

Головко В.А., Хацкевич М.В., Брич А.Л.

МЕТОД ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА

Введение. Прогнозирование временных рядов является важной задачей, решение которой позволяет предсказать эволюцию динамической системы на упреждающий промежуток времени. Особенно актуальным является прогнозирование экономических и финансовых временных рядов. В настоящее время существует большое количество подходов к прогнозированию временных рядов. Традиционный подход к прогнозированию временных рядов состоит в использовании статистических методов, например, ARIMA. Однако в последнее время все большую популярность приобретают нейросетевые методы, которые обладают большей гибкостью и легко позволяют учитывать нелинейную природу временных рядов. В последнее десятилетие появилось более 2000 статей, посвященных применению нейронных сетей для прогнозирования временных рядов [1]. Для того чтобы адекватно производить сравнение различных нейросетевых моделей, проводятся различного рода соревнования в области прогнозирования временных рядов [2]. Существуют различные нейросетевые подходы к прогнозированию временных рядов. Это многослойные перцептроны, рекуррентные нейронные сети, RBF сети, обобщенные регрессионные нейронные сети (GRNN) и т.д. [3–8]. Как правило, в настоящее время прогнозирующие системы объединяют различные типы нейронных сетей с целью эффективного прогнозирования.

В данной статье предлагается метод прогнозирования временных рядов на основе многослойного перцептрона, который включает как специальный способ представления входных данных, так и способ обучения нейронной сети для прогнозирования. Представленный здесь метод является частью системы прогнозирования временных рядов.

1. Стандартный подход к прогнозированию. Цель предсказания временного ряда может быть описана следующим образом: для данной последовательности $x(1), x(2), \dots, x(l)$ необходимо найти её продолжение $x(l+1), x(l+2), \dots$. Нелинейная модель для предсказания может быть представлена как $x(t) = Fx((t-1), x(t-2), \dots, x(t-k))$, где $t = k+1, N1$, F – нелинейная функция, построенная с помощью искусственной нейронной сети, и k – размер скользящего окна, который равен числу входных нейронов сети. В качестве базовой архитектуры для прогнозирования временных процессов в работе использован многослойный перцептрон (Multilayer Perceptron, MLP). Доказано, что данная нейронная сеть способна аппроксимировать любую непрерывную функцию со сколь угодно высокой точностью. Другим важным свойством MLP является способность к обобщению информации, представленной в множестве обучения. Следовательно, MLP является мощным инструментом для построения прогнозирующих систем (рисунок 1).

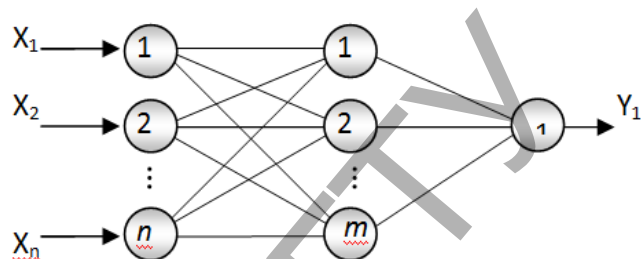


Рис. 1. Многослойный перцептрон

Как следует из представленного выше, стандартный подход к прогнозированию учитывает только предыдущие значения временного ряда для предсказания следующих значений.

Прогнозирование будем проводить на образцах, отличных от тех, которые подавались в обучающей выборке. Для того, чтобы нейронная сеть функционировала корректно, необходимо нормализовать входные данные. Вариантов нормализации довольно много, однако наилучшие результаты достигаются при проецировании области определения входного множества на отрезок $[0; 1]$.

2. Модифицированный подход к прогнозированию. Данный подход заключается как в специальном способе представления входных данных, так и в способе обучения нейронной сети для прогнозирования. На первом этапе временной ряд преобразуется следующим образом:

1. Во временном ряде выделяются только экстремальные значения (минимумы и максимумы), которые будут использоваться для прогнозирования.
2. Составляются пары значений (значение в экстремуме и промежуток времени от предыдущего экстремума), которые являются входными данными для прогнозирующей нейронной сети (рисунок 2).

