

# ПРИМЕНЕНИЕ НЕЙРОННЫХ ЭХО-СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

## APPLICATION OF ECHO STATE NEURAL NETWORKS FOR TIME SERIES FORECASTING

**A. Shelonik**  
**A. Smirnov**  
**A. Kalistratov**  
**V. Koldobskiy**

*Summary.* The architecture of an echo state neural network is described. An echo state neural network is a recurrent neural network. An echo state network allows effective time series forecasting. Having low computational requirements, an echo state network shows results that are highly dependent on randomly generated initial values. A solution of this dependency problem is proposed. Usage of the multilayer architecture of the network and a genetic algorithm for acquiring network parameters is described. Echo state neural network's performance is compared with other networks' results.

*Keywords:* neural network, neural network architecture, time series forecasting.

**Шелоник Антон Александрович**  
 Магистрант, МГТУ им. Н.Э. Баумана  
 antonshelonik@gmail.com

**Смирнов Артём Сергеевич**  
 Магистрант, МГТУ им. Н.Э. Баумана  
 artem.smirnov@hotmail.com

**Калистратов Алексей Павлович**  
 Магистрант, МГТУ им. Н.Э. Баумана  
 akalistratov@gmail.com

**Колдобский Владислав Игоревич**  
 Магистрант, МГТУ им. Н.Э. Баумана  
 admin@nevergone.ru

*Аннотация.* В статье описана архитектура эхо-сетей, рекуррентных нейронных сетей. Нейронные эхо-сети (НЭС) позволяют эффективно предсказывать последовательность значений временного ряда. Обладая невысокой вычислительной сложностью на этапе обучения, НЭС показывает результаты, сильно зависящие от начальных значений, заданных случайным образом. Рассматривается методика устранения проблемы зависимости результатов. Предлагается использование многослойной архитектуры и генетического алгоритма для подбора гиперпараметров сети. Производится сравнение результатов работы НЭС с нейронными сетями другой архитектуры.

*Ключевые слова:* нейронная сеть, архитектура нейронных сетей, прогнозирование временных рядов.

## Введение

**Н**ейронные сети состоят из связанных между собой простых элементов, называемых нейронами. С помощью простых автоматов можно моделировать нейроны и связи между ними [1]. Применение нейронных сетей позволяет решать такие задачи, как фильтрация данных по различным признакам, построение временных рядов на основе отфильтрованных данных и их анализ [2].

Анализ временных рядов имеет широкое применение: от предсказания продаж до прогнозирования погоды. Построение моделей основывается на предположении, что более близкие по времени состояния системы связаны сильнее, чем состояния, удаленные друг от друга во времени. Это необходимо учитывать при построении модели временного ряда на основе нейронной сети. Задача предсказания будущего состояния системы на основе прошлого, т.е., построение временных рядов, является одной из важнейших проблем машинного обучения [3].

Нейронные сети называются рекуррентными в случае, если они имеют обратные связи, одну или несколь-

ко. В наличии обратной связи заключается основное отличие рекуррентных сетей от сетей прямого распространения: в каждый момент времени сеть получает кроме входных данных информацию о прошлых состояниях сети. Форма обратной связи может быть сконфигурирована разными способами [4].

Преимущество использования рекуррентных ИНС заключается в большей гибкости прогноза, подстройке сети под изменяющуюся внешнюю среду и способности качественно прогнозировать временные ряды [5]. В настоящий момент у этой архитектуры есть недостаток, связанный с быстро возрастающей сложностью на этапе обучения при увеличении длины последовательности и при усложнении архитектуры. Данную проблему можно обойти, сохранив при этом обратную связь, при помощи нейронных эхо-сетей.

