

У статті надано опис основних особливостей архітектур нейронних мереж, що використовуються для передбачення часових рядів, вказані їх основні переваги та недоліки. Сформульовано вимоги до вибірок, що використовуються для навчання

Ключові слова: нейронні мережі, передбачення часових рядів

В статье описаны основные особенности различных нейросетевых структур, использующихся для предсказания временных рядов, показаны их основные достоинства и недостатки. Сформулированы требования к обучающим выборкам

Ключевые слова: нейронные сети, предсказание временных рядов

The article describes the main features of different neural structures that are used for time series prediction and shows their main advantages and disadvantages. The requirements to the training set are formulated

Keywords: neural networks, time series prediction

МЕТОДЫ ПРОГНОЗИРОВАНИЯ ПАРАМЕТРОВ ТЕЛЕ- КОММУНИКАЦИОННЫХ СИСТЕМ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

К. А. Овчинников

Аспирант

Кафедра телекоммуникационных систем

Харьковский национальный университет

радиоэлектроники

пр. Ленина, 14, г. Харьков, Украина, 61166

Контактный тел.: 096-828-83-39

E-mail: vonhaimek@gmail.com

1. Введение

Современные телекоммуникационные сети характеризуются значительным количеством элементов, большими объемами обрабатываемого трафика и множеством предоставляемых сервисов. Для эффективного функционирования таких сложных систем, как ТКС, необходима система автоматического управления, позволяющая гибко реагировать на изменения показателей трафика и качества обслуживания. Важной составляющей частью любой системы управления является блок прогнозирования состояния контролируемых параметров, который позволяет частично преодолеть устаревание статистической информации о состоянии сети и применять превентивные управляющие воздействия. Существует множество подходов для анализа и прогнозирования временных рядов (последовательности значений наблюдаемого параметра): метод Винтерса, ARMA-модель, ARIMA-модель и др. Значительный интерес привлекает прогнозирование при помощи нейронных сетей, дающее многообещающие результаты.

2. Анализ литературных данных и постановка проблемы

В последнее время наблюдается рост числа публикаций, посвященных оценке и предсказанию различных экономических показателей при помощи нейронных сетей (НС), что связано с развитием математического аппарата НС, появлением эффективных алгоритмов обучения и высокими показателями

качества формируемых НС предсказаний [1]. По результатам наблюдений формируется временной ряд значений некоторого параметра, на основании которого выбирается множество обучающих выборок для НС. Обученная нейронная сеть используется для предсказания изменений наблюдаемого параметра на основании текущих измеренных значений. Предлагается использовать данный подход для прогнозирования изменений параметров ТКС с целью повышения эффективности принимаемых управляющих воздействий (например, для борьбы с ростом задержек и потерь при перегрузке маршрутов передачи трафика).

3. Цель и задачи исследования

Целью данной работы является анализ возможностей различных нейросетевых структур в области предсказания значений временных рядов и формулирование рекомендаций к их применению для задач прогнозирования параметров ТКС.

При решении задач прогнозирования роль нейронной сети состоит в предсказании будущей реакции системы по ее предшествующему поведению. Обладая информацией о значениях переменной x в моменты, предшествующие прогнозированию $x(k-1)$, $x(k-2)$, ..., $x(k-N)$, сеть вырабатывает решение, каким будет наиболее вероятное значение последовательности $\hat{x}(k)$ в текущий момент k . Для адаптации весовых коэффициентов сети используются фактическая погрешность прогнозирования $\varepsilon = x(k) - \hat{x}(k)$ и значения этой погрешности в предшествующие моменты времени [2].

4. Основные типы НС

Для решения задачи прогнозирования могут применяться следующие типы сетей: многослойный перцептрон (Multilayer perceptron, MLP), радиально-базисная сеть (Radial Basis Function, RBF), обобщенно-регрессионная сеть (General Regression Neural Network, GRNN) и сеть Вольтерра.

