

## **Заключение.**

В работе описываются нейронные сети для автономного управления роботом на узких интервалах движения. Она характеризуется тем, что при соответствующем обучении, способна обеспечивать корректное управление роботом при неточной информации от сенсорных устройств. Проведено компьютерное моделирование нейронных сетей, которое показало устойчивое функционирование их для управления мобильным роботом. Натурные эксперименты планируется осуществить на реально действующем роботе "Walter"(Германия).

## **АРХИТЕКТУРА И АЛГОРИТМ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ ПРОГНОЗИРОВАНИЯ**

*Савицкий Ю.В.*

*Брестский политехнический институт*

## **Введение**

Одним из фундаментальных свойств нейронных сетей является их способность после обучения к обобщению и пролонгации результатов. Это создает предпосылки для создания на их базе различного рода систем прогнозирования [1, 2, 4]. Основными проблемами в технологии нейросетевого прогнозирования является выбор архитектуры нейронной сети, способной адекватно описывать прогнозируемый процесс и выполнять успешный прогноз. При этом важным вопросом является выбор типов нейроэлементов в архитектуре прогнозирующей нейросистемы, от которого во многом зависит способность модели сформировать оптимальную прогнозирующую функцию во время обучения. Другой важнейшей проблемой в технологии

создания прикладных нейросистем, является разработка эффективных алгоритмов обучения, обладающих глобальной сходимостью, высокой скоростью и точностью обучения [1, 3, 5]. При этом в большинстве случаев параметры обучения нейронной сети определяют эффективность применения нейросетевых технологий в целом. В данной работе обсуждаются вопросы выбора типов нейроэлементов в архитектуре нейронной сети прогнозирования, рассматривается адаптивный алгоритм обучения для нейронной сети выбранной архитектуры, анализируются достоинства и недостатки разработанных архитектур и методов обучения по сравнению с классическими.

## 1. Архитектура нейронной сети прогнозирования.

Одним из главных моментов в разработке прогнозирующих систем является выбор прогнозной модели, способной с минимальной ошибкой описать прогнозируемый процесс, так как от этого зависит точность прогноза. Поэтому нейросетевая модель должна быть такой архитектуры, чтобы быть способной с необходимой точностью описать исследуемый процесс и сформировать корректную прогнозирующую функцию. В качестве базовой была принята двухслойная нейронная сеть с прямыми связями, имеющая  $I$  входов,  $N$  скрытых нейроэлементов и один выходной нейроэлемент. Входной паттерн  $P$  размерностью  $I$  во время функционирования нейронной сети преобразуются в выходную активность в соответствии с выражением:

$$Y(P) = f \left( \sum_{j=0}^N w_{j0} \cdot g \left( \sum_{i=1}^I w_{ij} \cdot P_i - T_j \right) - T_0 \right), \quad (1)$$

Скрытые нейроэлементы имеют логарифмическую функцию активации:

$$g(x_j) = \ln(x_j + \sqrt{x_j^2 + 1}), \quad (2)$$

где

$$x_j = \sum_i w_{ij} \cdot Y_i - T_j \quad (3)$$

$Y_i$ -входная активность нейроэлемента;  $T_j$  -порог нейроэлемента  $j$ . Единственный выходной нейроэлемент использует линейную функцию активации.

$$F(x_0) = K \cdot x_0, K = 1 \quad (4)$$

где  $x_0 = \sum_i w_{i0} \cdot Y_i - T_0$ .

Использование неограниченной на всей области определения логарифмической функции активации вместо традиционной сигмоидной позволяет улучшить адаптивность и прогнозирующие свойства нейронных сетей, применяемых для прогнозирования нестационарных процессов.