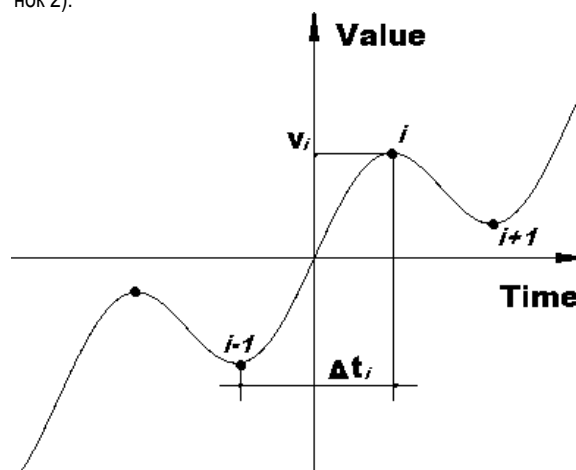


Рис. 2. Алгоритм составления пар «значение-время»

Головко Владимир Адамович, д.т.н., профессор, зав. кафедрой интеллектуальных информационных технологий Брестского государственного технического университета.

Хацкевич М.В., ст. преподаватель кафедры интеллектуальных информационных технологий Брестского государственного технического университета.

Брич А.Л., ассистент кафедры интеллектуальных информационных технологий Брестского государственного технического университета.

Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.

Таким образом, входными данными для нейронной сети являются значения временного ряда в точках экстремума и соответствующие им временные значения, которые формируются по методу скользящего окна. В результате на вход сети подается n пар значений временного ряда. Каждая пара состоит из значения в экстремуме и промежутка времени до предыдущего экстремума. Выходной слой многослойного персептрона содержит два значения: значение последующего экстремума и промежуток времени, через который данный экстремум будет достигнут. В результате выходные значения персептрона определяются следующим образом:

$$x(n+1) = F(x(1), x(2), \dots, x(n)),$$

$$t(n+1) = F(t(1), t(2), \dots, t(n)).$$

Архитектура сети представлена на рисунке 3. Рассмотрим модификацию алгоритма обучения нейронной сети. Модификация алгоритма обучения заключается в том, что сначала проводится классическое обучение нейронной сети на эталонной обучающей выборке, а затем осуществляется обучение на прогнозируемых данных. В результате минимизируется значение ошибки, полученной путем подачи на входы нейронной сети спрогнозированных ранее данных (итеративная ошибка). Таким образом, нейронная сеть сначала обучается на эталонных значениях, и, как только ошибка достигнет приемлемого уровня, начинается обучение сети на прогнозируемых данных. В результате ошибка прогнозирования должна уменьшиться и сеть должна с большей точностью проводить итеративное прогнозирование.

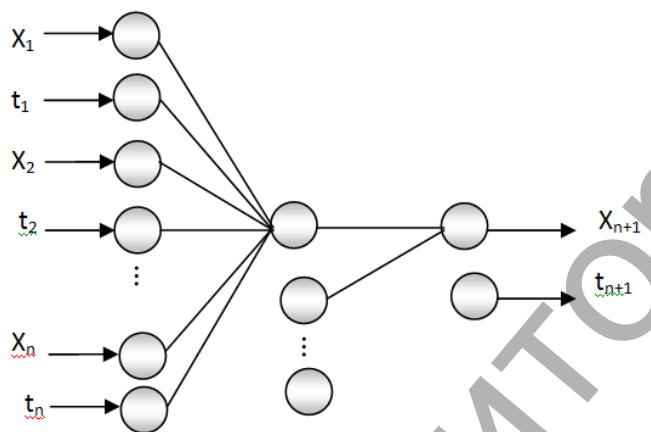


Рис. 3. Архитектура сети

3. Экспериментальные результаты на синтетических данных. Рассмотрим применение разработанного подхода для прогнозирования хаотических временных рядов на базе многослойного персептрона. В качестве исследуемых хаотических систем будем использовать аттрактор Энона. Процесс Энона описывается следующими уравнениями:

$$\begin{cases} x_{n+1} = 1 - \alpha x_n^2 + y_n \\ y_{n+1} = \beta x_n, \end{cases}$$

где $\alpha = 1,4$ и $\beta = 0,3$ для хаотического поведения системы.