Многослойный перцептрон

При использовании многослойного перцептрона с двумя скрытыми слоями (рис. 1) нейронная сеть строит отображение множества экспериментальных данных на множество параметров (весовых коэффициентов) НС с целью получения оптимального в силу некоторого критерия выходного сигнала [3]. Это отображение имеет вид:

$$\hat{Y}(t) = F(X(t)) = f_3(w_{30} + \sum_{k=1}^K w_{3k} f_2(w_{20m} + \sum_{m=1}^M w_{2k} f_1(\sum_{i=1}^n w_{1i} x_i(t) + w_{10i}))) \quad (1)$$

где $i = 1, n$ – число нейронов входного слоя; $m = 1, M$ – число нейронов первого скрытого слоя, $k = 1, K$ – число нейронов второго скрытого слоя; f_1, f_2, f_3 – функции активации первого, второго скрытого слоев и выходного слоя соответственно; w_{10i}, w_{20m}, w_{30} – начальное возбуждение i -го нейрона первого скрытого слоя, m -го нейрона второго скрытого слоя и выходного слоя соответственно; w_{1i}, w_{2m}, w_{3k} – весовые коэффициенты (сила синаптической связи); $x_i(t)$ – i -й элемент входного вектора.

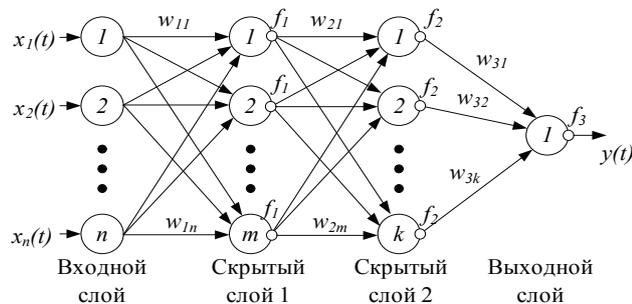


Рис. 1. Многослойный перцептрон с двумя скрытыми слоями

Функции активации одного слоя принимаются одинаковыми и зависят от типа решаемой задачи.

Для обучения MLP может применяться метод окон, предполагающий наличие двух окон W_i и W_0 с фиксированными размерами n и m соответственно. Эти окна способны перемещаться с некоторым шагом по временной последовательности наблюдаемых данных, начиная с первого элемента, и предназначены для доступа к данным временного ряда, причем первое окно W_i , получив такие данные, передает их на вход нейронной сети, а второе – W_0 – на выход. Т.о., на каждом шаге формируется обучающая пара $W_i \rightarrow W_0$, которая используется для подстройки параметров сети.

Многослойный перцептрон – одна из самых популярных и используемых НС, что обуславливается его универсальностью и развитостью математического описания. Однако одной из главных проблем при

построении MLP остается отсутствие однозначных подходов к определению связности сети и количеству скрытых слоев для различных задач.

Радиально-базисные сети

Радиально-базисные сети (RBF) широко используются в задачах классификации. Они содержат один входной слой, один скрытый слой нейронов, число которых обычно соответствует числу элементов в обучающей последовательности, и один выходной слой из одного или нескольких нейронов. Радиально-базисная нейронная сеть представлена на рис. 2.

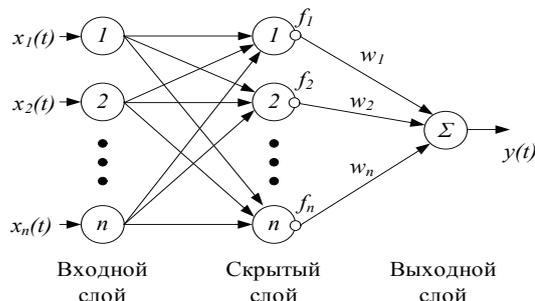


Рис. 2. Радиально-базисная нейронная сеть

Выход RBF-сети с одним выходным нейроном описывается формулой:

$$y(t) = \sum_{i=1}^m f_i(x(t)) \cdot w_i \quad (2)$$

где $\{f_i\} \in F, i = 1, m$ – множество функций активации; w_i – синаптические веса для каждой радиальной функции f_i . Функция активации может быть определена как гауссова:

$$f = \exp\left(-\frac{\|\bar{x} - \bar{c}\|^2}{r}\right),$$

где \bar{c} – центральная точка функции f , r – ее радиус; \bar{x} – вектор входных значений.

Обучение радиально-базисных сетей происходит при помощи кластеризации. Каждый кластер имеет свой центр, который может быть выбран как центр радиально-базисной функции. Существует множество алгоритмов для определения центров кластеров: метод k -средних, метод нечетких s -средних, LBG (алгоритм Linde–Buzo–Gray) и др. Алгоритмы кластеризации вычисляют центры кластеров таким образом, чтобы суммарная квадратическая ошибка представления данных на m центрах кластеров была минимальной.