## 2. Алгоритм обучения нейронной сети прогнозирования.

Наиболее популярным методом обучения для многослойных нейронных сетей является алгоритм обратного распространения ошибки (Back Propagation Error). Данный алгоритм базируется на методе градиентного спуска в пространстве весовых коэффициентов нейроэлементов и состоит из набора итерационных процедур преобразования весов и порогов нейроэлементов по следующим правилам:

$$\Delta w_{ij}(t) = -\alpha \frac{\partial E_p}{\partial w_{ij}(t)}, \quad (4)$$

$$\Delta T_j(t) = -\alpha \frac{\partial E_p}{\partial T_j(t)}, \quad (5)$$

где  $\frac{\partial E_p}{\partial w_{ij}}, \frac{\partial E_p}{\partial T_j}$  - градиенты функции ошибки для тренировочного эталона

$P$ :

$$E_p = \frac{1}{2} \sum_j (Y_j^P - D_j^P)^2, \quad (6)$$

### 3. Искусственный интеллект и нейронные сети

$Y_j^p$  - выходная активность сети;  $D_j^p$  - желаемое значение выхода для тренировочного эталона  $P$ . Во время обучения происходит процесс снижения общей ошибки сети:

$$E = \sum_{p=1}^P E_p . \quad (7)$$

С целью улучшения параметров обучения сети и устранения недостатков классического алгоритма обратного распространения ошибки, связанного с эмпирическим выбором постоянного шага обучения, используем метод скорейшего спуска для вычисления адаптивного шага обучения, в соответствии с которым:

$$\begin{aligned} \Delta w_y(t) &= -\alpha(t) \frac{\partial E}{\partial w_y(t)}, \\ \Delta T_j(t) &= -\alpha(t) \frac{\partial E}{\partial T_j}, \\ \alpha(t) &= \min \left\{ E(w_y(t+1), T_j(t+1)) \right\}, \end{aligned} \quad (8),$$

где  $\alpha(t)$  - значение шага, адаптируемое на каждой итерации  $t$  для каждого эталона  $P$ . Для линейной функции активации адаптивный шаг обучения определяется выражением [1]:

$$\alpha(t) = \frac{1}{K^2 \cdot (\sum_i Y_i^2 + 1)}, \quad (9)$$

где  $Y_i(t)$  - элементы входной активности для линейного нейроэлемента. Для логарифмической функции активации оценка активного шага может быть получена согласно выражению:

$$\hat{a}(t) = \frac{\sum_j g_j^2 \cdot Y_j'}{\left(1 + \sum_i Y_i^2\right) \cdot \sum_j g_j^2 \cdot (Y_j')^2}, \quad (10)$$

где  $Y_j'$  - производная логарифмической функции

$$Y_j' = \frac{1}{\sqrt{x_j^2 + 1}}$$

$$\gamma_i = \sum_j \gamma_j \cdot Y_j' \cdot w_{ij}$$

-для нейроэлементов скрытого слоя.

Использование алгоритма обучения с адаптивным шагом позволило значительно сократить число итераций обучения, увеличить точность и стабильность процесса обучения.

## Экспериментальные результаты.

Описанные выше архитектура нейронной сети и алгоритм обучения были использованы для прогнозирования стока вод реки Березины на территории Республики Беларусь. При этом временной ряд содержал 40 среднегодových значений стока вод с 1938г. по 1976г., взятых с коэффициентом 0.001. Из данного ряда было сформировано обучающее множество нейронной сети методом скользящего окна размерностью 10 и шагом 1. Выбранная архитектура нейронной сети содержала 10 входов, 15 скрытых и 1 выходной нейроэлемент. Результаты процедуры обучения показали значительное увеличение скорости обучения алгоритмом с адаптивным шагом по сравнению с традиционным алгоритмом обратного распространения ошибки. Точность обучения также повысилась. Так, используя классический алгоритм, минимальная ошибка обучения оставалась равной 0.38 и не

### 3. Искусственный интеллект и нейронные сети

уменьшалась при дальнейшем обучении сети (см. рис.1а). При использовании разработанного выше метода обучения погрешность обучения за 2000 итераций достигла  $10E-7$  (см. рис.1б). Результаты долгосрочного прогноза стока вод на период с 1977г. по 1991г. (см. рис.2.) демонстрируют потенциальные способности нейронной сети принятой архитектуры для выполнения анализа и прогнозирования сложных временных рядов.