На рисунке 4. изображен временной ряд Энона (координата x).

В качестве входных данных будем использовать координату (x) последовательности Энона. Для обучения возьмем 400 первых значений ряда и 300 значений для прогнозирования. Обучение будем проводить по методу скользящего окна. Прогнозирование осуществляется на образах, отличных от тех, которые подавались в обучающей выборке. Будем использовать многошаговый и одношаговый прогноз. Одношаговым прогнозированием называют краткосрочный прогноз (на один шаг), при этом для получения спрогнозированной величины используют только фактические (эталонные) данные. При данном методе прогнозирования на входы нейронной сети подаются эталонные значения (значения из выборки для прогнозирования) (рисунок 5).

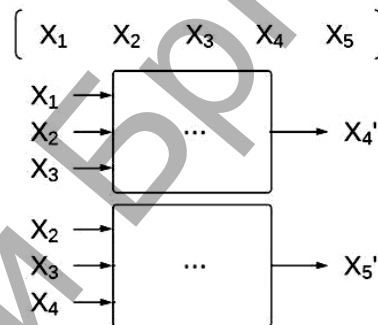


Рис. 5. Одношаговый прогноз

Многошаговым (итеративным) прогнозом называют долгосрочный прогноз, на основе данных, полученных с предыдущего этапа прогнозирования (рисунок 6).

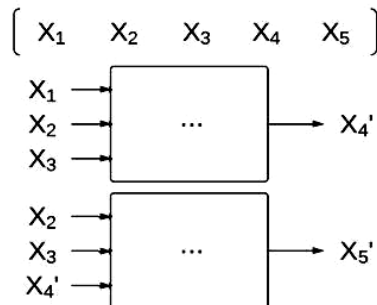


Рис. 6. Многошаговый прогноз

Оценку качества прогнозирования будем производить, используя ошибку MAPE (Mean absolute percentage error) и суммарную

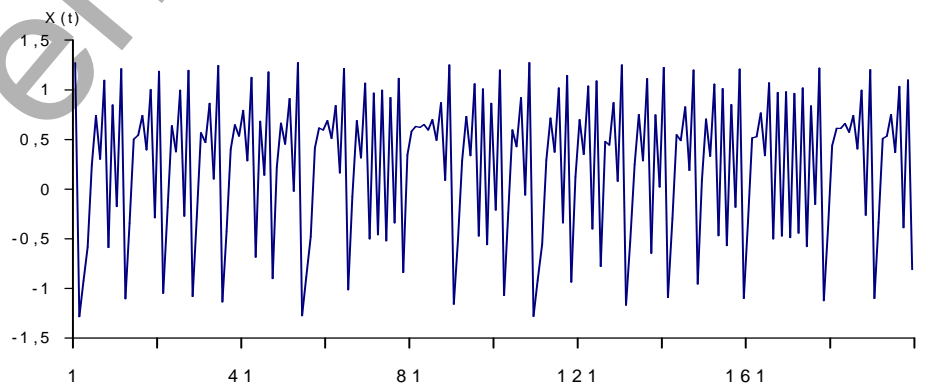


Рис. 4. Последовательность Энона (первые 200 элементов ряда)

среднеквадратичную ошибку MSE. MAPE – средняя абсолютная ошибка в процентах, используется для оценки точности построения прогнозов временных рядов в статистике. Рассчитывается MAPE по следующей формуле:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|,$$

где A_t – реальное значение, F_t – спрогнозированное значение.

