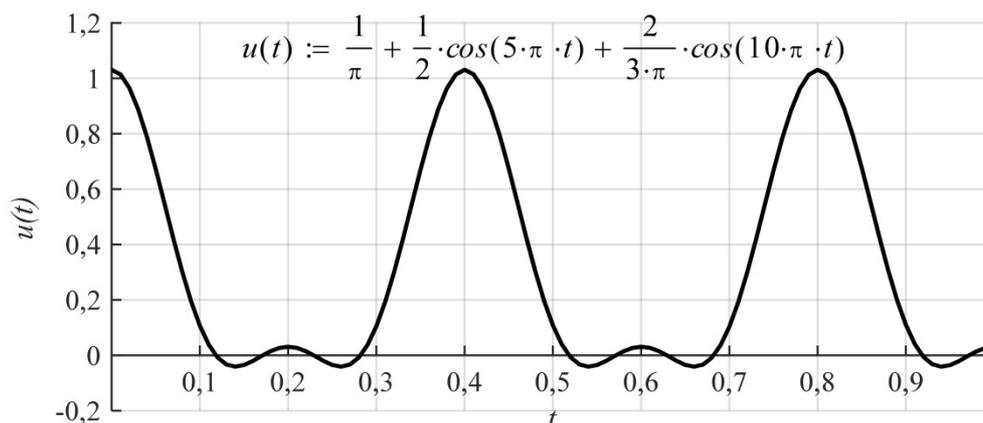
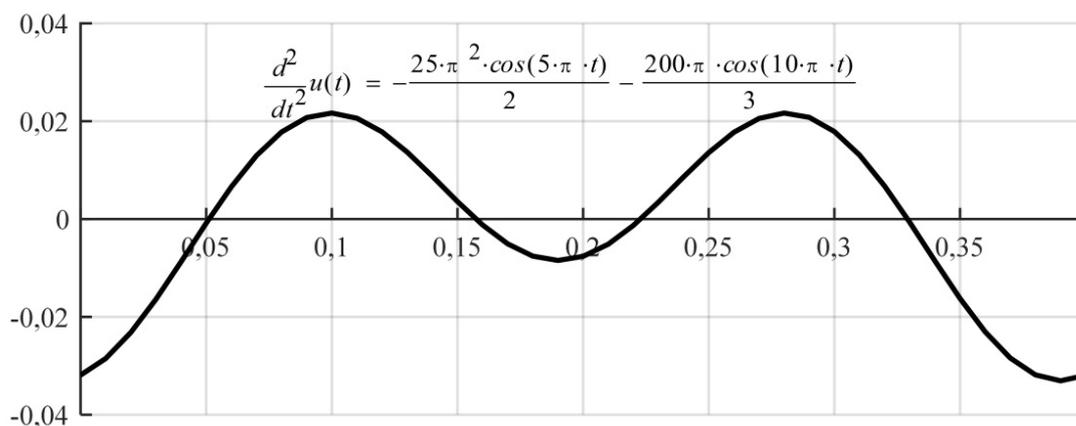


Рисунок 1 – Полигармонический сигнал с тремя гармониками  
Figure 1 – Polyharmonic signal with three harmonics



**Рисунок 2 – Определение точек перегиба по второй производной**  
**Figure 2 – Determination of inflection points from the second derivative**

В результате имеем, на примере сигнала на рисунке 1, количество нейронов предпоследнего слоя не меньше трех. Общее количество нейронов сети определяем по количеству нейронов УА путем подбора при минимизации суммарной абсолютной ошибки сети (1):

$$E^{\Sigma} = \sum_{i=1}^n |y_i - y_i^{net}|, \quad (1)$$

где  $E^{\Sigma}$  – суммарная абсолютная ошибка ИНС;  $n$  – количество значений сигнала после квантования, для случая на рисунке 1 принимаем 201 значение в диапазоне от 0 до 1;  $y_i$  – значения сигнала;  $y_i^{net}$  – значение сигнала по обученной ИНС.