Для корректировки весов используется метод наименьших квадратов или метод псевдообратных (сингулярное разложение).

RBF-сети имеют как ряд достоинств (компактность, быстрая обучаемость), так и недостатков. Например, неумение сетей RBF экстраполировать свои выводы за область известных данных приводит к падению отклика сети до нуля при удалении от обучающего множества [4].

Обобщенно-регрессионная сеть (GRNN)

Обобщенно-регрессионная сеть реализует методы ядерной аппроксимации [5]. В задачах регрессии выход сети может рассматриваться как ожидаемое значение модели в данной точке пространства входов. Это

ожидаемое значение связано с плотностью вероятности совместного распределения входных и выходных данных.

В точку расположения каждого обучающего наблюдения помещается гауссова ядерная функция:

$$K(x, x_i) = \frac{1}{(\sigma^2 \sqrt{2\pi})^d} \exp\left(-\frac{D^2(x, x_i)}{2\sigma^2}\right),$$

где σ – СКО вдоль каждой оси; d – количество размерностей в точке x ; $D(x, x_i)$ – расстояние между точкой x и экземпляром x_i . Считается, что каждое наблюдение свидетельствует о некоторой уверенности в том, что поверхность отклика в данной точке имеет определенную высоту, и эта уверенность убывает при отходе в сторону от точки. GRNN-сеть копирует внутрь себя все обучающие наблюдения и использует их для оценки отклика в произвольной точке. Окончательная выходная оценка сети получается как взвешенное среднее выходов по всем обучающим наблюдениям, где величины весов отражают расстояние от этих наблюдений до той точки, в которой производится оценивание (и, таким образом, более близкие точки вносят больший вклад в оценку).

Структура нейронной сети GRNN представлена на рис. 3.

GRNN-сеть имеет два скрытых слоя: слой радиальных элементов и слой элементов, которые формируют взвешенную сумму для соответствующего элемента выходного слоя. В выходном слое определяется взвешенное среднее путем деления взвешенной суммы $u_j(t) = \sum_{i=1}^n v_j(t)$, где $v_j(t)$ – выход j -го нейрона RBF слоя, на сумму весов $v_0 = \sum_{i=1}^n w_i$. В качестве радиальной функции применяется функция Гаусса. Таким образом, окончательный прогноз определяется как:

$$y(t) = \frac{u(t)}{v_0}. \tag{3}$$

Достоинством сети GRNN можно считать определенность структуры: сеть фактически вмещает в себя все обучающие данные. С другой стороны, такая структура НС и является ее самым большим недостатком: при большом объеме обучающих данных скорость работы сети падает, иногда очень существенно, по причине заметного увеличения сложности архитектуры.

Выходное значение сети имеет вероятностный смысл, поэтому его легче интерпретировать. При небольшом объеме входных данных сеть очень быстро обучается.

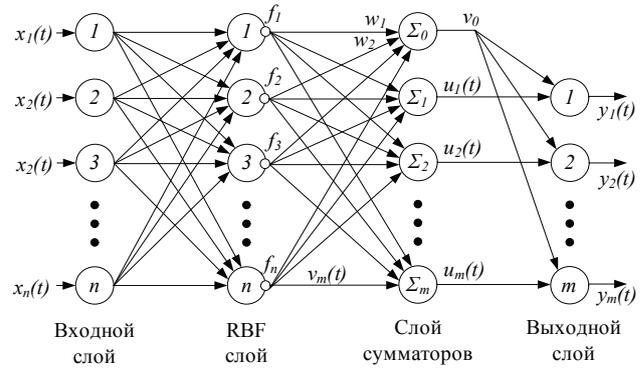


Рис. 3. Обобщенная структура сети GRNN

Нейронная сеть Вольтерра

Нейронная сеть Вольтерра – это динамическая сеть для нелинейной обработки последовательности сигналов, задержанных относительно друг друга. Обобщенный вид сети Вольтерра представлен на рис. 4. Возбуждением для сети в момент времени t служит вектор $\vec{x} = (x_t, x_{t-1}, \dots, x_{t-L})$, где L – количество единичных задержек, а $(L+1)$ означает длину вектора. В соответствии с определением ряда Вольтерра выходной сигнал y генерируется по формуле (4):

$$y(t) = \sum_{i_1=1}^L w_{i_1} x(t-i_1) + \sum_{i_1=1}^L \sum_{i_2=1}^L w_{i_1 i_2} x(t-i_1) x(t-i_2) + \dots + \sum_{i_1=1}^L \dots \sum_{i_k=1}^L w_{i_1 i_2 \dots i_k} x(t-i_1) x(t-i_2) \dots x(t-i_k) \tag{4}$$