Среднеквадратичная ошибка MSE рассчитывается по следующему выражению:

$$MSE = \frac{1}{2} \sum_{t=1}^n (A_t - F_t)^2.$$

Рассмотрим применение обычного многослойного перцептрона для прогнозирования ряда Энона. Наилучшие результаты сеть показала при архитектуре 28-13-1. Для прогнозирования на шаг вперед (рисунок 7) среднее значение MAPE составило 38,3%, а среднее значение MSE – 0,034.

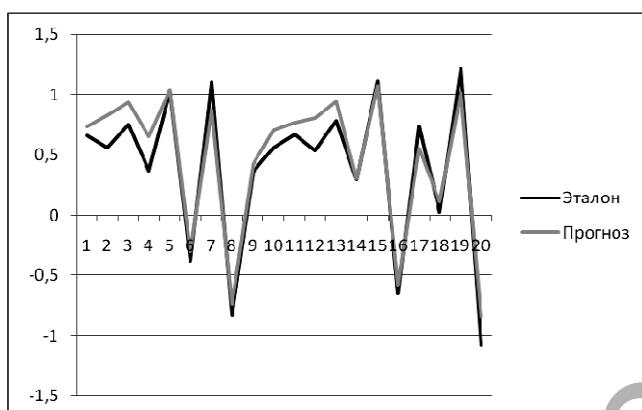


Рис. 7. Результат прогнозирования на шаг вперед для многослойного перцептрона (первые 20 значений)

Для итеративного прогноза (рисунок 8) среднее значение MAPE составило 237,3%, а среднее значение MSE – 0,950. Таким образом, для проведения долгосрочных прогнозов классическая архитектура многослойного перцептрона неприменима, поскольку уже на пятом шаге ошибка достаточно велика (126,1%).

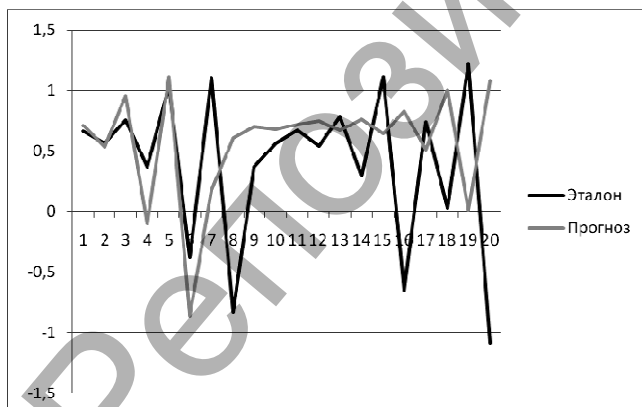


Рис. 8. Результат итеративного прогнозирования для многослойного перцептрона (первые 20 значений)

Рассмотрим предложенный подход к прогнозированию. Возьмем архитектуру сети 30-13-2 с сигмоидной функцией нелинейного преобразования в скрытом и выходном слоях. Рассмотрим прогнозирование на шаг вперед (рисунок 9). Среднее значение MAPE составило 36,7%, а среднее значение MSE – 0,085.

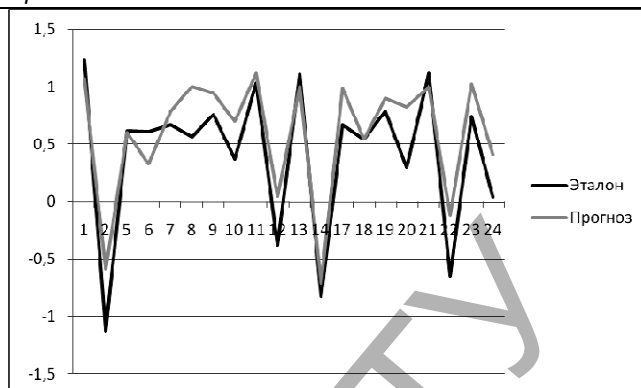


Рис. 9. Результаты прогнозирования на шаг вперед (первые 20 значений)

Среднее значение критерия MAPE и рисунок 9 свидетельствуют о достаточно высокой точности прогнозирования. Исходя из значения MAPE, предложенная сеть прогнозирует точнее, чем классический многослойный перцептрон (MAPE: 38,2%). Однако основной целью модифицированного обучения является минимизация итеративной ошибки. Для итеративного прогнозирования (рисунок 10), среднее значение MAPE составило 47,8%, а среднее значение MSE – 1,158.