Задачу подбора по минимуму (1) решаем в системе Matlab, последовательно обучая сеть newff (feed-forward backpropagation network) с одним скрытым слоем, содержащим нейроны со смещенной сигмоидальной функцией активации, и одним нейроном выходного слоя с линейной функцией активации. При обучении ИНС используем метод Levenberg-Marquardt backpropagation как наиболее точный в смысле аппроксимации данных. Максимальное количество нейронов в ИНС должно быть не более 200, поэтому будем последовательно изменять количество нейронов скрытого слоя от 20 до 70 и рассчитывать суммарную абсолютную ошибку. При обучении сети начальные веса выбираются случайным образом, поэтому, чтобы получить корректные результаты, проводим серию из 10 экспериментов и определяем тенденцию изменения  $E^{\Sigma}$  по усредненным данным (см. рисунок 3).

Анализ рисунка 3, *a* и *б* показывает, что после 40 нейронов абсолютная суммарная погрешность аппроксимации увеличивается. Таким образом, имеем ИНС прямого распространения без обратных связей с одним скрытым слоем из 40 нейронов и выходным слоем из одного нейрона для аппроксимации сигнала на рисунке 3, которая определяет общее количество нейронов многослойной сети. Следующим этапом необходимо распределить общее количество нейронов по слоям с тенденцией уменьшения, поскольку ИНС сужающаяся.

Согласно работе [17] для многослойных ИНС наиболее эффективно уменьшение количества нейронов по слоям в два раза, что подчиняется экспоненциальному закону с показателем степени, приблизительно равной 0,7 (0,691 по результатам аппроксимации). В этом случае для расчета количества нейронов в каждом слое ИНС решаем задачу оптимизации в следующей постановке: выбрать распределение нейронов по слоям ИНС, подчиняющееся экспоненциальному закону вида  $y = a \cdot e^{b \cdot x}$ , где коэффициент  $b$  должен быть максимально близким к 0,7. При этом необходимо учитывать такие ограничения: общее количество нейронов ИНС должно быть равно количеству нейронов УА, выбранному на основе минимизации (1); количество слоев ИНС должно быть равно целому наперед заданному числу; в предпоследнем слое количество нейронов не должно быть меньше количества гармоник сигнала; число нейронов каждого последующего слоя должно быть меньше предыдущего. Тогда имеем за-

дачу дискретной оптимизации, которую возможно решать с помощью ГА. Здесь хромосома представляет собой набор генов, в каждом из которых задается число нейронов отдельного слоя ИНС, а количество генов в хромосоме равно количеству слоев ИНС. Математическая модель этой задачи отражена в (2):

$$fs = ||b| - 0.7| \rightarrow \min, b = f(\bar{x})$$

$$\begin{cases} \sum_{i=1}^g x_i = N \\ x_g \geq h \\ x_i > x_{i+1}, \forall i \in I, I = \overline{1, g} \end{cases}, \quad (2)$$

где  $g$  – количество слоев ИНС;  $N$  – общее количество нейронов в ИНС;  $f(\bar{x})$  – функция, возвращающая результаты аппроксимации экспоненциальной зависимостью;  $\bar{x}$  – вектор, содержащий значения чисел нейронов в каждом слое ИНС;  $h$  – число гармоник в сигнале;  $fs$  – целевая функция.

