

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

УДК 681.324

Савицкий Юрий Викторович

МЕТОДЫ И СРЕДСТВА ГЕНЕРАЦИИ ИСКУССТВЕННЫХ
НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ПРОГНОЗИРОВАНИЯ

Специальность 05.13.15 – вычислительные машины и системы

Автореферат диссертации
на соискание ученой степени кандидата технических наук

Минск - 2000

Работа выполнена в Брестском государственном техническом университете

Научный руководитель –
кандидат технических наук, доцент Головки В. А.

Официальные оппоненты:
доктор технических наук, профессор Садыхов Р. Х.
кандидат технических наук, с.н.с. Дудкин А. А.

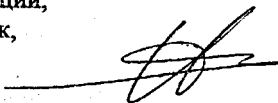
Опонирующая организация – Белорусский государственный университет

Защита состоится 28 декабря 2000 года в 15.00 часов на заседании совета по защите диссертаций Д 02.15.04 в Белорусском государственном университете информатики и радиоэлектроники по адресу: 220027, Минск, ул. П. Бровки, 6 (239-89-89).

С диссертацией можно ознакомиться в библиотеке Белорусского государственного университета информатики и радиоэлектроники.

Автореферат разослан ” 14 “ ноября 2000 г.

Ученый секретарь
совета по защите диссертаций,
кандидат технических наук,
доцент



Б. В. Никульшин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы диссертации. Интенсивные исследования в области нейронных сетей и полученные теоретические и практические результаты привели к возникновению принципиально новых технологий создания интеллектуальных систем в различных областях знаний. Способность нейронных сетей к самоорганизации, адаптации и обучению позволяет создавать качественно новые прикладные системы, обладающие рядом преимуществ по сравнению с традиционными: меньшей процедурной сложностью, толерантностью к ошибкам, высокой точностью решения сложных задач. Одним из фундаментальных свойств нейронных сетей является способность их после обучения к обобщению и пролонгации результатов. Это создает предпосылки для создания на базе их различного рода прогнозирующих систем в разных сферах применения: экономика, финансы, медицина, экология, метеорология, техника и т.д.

Нейронные сети позволяют организовывать высокоадаптивные нелинейные прогнозные модели, позволяющие во многих случаях моделировать и выполнять прогнозы сложных временных зависимостей с большей эффективностью и точностью по сравнению с известными статистическими. Об актуальности проблем нейросетевого прогнозирования свидетельствует возрастающий поток публикаций и конференций. Большинство этих работ ориентировано на исследование базовых аспектов функционирования и обучения архитектур многослойных нейронных сетей с прямыми и рекуррентными связями, являющимися доминирующими в нейронных системах прогнозирования. Перспективность и популярность данного типа нейронных сетей обусловлена большим количеством аналитических и практических результатов применения их в реальных задачах. Использование для обучения алгоритма обратного распространения ошибки, базирующегося на технике градиентного спуска, позволяет выполнять точную настройку параметров сети и тем самым обеспечить высокое качество решения задач.

Вместе с тем детальный анализ вышеуказанных работ выявил ряд недостатков в базовых процедурах синтеза нейронных систем прогнозирования. Ключевой проблемой являются недостатки алгоритмов обучения нейронной сети. При этом классический алгоритм обратного распространения ошибки требует трудоемкого процесса подгонки параметров обучения экспериментальным путем, что в большинстве случаев не приводит к оптимальному обучению сети. Вследствие этого такие обучающие процедуры обладают большой временной сложностью и неустойчивостью. Неустойчивость обучения заключается в том, что сходимости алгоритма обратного распространения ошибки сильно зависит от режима обучения (определяемого константами обучения), а также от начальной инициализации синаптических связей. Вследствие этого велика вероятность попадания в локальные минимумы целевой функции. Это

приводит к неприемлемо большой погрешности обучения, а, следовательно, к большой погрешности прогноза. Кроме этого, при использовании в архитектуре нейронной сети прогнозирования разнотипных функций активации нейроэлементов процедура подгонки параметров обучения является еще более сложной из-за различия в передаточных характеристиках нейроэлементов сети. Другой важной проблемой является отсутствие эффективных методов синтеза базовых архитектур нейронных сетей прогнозирования. Под архитектурным синтезом здесь понимается выбор количества слоев, количества нейронных элементов в слое и функции активации нейронов каждого слоя. Существующие теоретические результаты носят рекомендательный характер и ориентированы на иные практические приложения.

Связь работы с научными программами, темами. Результаты диссертационной работы связаны с выполнением госбюджетных НИР, проводившихся в Брестском государственном техническом университете по темам "Проектирование нейронных сетей и самоорганизующихся схем" в 1995-1996 г. г. (№ ГР1995739), "Разработка и исследование интеллектуальных нейронных систем для многофакторного прогнозирования" в 1997-1998 г. г. (№ ГР1997979), а также в рамках международного проекта программы INTAS по теме "Разработка интеллектуальной сенсорной инструментальной структуры" в 1997-2000 г. г. (INTAS-97-0606).

Цель и задачи исследования. Цель диссертационной работы заключается в разработке методов и средств синтеза прогнозирующих нейронных сетей и эффективных алгоритмов их обучения для повышения точности прогноза. Для достижения цели работы решаются следующие задачи:

- разработать методы и алгоритмы обучения базовых архитектур нейронных сетей прогнозирования с различными функциями активации нейроэлементов, обладающие высокой точностью, стабильностью и позволяющие уменьшить временную сложность процесса обучения;
- разработать метод и алгоритм начальной инициализации синаптических связей нейроэлементов в архитектурах нейронных сетей прогнозирования;
- исследовать эффективность методов и алгоритмов начальной инициализации и обучения нейронных сетей с прямыми и обратными связями;
- разработать алгоритмы архитектурного синтеза нейронных систем прогнозирования временных рядов;
- исследовать свойства различных функций активации нейроэлементов в архитектурах нейронных сетей для решения задач прогнозирования временных последовательностей;
- разработать нейросетевые модели прогнозирования и идентификации сложных временных последовательностей;

- разработать программное обеспечение для нейронных систем прогнозирования временных последовательностей.