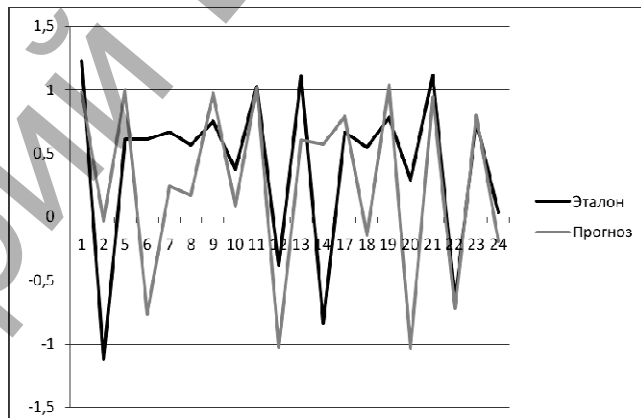


Рис. 10. Итеративное прогнозирование с помощью Value-Time сети и комбинированного обучения (первые 20 значений)

Судя по графику, точность прогноза не слишком велика. Однако это только первых 20 значений ряда (рисунок 10). Согласно ожиданиям, при модифицированном методе обучения ошибка не будет накапливаться (увеличиваться) в процессе прогнозирования. Для того, чтобы убедиться в этом приведем прогнозные и эталонные значения для 100–120 итераций, т.е. оценим ошибку прогнозирования на 120 значений вперед. Результат представлен на рисунке 11.

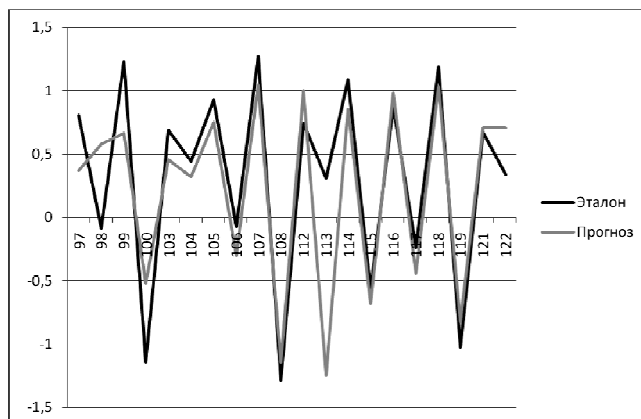


Рис. 11. Итеративное прогнозирование (100–120 значения)

Судя по графику, представленному на рисунке 11, можно сделать вывод о весьма высоком качестве итеративного обучения. Об этом также свидетельствует среднее значение MAPE, равное 47,8%. Таким образом, модифицированный подход показывает значительно лучшие результаты при итеративном прогнозировании, чем классический многослойный персептрон. Результаты проведенных испытаний сведены в таблицы 1 и 2. В таблице 1 сведены значения критерия MAPE, а в таблице 2 – значения критерия MSE.

Таблица 1. Значения критерия MAPE (%) для проведенных испытаний

	Способ прогнозирования	
	На шаг вперед	Итеративно
Алгоритм 1 - Многослойный персептрон	38,3	237,3
Алгоритм 2 - Модифицированный подход	36,7	47,8

Таблица 2. Значения критерия MSE для проведенных испытаний

	Способ прогнозирования	
	На шаг вперед	Итеративно
Алгоритм 1 - Многослойный персептрон	0,034	0,95
Алгоритм 2 - Модифицированный подход	0,085	1,158