Здесь x обозначает входной сигнал, а веса $w_i, w_{i,j}, \dots, w_{i,j,k}$ и т.д., называемые также ядрами Вольтерра, соответствуют реакциям высших порядков, порядок данного полинома K называется степенью ряда Вольтерра. В случае адаптации реакции системы Вольтерра к заданной последовательности значений необходимо определить соответствующую целевую функцию, например $E = 0.5[y(t) - d(t)]^2$ и минимизировать её значение с

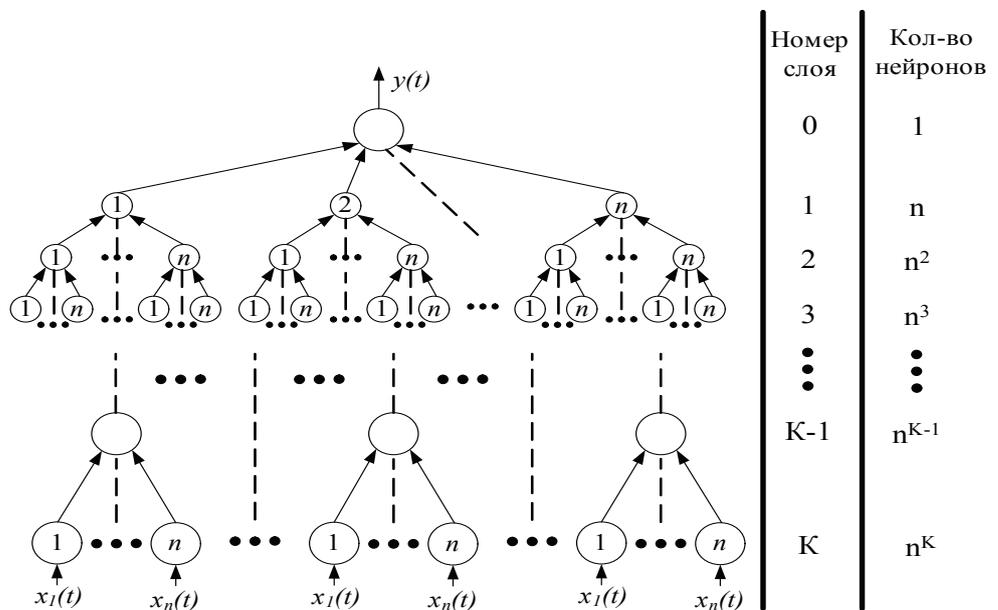


Рис. 4. Обобщенный вид сети Вольтерра

использованием универсальных способов оптимизации нейронных сетей [6].

На данном рисунке p – количество переменных обучающей выборки, т.е. размерность вектора входных параметров $[x_1, \dots, x_n]$; k – количество заданных слоев сети. Как видно из рисунка, сеть Вольтерра геометрически возрастает при увеличении количества переменных и количества слоев. Количество нейронов в слое равно p^s , где s – номер слоя.

Структуры Вольтерра дают неплохие результаты при прогнозировании временных рядов, однако процесс обучения этих сетей достаточно трудоемкий и некачественные обучающие выборки могут привести к значительному снижению точности предсказания.

5. Формирование обучающей выборки

Для обеспечения качественного обучения, а также достоверного прогноза обучающая выборка должна отвечать следующим требованиям:

1. Представительность. В общем случае для задач классификации под представительностью понимается наличие достаточного количества примеров, описывающих все многообразие распознаваемых классов. Для прогнозирующих НС под представительностью ОБ понимается то, что данные об изменении прогнозируемой величины должны быть взяты за период времени, достаточный для построения достоверного прогноза.