В данной работе рассматривается нейронная эхо-сеть (НЭС), архитектура рекуррентных нейронных сетей, способная эффективно предсказывать последовательность значений временного ряда. Обладая невысокой вычислительной сложностью на этапе обучения, НЭС показывает результаты, сильно зависящие от случайно сге-

нерированных начальных значений. Предметом сравнения в данной работе является величина погрешности сгенерированных НС данных относительно реальных показателей. Объектом обработки для НЭС и ее аналогов выступают временные ряды, представляющие собой суперпозицию гармонических сигналов разных частот. Результаты работы НЭС будут сравниваться с другими архитектурами нейронных сетей [6, 7]:

- ◆ рекуррентная нейронная сеть Элмана (РНЭС, далее используется англоязычная аббревиатура ERNN)
- ◆ сеть с долгой краткосрочной памятью (ДКП, далее LSTM)
- ◆ сеть с управляемыми рекуррентными нейронами (УРН, далее GRU)

### Нейронные эхо-сети

Нейронные эхо-сети — нестандартная архитектура рекуррентных сетей, главное преимущество которой — очень низкая вычислительная сложность на этапе обучения [8]. Основная идея НЭС состоит в применении скрытого слоя, называемого резервуаром, внутри которого нейроны соединены случайно сгенерированными связями. В классическом описании вводится понятие состояния резервуара, вычисляемое по формуле:

$$x(n) = \tanh(W_{reservoir}x(n-1) + W_{input}i(n-1) + W_{feedback}y(n-1)) \quad (1)$$

Здесь  $n$  — номер шага по времени,  $i(n)$  — входной сигнал,  $y(n)$  — выходной сигнал.  $W_{input}$ ,  $W_{reservoir}$  и  $W_{feedback}$  — входная матрица, матрица весов связей в резервуаре и матрица обратной связи, соответственно. Все три матрицы наполняются случайно сгенерированными числами и не изменяются в процессе тренировки — за счет этого достигается прирост в скорости обучения.

На основе текущего состояния резервуара система вычисляет выходной вектор значений, путем умножения на выходную матрицу:

$$y(n) = W_{output}x(n) \quad (2)$$

Выходной вектор значений затем передается обратно в систему, обеспечивая обратную связь. Параметры матрицы  $W_{output}$  — это единственные веса, изменяющиеся в процессе обучения НЭС.

### Эмуляция полосного фильтра нейронами

Одной из проблем НЭС является ее зависимость от начальных значений, случайно задаваемых при гене-

рации. Чтобы увеличить точность предсказания модели, требуется множество итераций генерации начальных параметров для нахождения комбинации, наиболее подходящей под данные и условия задачи. Один из способов устранения этой проблемы при построении модели временных рядов описывает новый тип нейрона, реализующий полосный фильтр как разность двух низкочастотных:

$$x_1(n) = (1 - \alpha_1)x(n-1) + \alpha_1 \tanh(W_{reservoir}x(n-1) + W_{input}i(n-1) + W_{feedback}y(n-1)) \quad (3)$$

$$x_2(n) = (1 - \alpha_2)x_2(n-1) + \alpha_2 x_1(n-1) \quad (4)$$

где коэффициенты

$\alpha_1$  и  $\alpha_2$  пропорциональны частоте среза фильтра [8].

Итоговая формула выглядит следующим образом:

$$x(n) = x_1(n) - x_2(n) \quad (5)$$

На практике не всегда просто определить частоту среза на основе массива неотфильтрованных данных. В случае ошибочного подбора данных эти значения ухудшают качество предсказания модели. Поэтому мы добавили эти коэффициенты к изучаемым параметрам.

Из-за наличия этих параметров становится невозможным использовать традиционный метод обучения НС. Вместо него для обучения будет использован алгоритм Adam [10].

### Архитектура глубокой НЭС

Другой метод улучшения точности модели, предлагает использовать сочетание из нескольких сетей, каждая из которых использует отличный от других набор значений параметров «резервуара» [11]. Для нашей модели воспользуемся архитектурой с последовательно передающими друг другу состояние «резервуара» слоями. Каждый слой представляет собой полосный фильтр определенной частоты, поэтому коэффициенты  $\alpha_1$  и  $\alpha_2$  внутри каждого слоя одинаковы. На рисунке 1 изображено графическое представление архитектуры глубокой НЭС.