4. Экспериментальные результаты на реальных данных

Тестирование системы производилось также на реальных значениях биржевого индекса. Рассмотрим персептрон с архитектурой 28-13-1 с сигмоидной функцией активации нейронов в скрытом и выходном слоях. В качестве данных для обучения используются первые 400 элементов ряда (60% всех элементов). Для оценки качества прогнозирования используются оставшиеся 40% значений ряда. При прогнозировании на шаг вперед (рисунок 12) среднее значение MAPE составило 63,7%, а среднее значение MSE – 298.917. Среднее значение MAPE и форма графика на рисунке 12 свидетельствуют о том, что данная архитектура неприменима к подобного рода прогнозам ввиду значительного уровня ошибки уже на первом шаге. Рассмотрим применение модифицированного подхода к прогнозированию на шаг вперед. Будем использовать архитектуру сети 30-13-2. Среднее значение MAPE составило 31,4%, а среднее значение MSE – 71,890.

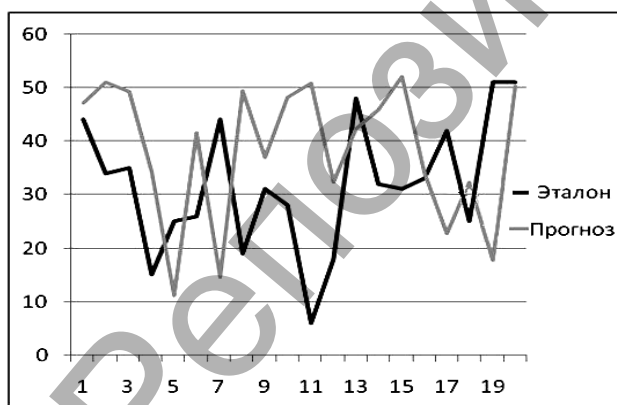


Рис. 12. Результаты прогнозирования на шаг вперед при помощи многослойного персептрона (первые 20 значений)

При данном алгоритме обучения, результаты прогнозирования на реальных данных получаются лучше. Об этом свидетельствует низкое значение индекса MAPE в сочетании с результатами, представленными на рисунке 13. Таким образом, данная архитектура может использоваться для краткосрочного прогнозирования временных рядов.

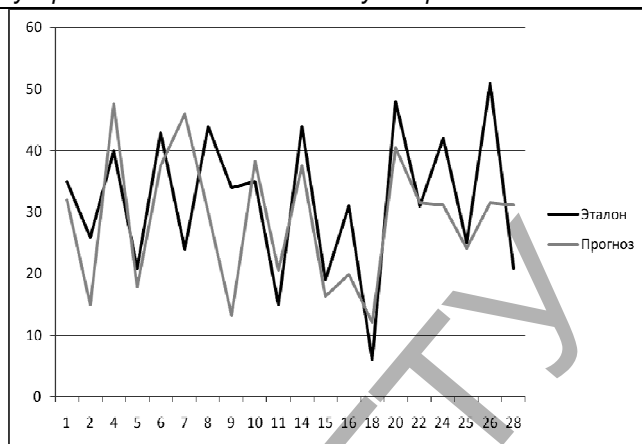


Рис. 13. Результаты прогнозирования на шаг вперед при помощи модифицированного подхода (первые 20 значений)

Рассмотрим итеративное прогнозирование (рисунок 14). Среднее значение MAPE составило 27,4%, а среднее значение MSE – 33,969, в то время как для обычного персептрона значение MAPE составило 67,754%, а среднее значение MSE – 330,219.

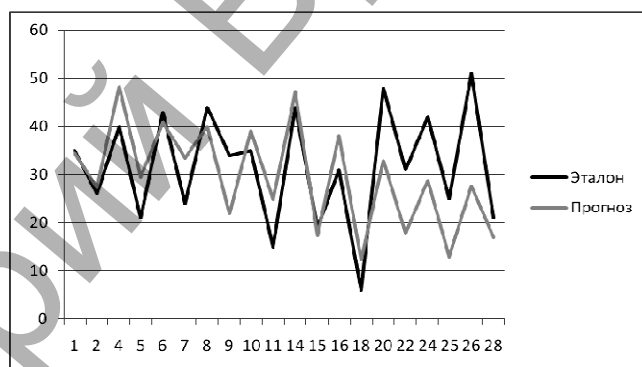


Рис. 14. Результаты итеративного прогнозирования при помощи Value-Time сети с комбинированным обучением (первые 20 значений)

Таким образом, полученные результаты наглядно демонстрируют преимущество предложенного подхода к прогнозированию.