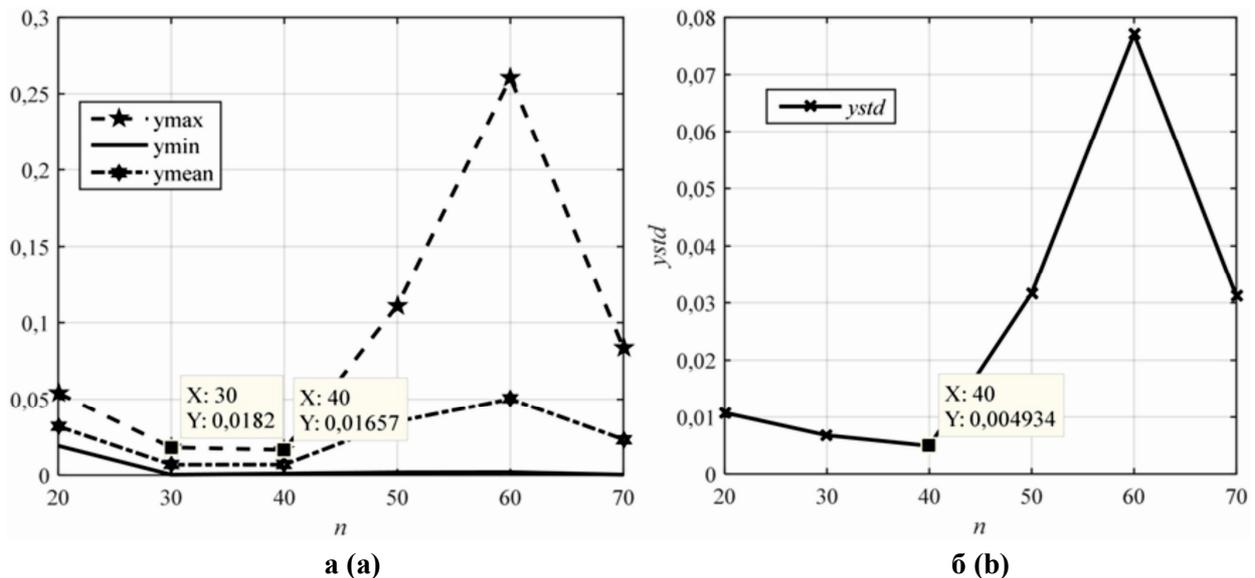


Рисунок 3 – Изменение  $E^Z$  при увеличении  $n$  количества нейронов в скрытом слое по 10 экспериментам: а –  $y_{max}$  – максимальное,  $y_{min}$  – минимальное,  $y_{mean}$  – среднее значение  $E^Z$ ; б –  $y_{std}$  – среднее квадратическое отклонение значений  $E^Z$

Figure 3 – Change  $E^Z$  with an increase in neuron number in a hidden layer for 10 experiments: а –  $y_{max}$  – maximum,  $y_{min}$  – minimum,  $y_{mean}$  – average value  $E^Z$ ; б –  $y_{std}$  – standard deviation of values  $E^Z$

Оптимизационную задачу решаем в Matlab, используя функции: fit – аппроксимация экспоненциальной зависимостью (exp1) с использованием метода Trust-region; ga – ГА с целыми значениями генов и ограничениями из (2), оформленными в виде отдельной функции.

В результате, для сигнала на рисунке 1, чтобы выполнить серию экспериментов, сформированы на основе (2) такие данные для 5 архитектур ИНС: [27 13 1]; [23 11 6 1]; [21 12 4 3 1]; [20 8 5 4 3 1]; [15 7 6 5 4 3 1]. В серии экспериментов сравниваются УА и многослойная ИНС с одним и тем же количеством нейронов по критерию (1).

Последовательность выполнения единичного эксперимента.

1. Обучаем УА и вычисляем модуль абсолютной ошибки приближения по каждому замеру сигнала  $one = \{one_i : one_i = |y_i - y_i^{ONET}|, i = \overline{1, m}\}$ , где  $y_i^{ONET}$  – результат ИНС.

2. Вычисляем среднее ошибки  $\overline{one} = \frac{1}{m} \cdot \sum_{i=1}^m one_i$

3. Выполняем п. 1 для многослойной ИНС  $multi = \{multi_i : multi_i = |y_i - y_i^{MNET}|, i = \overline{1, m}\}$ , где  $y_i^{MNET}$  – результат ИНС.