2. Равномерность: количество обучающих наборов в классах должно быть соизмеримо. Для оценки неравномерности обучающей выборки используется среднеквадратическое отклонение количества обучающих наборов в классе:

$$R_{ОВ} = \sigma(N_{ОН}) .$$

3. Непротиворечивость. Под непротиворечивостью понимается отсутствие в обучающей выборке одинаковых объектов, принадлежащих к различным классам. Оценка противоречивости выборки [7]:

$$C_{ОВ} = \frac{1}{n} \sum_{i=1}^n C_{ij} ,$$

где C_{ij} – парная противоречивость (противоречивость двух обучающих наборов, i -го и j -го, принадлежащих соответственно классам A_m и A_n):

$$C_{ij} = \frac{\Delta A_{mn}}{\Delta a_{ij} + \Delta A_{mn}} ,$$

где $\Delta A_{mn} = |A_m - A_n|$ – расстояние между центроидами классов m и n ; Δa_{ij} – расстояние между объектами этих классов.

Выводы

Выбор типа нейронной сети для прогнозирования значения параметров ТКС обуславливается особенностями самих сетей. Так, строя систему прогнозирования при помощи многослойного персептрона, необходимо дополнительно исследовать эффективность сетей с различным количеством слоев.

Радиально-базисные сети обладают высоким быстродействием в силу простоты своей структуры, что

очень важно для своевременного принятия решения, но при этом необходимо особенно тщательно подходить к формированию обучающей выборки.

При использовании обобщенно-регрессионных сетей необходимо соблюдать баланс между объемом знаний, хранящимся в сети, и полнотой обучающей выборки.

Эффективность сетей Вольтера в общем случае значительно превосходит таковую MLP, но при этом обладает более сложной структурой. Тем не менее, при тщательном выборе параметров эти сети эффективно справляются с узкоспециализированными задачами, где MLP дают худшие результаты.

Ключевым аспектом построения НС является обучение, поэтому при формировании выборки необходимо соблюдать требования к полноте, равномерности и непротиворечивости.

Литература

1. Крючин, О.В. Прогнозирование временных рядов с помощью искусственных нейронных сетей и регрессионных моделей на примере прогнозирования котировок валютных пар / О.В. Крючин, А.С. Козадаев, В.П. Дудаков. – Электронный научный журнал «Исследовано в России», 2010. – Т.13. – С. 354-362. – ISSN 1819-4192. – Режим доступа: <http://zhurnal.ape.relarn.ru/articles/2010/030.pdf>.
2. Горбань, А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей. – Сибирский журнал вычислительной математики, 1998. – Т.1, №1. – С.12-24.
3. Султанов, А.Х. Использование нейросетевых моделей для прогнозирования интенсивностей самоподобного трафика телекоммуникационной сети / А.Х. Султанов, А.А. Габдрахманов, Н.Т. Габдрахманова. – Уфа: Вестник УГАТУ, 2011. – № 2(42), Т. 15. – С. 93-98.
4. Amin, S.M. Traffic prediction and management via RBF neural nets and semantic control / S. Massoud Amin, E.Y. Rodin, A-P. Liu, K. Rink. – Computer-Aided Civil and Infrastructure Engineering, 1998. – №13. – pp. 315-327.
5. Солдатова, О.П. Применение нейронных сетей для решения задач прогнозирования. / О.П. Солдатова, В.В. Семенов. – Электронный научный журнал «Исследовано в России», 2006. – Т.9. – С. 1270-1276. – ISSN 1819-4192. – Режим доступа: <http://zhurnal.ape.relarn.ru/articles/2006/136.pdf>.
6. Крючин, О.В. Прогнозирование временных рядов с использованием нейронных сетей типа Вольтерри и сравнение показателей с многослойным персептроном / О.В. Крючин, О.В. Кондратов. – Электронный научный журнал «Исследовано в России», 2010. – Т.13. – ISSN 1819-4192. – С. 953-967. – Режим доступа: <http://zhurnal.ape.relarn.ru/articles/2010/082.pdf>.
7. Олешко, Д.Н. Построение качественной обучающей выборки для прогнозирующих нейросетевых моделей / Д.Н. Олешко, В.А. Крисилов, А.А. Блажко. – Штучний інтелект, 2004. – № 3. – С. 567-573.
8. Хайкин, С. Нейронные сети: полный курс / Саймон Хайкин; пер. с англ.; 2-е издание. – М.: Издательский дом «Вильямс», 2006. – 1104 с.: ил. – Парал. тит. англ. – ISBN 5-8459-0890-6 (рус.).