Для данной архитектуры формула, описывающая состояние «резервуара» каждого слоя, принимает следующий вид:

$$x_j^l(n) = (1 - \alpha_j^l)x_j^l(n-1) + \alpha_j^l \tanh(W_{reservoir}^l x_j^l(n-1) + W_{input}^l i(n-1) + W_{feedback}^l y^l(n-1)) \quad (6)$$

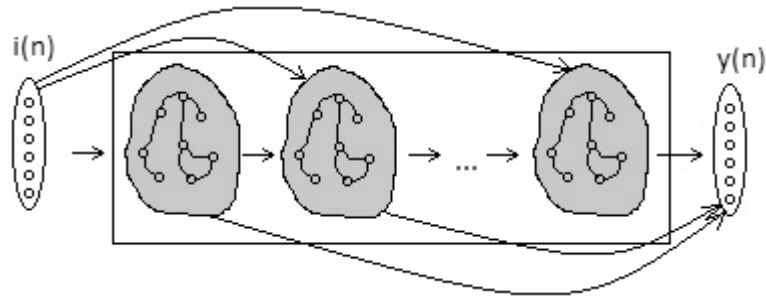


Рис. 1. Архитектура глубокой НЭС

$$x_2^l(n) = (1 - \alpha_2^l)x_2^l(n - 1) + \alpha_2^l x_1^l(n - 1) \quad (7)$$

$$x^l(n) = x_1^l(n) - x_2^l(n) \quad (8)$$

Здесь  $l$  — номер слоя.

Для вычисления результирующего значения после слоев эхо-сетей были добавлены два слоя прямого пространства. Первый слой состоит из нейронов, которые вычисляют выходное значение каждой эхо-сети, получая вектор  $y(n)$ . Значение каждого элемента рассчитывается по формуле:

$$y^l(n) = W_{output}^l x^l(n) \quad (9)$$

Второй слой является простым полносвязным слоем, выполняющим следующую операцию:

$$Y(n) = W_Y y(n) \quad (10)$$

### Оптимизация процесса подбора параметров

В большинстве случаев структурные параметры НС выбираются вручную, основываясь на опыте и конкретной задаче [12]. Это может привести, например, к слишком большому размеру резервуара, что влечет за собой необоснованное увеличение вычислительной сложности. Слишком малый размер резервуара, в свою очередь, снижает качество модели и точность прогноза. Для оптимизации процесса подбора параметров и повышения точности вычислений предлагается использовать алгоритм недоминирующей сортировки (АНС, далее NGSA-II) [13]. Структура сети представляется в виде вектора гиперпараметров, состоящего из:

- ◆ вероятность формирования связи в резервуаре;
- ◆ вероятность формирования связи в матрице обратной связи;
- ◆ размер резервуара;
- ◆ диапазон значений резервуара;
- ◆ диапазон значений входной матрицы;
- ◆ диапазон значений матрицы обратной связи;

- ◆ число слоев — полосных фильтров.

Для каждого набора параметров составлялся вектор значений, по значению вектора строилась нейронная сеть, которая затем обучалась и тестировалась.

### Эксперимент

В процессе тестирования было предложено предсказать несколько значений из набора временных рядов. Каждый временной ряд получаем путем суммирования  $M$  синусоид:

$$Y = \sum_{i=1}^M \sin(\omega_i n) \quad (11)$$

Были использованы следующие значения частот:  $\omega_1 = 0.2, \omega_2 = 0.311, \omega_3 = 42, \omega_4 = 0.51, \omega_5 = 0.63, \omega_6 = 0.74, \omega_7 = 0.85, \omega_8 = 0.97, \omega_9 = 1.08, \omega_{10} = 1.19, \omega_{11} = 1.27, \omega_{12} = 1.32$ .

В качестве тренировочных данных были использованы 500 заранее подготовленных эталонных результатов вычислений, а в качестве тестовых — неизвестные последующие 20 результатов. Для увеличения сложности задачи, также сеть тестировалась на тех же данных с добавлением случайного шума, сгенерированного с использованием нормального распределения  $N(0, 0.2)$ .