4. Вычисляем среднее ошибки  $\overline{multi} = \frac{1}{m} \cdot \sum_{i=1}^m multi_i$ .

5. Вычисляем коэффициент  $Km = \frac{\overline{one}}{\overline{multi}}$ , показывающий во сколько раз средняя ошибка

УА больше чем многослойной ИНС.

При обучении ИНС, начальные веса синаптических коэффициентов выбираются случайным образом, поэтому для оценки усредненной тенденции изменения  $\overline{Km}$  проведем серию из 200 экспериментов, для каждой пары ИНС. При этом необходимо учесть, что могут быть аномальные результаты (выбросы), которые выпадают из общей тенденции. Удалять выбросы будем в том случае, если выполняется условие

$$|Km_i - Me| \geq 3 \cdot \sigma, i = \overline{1, 200}, \quad (3)$$

где  $Me$  – медиана по всем экспериментам;  $\sigma$  – среднеквадратичное отклонение.

Таким образом, общий алгоритм численного эксперимента будет.

1. Для каждой из пяти архитектур выполнить алгоритм единичного эксперимента, получив матрицу из  $Km$  размерностью  $200 \times 5$ .

2. Выполнить фильтрацию данных по выбросам на основе (3) для каждого столбца матрицы.

3. Вычислить  $\overline{Km}$  по каждому столбцу матрицы без учета выбросов.

4. Отобразить тенденцию изменения коэффициента на графике, где по оси ординат откладываем усредненное значение  $\overline{Km}$  для конкретной конфигурации ИНС, а по оси абсцисс – номер архитектуры ИНС, совпадающий с количеством скрытых слоев в ней.

Выполнив указанный алгоритм для обработки сигнала на рисунке 1, получаем тенденцию изменения  $\overline{Km}$  в зависимости от архитектуры ИНС от 2 до 6 слоев на рисунке 4. По результатам этой серии экспериментов можно предположить, что наиболее оптимальная архитектура ИНС будет [21 12 4 3 1], поскольку для нее коэффициент  $\overline{Km}$  имеет максимальное значение. Причем количество скрытых слоев совпадает с количеством точек перегиба на рисунке 2.

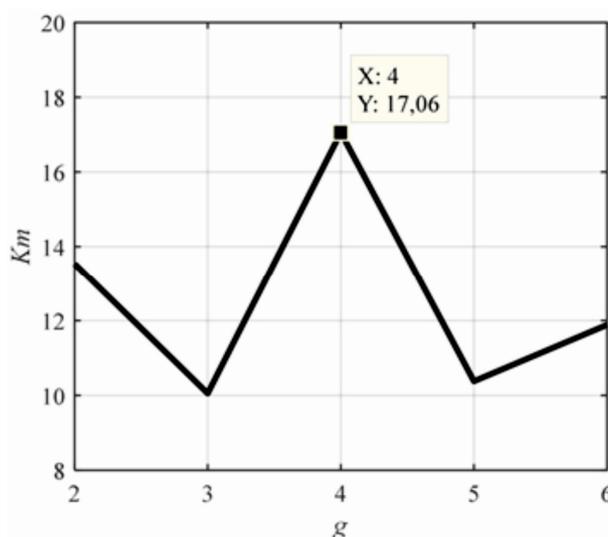


Рисунок 4 – Результаты численных экспериментов для сигнала на рисунке 1:  
 $\overline{Km}$  – среднее значение коэффициента по 200 экспериментам;  $g$  – число слоев ИНС  
 Figure 4 – The results of numerical experiments for the signal in figure 1:  
 $\overline{Km}$  – average value of the coefficient over 200 experiments;  $g$  – number of ANN layers

Проведенные серии таких экспериментов (см. рисунки 5, 6) для различных сигналов, подтверждают гипотезу, что количество слоев сужающейся многослойной ИНС прямого распространения без обратных связей с сигмоидальными функциями активации, используемой для аппроксимации сигнала, приближенно может быть равно числу точек перегиба сигнала.