Заключение. В данной статье предложен метод прогнозирования временных рядов на основе многослойного персептрона, который включает в себя как специальный способ представления входных данных, так и способ обучения нейронной сети для прогнозирования. Представленный метод является частью системы прогнозирования временных рядов. Полученные результаты наглядно демонстрируют преимущество предложенного подхода к прогнозированию.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Crone, S. An Evaluation Framework for Publications on Artificial Neural Networks in Sales Forecasting / S. Crone, P. Grafeille // H. Arabnia; R. Joshua; Y. Mun (eds.): proceedings of the International Conference on Artificial Intelligence, IC-AI'04. – Las Vegas: CSREA Press: Athens. – Vol. 1. – P. 221–227.
2. Makridakis, S. The M3-competition: results, conclusions, and implications / S. Makridakis, M. Hibon // International Journal of Forecasting, 2000. – Vol. 16. – P. 451–476.
3. Lim, C.P. The Application of an Ensemble of Boosted Elman Networks to Time Series Prediction: A Benchmark Study / C.P. Lim, W.Y. Goh // International Journal of Computational Intelligence, 2006. – Vol. 3. – No. 2. – P. 119–126.
4. Sorjamaa, A. Methodology for long-term prediction of time series / A. Sorjamaa, H. Jin, N. Reyhanian, Y. Jia, A. Lendasse // Neurocomputing, 2007. – Vol. 70. – Issues 16–18. – P. 2861–2869.

- Tripathi, M.M. Short-Term Load Forecasting Using Generalized Regression and Probabilistic Neural Networks in the Electricity Market / M.M. Tripathi, K.G. Upadhyay, S.N. Singh // The Electricity Journal, 2008. – Vol. 21. – Issue 9. – P. 24–34.
- Tsai, C.L. Federal Funds Rate Prediction Using Robust Radial Basis Function Neural Networks / C.L. Tsai, C.C. Lee, Y.C. Chiang // The 2nd International Conference on Innovative Computing, Information and Control (ICIC'07), Aug. 12–17, 2007. – P. 696–699.

Материал поступил в редакцию 19.11.13

GOLOVKO V.A., KHATSKEVICH M.V., BRICH A.L. A forecasting approach of time series basis on multilayer perceptron

The forecasting approach, base on multilayer perceptron is presentea. In consist of special technique for input data presentation and training approach for time series predition.

УДК 004.89

Иванюк Д.С., Головко В.А.

НЕЙРО-ПИД-КОНТРОЛЛЕР ПАСТЕРИЗАЦИОННОЙ УСТАНОВКИ

Введение. Нечеткая логика, нейруправление и оптимальное управление в настоящее время являются глубоко проработанными и успешно применяемыми подходами для эффективного решения многих реальных производственных задач управления [1]–[2]. Рост исследований в области нейруправления произошел после 1987 года, когда в Сан-Диего была проведена первая IEEE конференция по нейронным сетям. На данном мероприятии было предложено большое количество нейросетевых систем управления [10].

Однако нейросетевые подходы имеют целый ряд серьезных недостатков. Физический смысл полученной нейронной сети, которая осуществляет управление, не является четким для понимания (особенно для инженеров по автоматизации). Также возникают вопросы об устойчивости такой системы управления. Поэтому в настоящее время, даже несмотря на их недостатки ПИД-контроллеры наиболее широко распространены. Они требуют экспертной настройки при запуске системы, более того во время работы из-за изменяющихся внешних условий их также необходимо подстраивать для получения наилучшего результата. ПИД-контроллеры плохо подходят для управления нелинейными системами, для условий быстро меняющихся параметров объекта управления и очень требовательных к точности управления систем. Например, для процесса пастеризации очень важен точный температурный диапазон нагрева молока, так как это напрямую влияет на качество полученного продукта (даже кратковременный перегрев может быть опасен).