Для оценки качества работы нашей модели мы сравнили ее с другими архитектурами, перечисленными ранее. Для обучения сетей использовался алгоритм Adam. В качестве функции потерь была выбрана функция приведенной среднеквадратичной ошибки:

$$Loss = \sqrt{\frac{\sum_{t=1}^n (y - \hat{y})^2}{n(y_{max} - y_{min})^2}} \quad (12)$$

Генетический алгоритм стартовал с популяции в 1000 особей и сошелся через примерно 140 поколений. Каждая сеть тренировалась и тестировалась на 8 временных рядах с  $M = 2, 5, 8, 12$  с добавлением шума и без него.

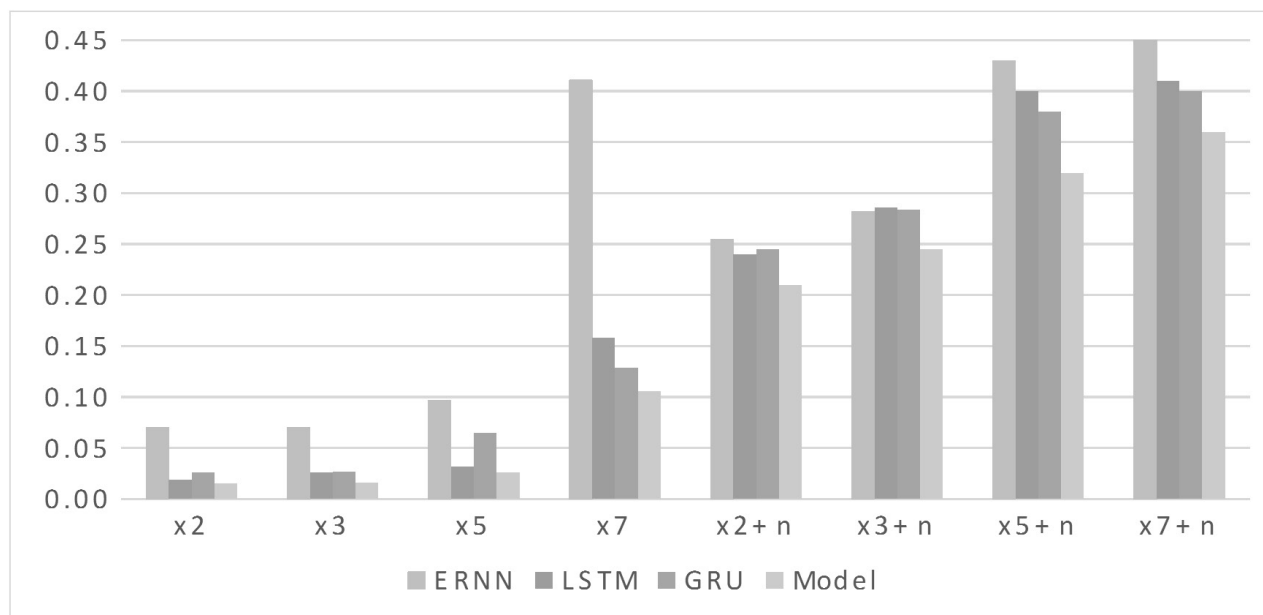


Рис. 2. Средняя величина ошибки для каждой НС

На рисунке 2 показана средняя величина ошибки 20 независимых испытаний для каждого временного ряда. Предложенная модель НЭС показала лучшую результативность в каждом испытании. Как и ожидалось, на зашумленных данных результат предсказаний значительно падает у каждой модели.

Архитектуры сетей, основанные на методе градиентной оптимизации (ERNN, LSTM, GRU), демонстрируют в среднем большее значение ошибки прогнозирования, чем наша модель. Также следует отметить, что сеть ERNN показывает заметно худшие результаты, чем модели LSTM и GRU, что особенно очевидно при экспериментах на незашумленных рядах. Такой результат связан с более простой архитектурой, которая неспособна корректно обрабатывать долгосрочные корреляции в анализируемых временных рядах. Результаты LSTM и GRU достаточно близки, и сложно сделать однозначный вывод о превосходстве в точности одной модели над другой.