### Заключение

Метод выбора архитектуры многослойной ИНС прямого распространения без обратных связей с сигмоидальными функциями активации для целей аппроксимации заключается в следующем.

1. Определение количества гармоник исходного полигармонического сигнала.
2. Нахождение точек перегиба сигнала.
3. Расчет оптимального количества нейронов для УА.

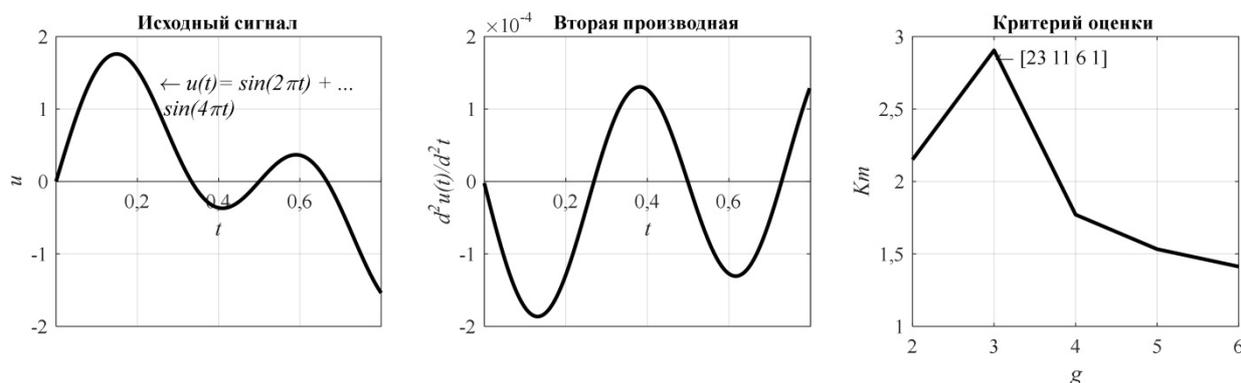


Рисунок 5 – Серия экспериментов для сигнала с двумя гармониками и тремя точками перегиба  
Figure 5 – Series of experiments for a signal with two harmonics and three inflection points



Рисунок 6 – Серия экспериментов для сигнала с пятью гармониками и двумя точками перегиба  
Figure 6 – Series of experiments for a signal with five harmonics and two inflection points

4. Формирование архитектуры ИНС на основании следующего: количество скрытых слоев должно быть равно количеству точек перегиба; общее количество нейронов скрытых слоев должно быть равно количеству нейронов из п. 3; распределение нейронов между слоями должно подчиняться экспоненциальному закону с показателем степени, близким к 0,7; последний скрытый слой должен иметь количество нейронов, не меньшее, чем количество гармоник сигнала.

Предложенный в данной работе метод, позволяет на ранних этапах проектирования ИНС задать архитектуру, потенциально содержащую лучшее решение, и в зависимости от конкретных условий, либо выполнять дальнейшую оптимизацию, либо принять это решение за конечное с определенной долей вероятности.