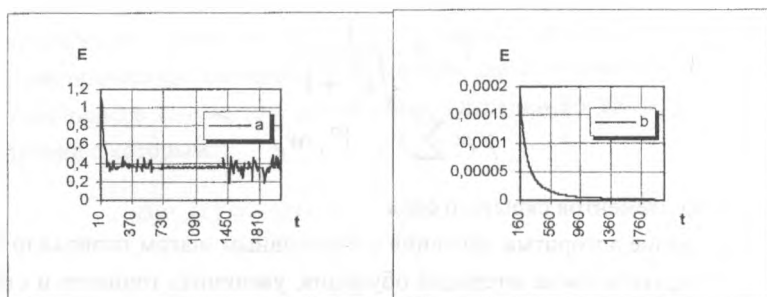


Рис 1. Результаты обучения нейронной сети: (а) – с постоянным шагом обучения; (б) – с адаптивным шагом обучения

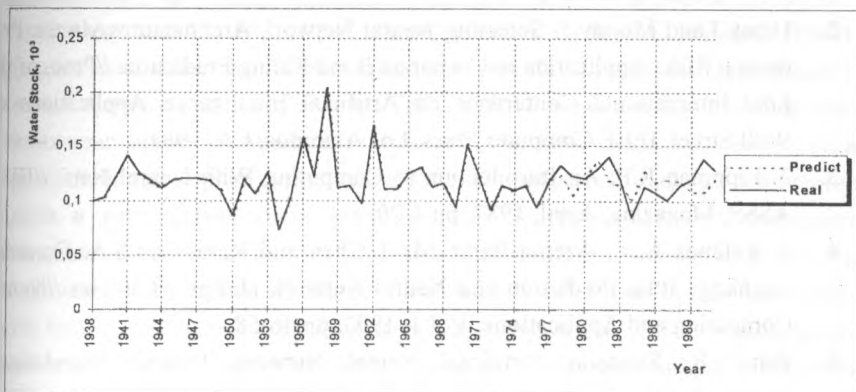


Рис 2. Временной ряд и результаты прогноза уровня стока вод реки Березина

## Заключение.

В данной работе рассмотрены вопросы выбора архитектуры нейронных сетей прогнозирования, позволяющих адекватно описывать исследуемые процессы и выполнять корректный прогноз. Применение логарифмической функции активации вместо традиционной сигмоидной позволяет улучшить прогнозирующие способности нейронных сетей для задач прогнозирования нестационарных процессов. Алгоритмы обучения с адаптивным шагом, разработанные в рамках данной работы для логарифмической функции активации нейроэлементов, позволили значительно улучшить параметры обучения нейронной сети. На реальных данных продемонстрированы потенциальные возможности нейронных сетей разработанной архитектуры выполнять прогнозы сложных процессов.

## Литература

1. Vladimir A. Golovko, Jury V. Savitski, Vitaly B. Gladischuk. Predicting Neural Net // Proc. of the CMNDT-95, Minsk, Belarus, 1995, pp.348-353.

2. Utans J. and Moody J. Selecting Neural Network Architectures via the Prediction Risk: Application to Corporate Bond Rating Prediction. //Proc. of the First International Conference on Artificial Intelligence Applications on Wall Street. IEEE Computer Press, Los Alamitos, CA, 1991.
3. 3. Lippman R.P. An Introduction to Computing With Neural Nets. //IEEE ASSP, Magazine, April, 1987, pp.4-20.
4. 4. Refenes A.N., Azema-Barac M., L.Chen and Karaussos S.A. Currency Exchange Rate Prediction and Neural Network Design Strategies //Neural Computing and Applications, Vol.1, 1993, pp.46-58.
5. Patric R. Simpson. Artificial Neural Network System Foundations, paradigms, applications and implementations. //San Diego, Pergamon Press,-1990

## НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ

*Савицкий Ю.В., Ваткин М.Е., Поляков М.М*

*Брестский политехнический институт*

Четыре года назад на кафедре ЭВМ и Систем были инициированы научные исследования в области нейронных сетей (НС). Одним из направлений научной деятельности является создание на базе НС различного рода прогнозирующих систем. При этом в качестве базовой использовалась многослойная НС с обучением по методу обратного распространения ошибки. Для НС данной архитектуры были получены новые аналитические и практические результаты, связанные с разработкой эффективных алгоритмов обучения, выбора архитектуры сети и типа нейроэлементов, с целью повышения точности прогнозирования. В настоящее время рабочей группой по нейронным сетям начаты исследования по использованию альтернативных профилей нейронных сетей в задачах прогнозирования: ре-