ПИД-контроллер может быть подстроен во время работы за счет изменения его коэффициентов. Существует большое количество схем таких самонастраивающихся ПИД-контроллеров. В данной работе мы использовали для настройки коэффициентов нейронную сеть (НС).

Такие нейро-ПИД-контроллеры в настоящее время используются для построения различных систем управления. Рис. 1 поясняет базовые идеи данного подхода.

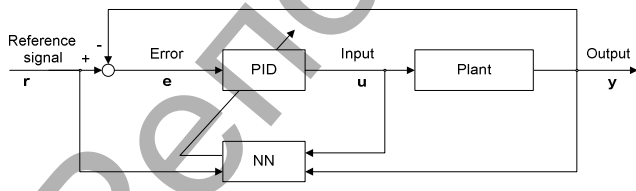


Рис. 1. Общая схема нейро-ПИД контроллера

Общая схема выглядит так – две многослойные нейронные сети (МНС), идентификатор системы (НС1) и настройщик коэффициентов ПИД (НС2) используются для «тонкой» подстройки ПИД [3, 6, 14]. Также используются рекуррентные или RBF нейронные сети [12]. Интересный подход используется в работе Ниша Джха [7] – весовые коэффициенты выходного слоя многослойного перцептрона соответствуют коэффициентам ПИД.

Для обучения НС обычно используется алгоритм обратного распространения ошибки (BPalgorithm) [3] или его модификации. Для поиска начальных значений порогов и весовых коэффициентов Сигеру Омату [3] предложил использовать генетические алгоритмы (GA). Для симуляции систем управления широко применяется пакеты MATLAB и Simulink, в некоторых работах можно встретить реализацию в качестве программного модуля для персонального компьютера (PC) [7].

Нейро-ПИД-контроллер. Общая структура самонастраивающегося нейро-ПИД-контроллера показана на рис. 2, где выходы нейронной сети – пропорциональный (K_P), интегральный (K_I) и дифференциальный (K_D) коэффициенты. Вход же определяются в зависимости от конкретной задачи управления.

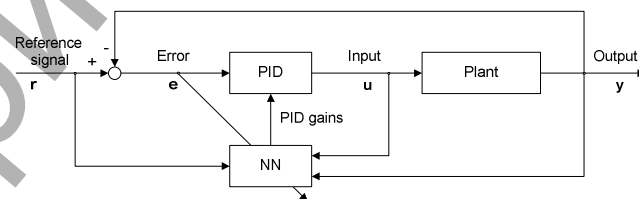


Рис. 2. Схема ПИД с самонастройкой коэффициентов

Для построения модели нейро-ПИД-контроллера опишем алгоритм функционирования нейронной сети. Вход нейрона j определяется как net_j и равен

$$net_j = \sum_{i=1}^n w_{ij} o_i + \theta_j,$$

где θ_j – пороговое значение. Выход нейрона j определяется по следующей формуле

$$o_j = f(net_j), \quad f(x) = \frac{1}{1 + e^{-x}}.$$

Для обучения будем использовать алгоритм ВР.

ПИД-контроллер в дискретном времени описывается следующим выражением:

$$u_n = u_{n-1} + K_P(e_n - e_{n-1}) + K_I e_n + K_D(e_n - 2e_{n-1} + e_{n-2}),$$

где K_P , K_I и K_D – пропорциональный, интегральный и дифференциальный коэффициенты соответственно, u_n определяет вход объекта управления в момент tnT и e_n – ошибка между желаемым значением выхода r_n и реальным, то есть

$$e_n = r_n - y_n.$$

T_0 пределяет единичный интервал времени.

Иванюк Дмитрий Сергеевич, старший преподаватель кафедры интеллектуальных информационных технологий Брестского государственного технического университета.
Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.