## Библиографический список

1. **Хайкин С.** Нейронные сети: полный курс / С. Хайкин. 2-е изд.: пер. с англ. М.: Издательский дом «Вильямс», 2006. 1104 с.
2. **Cybenko G.** Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems*. 1989, no. 2, pp.303-314. DOI 10.1007/BF02551274.
3. **Калистратов Т. А.** Методы и алгоритмы создания структуры нейронной сети в контексте универсальной аппроксимации функций. *Вестник Тамбовского университета. Серия: Естественные и технические науки*. 2014. Т. 19. № 6. С. 1845-1848. [eLIBRARY ID: 22902900].
4. **Лукьянова О. А.** Применение матричных фильтров и теории кос для процедурной генерации архитектур нейронных сетей. *Вычислительные технологии*. 2019. Т. 24. № 6. С. 69-78. [eLIBRARY ID: 41602939]. DOI 10.25743/ICT.2019.24.6.009.
5. **Воевода А. А.** Метод синтеза нейронной сети для аппроксимации поверхностей двух переменных / А. А. Воевода, Д. О. Романников, Г. В. Трошина. *Международная конференция по мягким вычислениям и измерениям*. 2018. Т. 1. С. 732-735. [eLIBRARY ID: 35619918].
6. **Винокуров И. В.** Интегрирование систем дифференциальных уравнений в нейросетевом логическом базисе. *Инновации и инвестиции*. 2017. № 11. С. 114-117. [eLIBRARY ID: 41505462].
7. **Созыкин А. В.** Обзор методов обучения глубоких нейронных сетей. *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*. 2017. Т. 6. № 3. С. 28-59 [eLIBRARY ID: 30016527]. DOI 10.14529/cmse170303.
8. **Николенко С.** Глубокое обучение. Погружение в мир нейронных сетей / С. Николенко, А. Кадурин, Е. Архангельская. СПб: Питер, 2018. 480 с.
9. **Миркин Е. Л.** Метод формирования топологии самоорганизующейся нейронной сети в процессе её обучения / Е. Л. Миркин, С. С. Нежинских // *Проблемы автоматизации и управления*. 2014. № 2(27). С. 28-36.
10. **Скуратовский Н. И.** Систематизация методов и алгоритмов структурно-параметрической оптимизации нейросетевых классификаторов в задачах распознавания / Н. И. Скуратовский, А. Ю. Темляков, М. Д. Алёхин // *Материалы VIII международной научной конференции «Системный анализ в медицине» (САМ 2014)* / под общ. ред. В. П. Колосова. Благовещенск: ФГБУ «ДНЦ ФПД» СО РАМН. 2014. С. 30-33.
11. **Шумков Е. А.** Генетическая топология поиска нейросетевых моделей / Е. А. Шумков, И. К. Чистик // *Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета*. 2013. № 93. С. 564-579. [eLIBRARY ID: 20929353].
12. **Мищенко В. А.** Использование генетических алгоритмов в обучении нейронных сетей / В. А. Мищенко, А. А. Коробкин // *Современные проблемы науки и образования*. 2011. № 6. С. 116. [eLIBRARY ID: 17388410].
13. **Королев Д. А.** Нейроэволюционный подход к оптимизации внутренней структуры нейронных сетей / Д. А. Королев, В. Г. Суфиянов // *Вестник Тихоокеанского государственного университета*. 2007. № 4(7). С. 107-122. [eLIBRARY ID: 11686604].
14. **Виноградова Е. Ю.** Принципы выбора оптимальной топологии нейронной сети для поддержки принятия управленческих решений. *Управленец*. 2012. № 7-8(35-36). С. 74-78. [eLIBRARY ID: 18026450].
15. **Хмелёв А. Г.** Эволюционные методы формирования нейросетевых моделей сложных экономических систем. *БИЗНЕСИНФОРМ*. 2014. № 1. С. 69-73.
16. **Ермаков Б. С.** Влияние структуры искусственной нейронной сети на ее эффективность на примере прогнозирования макроэкономических показателей. *Системный анализ и логистика: журнал*. выпуск № 2 (28), ISSN 2077-5687. СПб.: ГУАП., 2021. С. 44-51. [eLIBRARY ID: 45831386]. DOI: 10.31799/2077-5687-2021-2-44-51.
17. **Аль-Хайдри В. А.** Возможности применения искусственных нейронных сетей для обнаружения искажений ээг / В. А. Аль-Хайдри, Р. В. Исаков, Л. Т. Сушкова // *Биотехносфера*. 2016. № 1(43). С. 8-13. [eLIBRARY ID: 25993052].