Помимо лучших результатов эксперимента предложенной модели среди рассматриваемых, следует отметить, что время обучения нашей модели значительно

ниже, чем у сетей других архитектур, так как обучается меньшее число параметров.

## Заключение

В данной работе была описана архитектура рекуррентной нейронной сети, отличающейся от традиционных НС наличием скрытого слоя, названного резервуаром. В НЭС каждый нейрон представляет собой полосный фильтр, полученный путем комбинирования двух фильтров низких частот. Случайно соединенные нейроны образуют резервуар.

В работе была представлена структура нейронной сети в виде вектора гиперпараметров и описаны составляющие этого вектора. Число слоев, размер резервуара и другие составляющие можно вычислить при помощи генетического алгоритма.

Описанные улучшения создают дополнительные вычислительные расходы, повышая вычислительную сложность модели. Но они компенсируются простотой обучения НЭС и позволяют значительно повышать точность моделирования временных рядов и их анализа.

## ЛИТЕРАТУРА

1. А. Н. Горбань, Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей, Сиб. журн. вычисл. матем., 1998, том 1, номер 1, 11–24
2. Терехов В.И., Жуков Р.В. Современное состояние и перспективы применения импульсных нейронных сетей // Информационно-измерительные и управляющие системы. Москва: Радиотехника. 2016 — № 12 — С. 143–145.

3. Кузовлев В.И., Орлов А. О. Метод выявления аномалий в исходных данных при построении прогнозной модели решающего дерева в системах поддержки принятия решений // Наука и образование. МГТУ им. Н.Э. Баумана. Электрон. журн. 2012. № 9.
4. Нейронные сети: полный курс, 2 е издание.: Пер. с англ. М. Издательский дом «Вильямс», 2006. 1104 с.
5. Воронов Иван Викторович, Политов Евгений Александрович, Ефременко Владимир Михайлович Обзор типов искусственных нейронных сетей и методов их обучения // Вестник КузГТУ. 2007. № 3. URL: <http://cyberleninka.ru/article/n/obzor-tipov-iskusstvennyh-neyronnyh-setey-i-metodov-ih-obucheniya> (дата обращения: 03.04.2017).
6. Рекуррентная нейронная сеть Элмана и ее моделирование в MATLAB/ Константинова Е. А. Хлевицкий А. В. // Вестник Морского государственного университета. Сер. Автоматическое управление, математическое моделирование и информационные технологии. — Владивосток: Мор. гос. ун-т, 2010. — Вып. 37.
7. Будыльский Д. В. GRU и LSTM: современные рекуррентные нейронные сети // Молодой ученый. — 2015. — № 15. — С. 51–54.
8. H. Jaeger. Short term memory in echo state networks. GMD Report 152, GMD -German National Research Institute for Computer Science, 2002. <http://www.gmd.de/People/Herbert.Jaeger/Publications.html>.
9. Siewert, U., and Wustlich, W. Echo-state networks with band-pass neurons: Towards generic time-scale-independent reservoir structures. Internal status report, PLANET intelligentsystems GmbH (2007).
10. Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. ICLR, 2015. arXiv:1412.6980
11. C. Gallicchio, A. Micheli (2016). Deep Reservoir Computing: A Critical Analysis Claudio Gallicchio and Alessio Micheli // ESANN2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27–29 April 2016, publ., ISBN978–287587027–8.
12. Лабунец Л.В., Лабунец Н.Л., Чижов М. Ю. Рекуррентные статистики нестационарных временных рядов // Радиотехника и электроника. 2011. Т. 56. № 12. С. 1468–1489.
13. Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. A Fast Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2):182–197, April 2002.

---

© Шелоник Антон Александрович ( [antonshelonik@gmail.com](mailto:antonshelonik@gmail.com) ), Смирнов Артём Сергеевич ( [artem.smirnov@hotmail.com](mailto:artem.smirnov@hotmail.com) ),  
Калистратов Алексей Павлович ( [akalistratov@gmail.com](mailto:akalistratov@gmail.com) ), Колдобский Владислав Игоревич ( [admin@nevergone.ru](mailto:admin@nevergone.ru) ).  
Журнал «Современная наука: актуальные проблемы теории и практики»