UDC: 004.8: 004.94

## METHOD OF PRELIMINARY SELECTION OF MULTILAYER NEURAL NETWORK ARCHITECTURE FOR POLYHARMONIC SIGNAL APPROXIMATION

**V. V. Frolov**, Dr. Sc. (Tech.), associate professor, professor, V. N. Karazin Kharkiv National University, Kharkov, Ukraine;

orcid.org/0000-0002-2770-3385, e-mail: vviefrol@rambler.ru

**G. N. Zholtkevych**, Dr. Sc. (Tech.), full professor, V. N. Karazin Kharkiv National University, Kharkov, Ukraine;

orcid.org/0000-0002-7515-2143, e-mail: g.zholtkevych@gmail.com

**O. Yu. Prihodko**, Ph.D. (Tech.), associate professor, BSTU named after V. G. Shukhov, Belgorod, Russia;

orcid.org/0000-0002-6452-0465, e-mail: prihodko.o.u@gmail.com

**Yu. V. Skuryatin**, Ph.D. (Tech.), associate professor, BSTU named after V. G. Shukhov, Belgorod, Russia;

orcid.org/0000-0001-5555-8691, e-mail: yvs-work@mail.ru

*The proposed method for preliminary selection of the architecture for narrowing multi-layer artificial neural network of direct propagation without feedback with sigmoid activation functions is based on determining the number of layers by the shape of a polyharmonic signal. The authors have proposed to characterize signal shape by the number of inflection points. The article has experimentally proved that the number of layers in multilayer network correlates with the number of inflection points according to the criterion of minimizing absolute error when compared with a universal approximator. The total number of neurons in a multilayer network is determined from the condition that it cannot exceed the number of neurons in a universal approximator. The essence of the method lies in the comparative analysis of absolute error for single-layer and multi-layer network. Resulting configuration of multilayer network can be used as an initial one for further optimization of the structure.*

**Key words:** artificial neural network, activation function, approximation, genetic algorithm, feedforward neural network without feedbacks, discrete optimization, median, mean square, absolute error.

**DOI:** 10.21667/1995-4565-2021-78-102-111

### References

1. **Hajkin C.** *Nejronnye seti: polnyj kurs.* 2-e izdanie: per. s angl. M.: Izdatel'skij dom Vil'jams, 2006. 1104 p. (in Russian).
2. **Cybenko G.** Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems.* 1989, 2, pp. 303-314. DOI 10.1007/BF02551274.
3. **Kalistratov T. A.** Metody i algoritmy sozdaniya struktury nejronnoj seti v kontekste universal'noj ap-proksimacii funkcij. *Vestnik Tambovskogo universiteta. Serija: Estestvennye i tehnicheckie nauki.* 2014, vol. 19, no 6, pp. 1845-1848. [eLIBRARY ID: 22902900]. (in Russian).
4. **Luk'janova O. A.** Primenenie matrichnyh fil'trov i teorii kos dlja procedurnoj generacii arhitektur nejronnyh setej. O. A. Luk'janova, O. Ju. Nikitin, A. S. Kunin. *Vychislitel'nye tehnologii.* 2019, vol. 24, no. 6. pp. 69-78. [eLIBRARY ID: 41602939]. DOI 10.25743/ICT.2019.24.6.009. (in Russian).
5. **Voevoda A. A.** Metod sinteza nejronnoj seti dlja approksimacii poverhnostej dvuh peremennyh. A. A. Voevoda, D. O. Romannikov, G. V. Troshina. *Mezhdunarodnaja konferencija po mjagkim vychislenijam i izmerenijam.* 2018, vol. 1, pp. 732-735. [eLIBRARY ID: 35619918]. (in Russian).
6. **Vinokurov I. V.** Integrirovaniye sistem differencial'nyh uravnenij v nejrosetevom logicheskom bazise. *Innovacii i investicii.* 2017, no. 11, pp. 114-117. [eLIBRARY ID: 41505462]. (in Russian).
7. **Sozykin A. V.** Obzor metodov obucheniya glubokih nejronnyh setej. *Vestnik Juzhno-Ural'skogo gosudarstvennogo universiteta. Serija: Vychislitel'naja matematika i informatika.* 2017, vol. 6, no. 3, pp. 28-59 [eLIBRARY ID: 30016527]. DOI 10.14529/cmse170303. (in Russian).
8. **Nikolenko S.** *Glubokoe obuchenie. Pogruzhenie v mir nejronnyh setej.* S. Nikolenko, A. Kadurin, E. Arhangel'skaja. Sankt-Peterburg: Piter, 2018. 480 p. (in Russian).
9. **Mirkin E. L.** Metod formirovaniya topologii samoorganizujushhejsja nejronnoj seti v processe ejo obucheniya. E. L. Mirkin, S. S. Nezhinskih. *Problemy avtomatiki i upravleniya.* 2014, no. 2(27), pp. 28-36. (in Russian).

10. **Skuratovskij N. I.** Sistematizacija metodov i algoritmov strukturno-parametricheskoj optimizacii nejrosetevyh klassifikatorov v zadachah raspoznavanija. N. I. Skuratovskij, A. Ju. Temljakov, M. D. Aljohin. *Materialy VIII mezhdunarodnoj nauchnoj konferencii «Sistemnyj analiz v medicine» (SAM 2014)* / pod obshh. red. V. P. Kolosova. Blagoveshensk: FGBU «DNC FPD» SO RAMN. 2014, pp. 30-33. (in Russian).

11. **Shumkov E. A.** Geneticheskaja topologija poiska nejrosetevyh modelej. E. A. Shumkov, I. K. Chistik. *Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta*. 2013, no. 93, pp. 564-579. [eLIBRARY ID: 20929353]. (in Russian).

12. **Mishhenko V. A.** Ispol'zovanie geneticheskikh algoritmov v obuchenii nejronnyh setej. V. A. Mishhenko, A. A. Korobkin. *Sovremennye problemy nauki i obrazovanija*. 2011, no. 6. 116 p. [eLIBRARY ID: 17388410]. (in Russian).

13. **Korolev D. A.** Nejroevoljucionnyj podhod k optimizacii vnutrennej struktury nejronnyh setej. D. A. Korolev, V. G. Sufijanov. *Vestnik Tihookeanskogo gosudarstvennogo universiteta*. 2007, no. 4(7), pp. 107-122. [eLIBRARY ID: 11686604]. (in Russian).

14. **Vinogradova E. Ju.** Principy vybora optimal'noj topologii nejronnoj seti dlja podderzhki prinjatija upravlencheskih reshenij. *Upravlenec*. 2012, vol. 7-8(35-36), pp. 74-78. [eLIBRARY ID: 18026450]. (in Russian).

15. **Hmel'ov A. G.** Jevoljucionnye metody formirovanija nejrosetevyh modelej slozhnyh jekonomicheskikh system. *BIZNESINFORM*. 2014, no.1, pp. 69-73. (in Russian).

16. **Ermakov B. S.** Vlijanie struktury iskusstvennoj nejronnoj seti na ee jeffektivnost' na primere prognozirovaniya makrojekonomicheskikh pokazatelej. *Sistemnyj analiz i logistika: zhurnal: vypusk no. 2(28)*, ISSN 2077-5687. SPb.: GUAP., 2021, pp. 44-51. [eLIBRARY ID: 45831386]. DOI: 10.31799/2077-5687-2021-2-44-51. (in Russian).

17. **Al'-Hajdri V. A.** Vozmozhnosti primenenija iskusstvennyh nejronnyh setej dlja obnaruzhenija iskazhenij jekg. V. A. Al'-Hajdri, R. V. Isakov, L. T. Sushkova. *Biotehnosfera*. 2016, no. 1(43), pp. 8-13. [eLIBRARY ID: 25993052]. (in Russian).