Методология и методы проведенного исследования. Для решения рассматриваемых в диссертационной работе задач использовались методы теории нейронных сетей, методы безусловной оптимизации, методы математического и программного моделирования.

Научная новизна полученных результатов.

Проведен анализ различных функций активации нейроэлементов в архитектурах нейронных сетей с целью повышения точности прогноза для разных классов реальных и математических временных рядов. Предложен новый вариант нелинейной активационной функции, позволяющей улучшить свойства базовых архитектур нейронных сетей прогнозирования.

Разработаны адаптивные алгоритмы обучения нейронных сетей для различных функций выходной активности нейронных элементов, позволяющие вычислять шаг обучения на каждой итерации обучения, обладающие высокой сходимостью и точностью обучения и позволяющие сократить время обучения. При этом получены конечные выражения и границы для определения адаптивного шага обучения.

Разработаны алгоритмы адаптивной инициализации параметров нейронных сетей, позволяющие снизить вероятность попадания сети в локальные минимумы на начальной стадии процесса обучения и тем самым увеличить стабильность и точность обучения.

Разработан метод последовательного послойного обучения гетерогенных нейронных сетей с прямыми и обратными связями. Данный метод позволяет улучшить стабильность процесса обучения нейронных структур, обладающих разнотипными функциями активации нейроэлементов.

Предложен метод и алгоритм проектирования скрытого слоя базовых архитектур нейронных сетей прогнозирования, позволяющий выполнять оценку необходимого количества скрытых нейронов с целью формирования неизбыточной модели.

Практическая значимость полученных результатов состоит в том, что они обеспечивают высокое качество решения актуальных задач моделирования и прогнозирования сложных временных процессов в различных сферах применения. Основные результаты данной диссертационной работы, используются при разработке интеллектуальной сенсорной системы для прогнозирования погрешностей дрейфа сенсорных датчиков в рамках международного проекта INTAS, при выполнении госбюджетных НИР, проводимых на кафедре "ЭВМ и системы" Брестского государственного технического университета. Теоретические результаты и алгоритмы используются в учебном процессе на кафедре "ЭВМ и системы" Брестского государственного технического университета в курсах "Интеллектуальные системы принятия решений", "Структура и архитектура супер-ЭВМ".

Основные положения диссертации, выносимые на защиту, заключаются в следующем:

- предложена методика и алгоритмы обучения базовых архитектур нейронных сетей прогнозирования для различных функций выходной активности нейроэлементов. Данная методика обеспечивает более высокую сходимость, устойчивость и точность обучения по сравнению с классическим алгоритмом обратного распространения ошибки;

- предложен адаптивный алгоритм последовательного послойного обучения гетерогенных архитектур многослойных нейронных сетей с прямыми и обратными связями, применяемых в прогнозирующих нейронных системах (данный алгоритм обладает высокими параметрами обучения и исключает эмпирический подбор констант обучения и балансировки процесса обучения разнотипных слоев нейронной сети);

- исследована эффективность предложенных методов и алгоритмов обучения нейронных сетей прогнозирования;

- предложена методика адаптивной инициализации параметров нейронной сети (обеспечивает снижение вероятности попадания целевой функции в локальные экстремумы на начальных стадиях процесса обучения, что повышает точность и стабильность обучения);

- на основании результатов исследований различных функций активации нейроэлементов сформулированы требования к архитектуре нейронных сетей прогнозирования; предложен новый вариант функции выходной активности нейронных элементов, позволяющей как оптимизировать процедуры обучения, так и снизить ограничения традиционных нейросетевых моделей прогнозирования;

- предложен метод синтеза скрытого слоя нейронных сетей прогнозирования, позволяющий уменьшить структурную избыточность сети и исключить произвольный выбор архитектуры, обеспечивая при этом необходимую точность обучения и прогнозирования.

Личный вклад соискателя. Основные положения диссертации получены соискателем лично и отражены в 35 публикациях. Соавтором в работах [1-7, 11-31; 33, 35] является научный руководитель к.т.н., докторант ИТК НАН РБ В.А. Головки, которому принадлежит постановка задачи и обсуждение методов ее решения.

Апробация результатов диссертации. Основные теоретические и практические результаты работы докладывались:

- на конференции "Современные проблемы радиотехники, электроники и связи" (Минск, 1995);

- на международной конференции Third International Conference "Pattern Recognition and Information Analysis" (Минск, 1995);

- на международной конференции CMNDT-95 (Минск, 1995);

- на международной конференции International Conference on Pattern Recognition (Щецин, Польша, 1995);

- на международной конференции International Conference on Design Methodologies for Signal Processing (Щецин, Польша, 1996);

- на международной конференции International Conference on Technical Information (Тимисоара, Румыния, 1996);
- на конференции "Актуальные проблемы информатики: математическое, программное и информационное обеспечение" (Минск, 1996);
- на международной конференции Second International Conference CMNDT-98 (Минск, 1998);
- на международной конференции УПК-98 (Брест, 1998);
- на научно-технической конференции, посвященной 30-летию БПИ (Брест, 1996);
- на X научно-технической конференции "Новые технологии в машиностроении и вычислительной технике", (Брест, 1998);
- на V международной конференции "Pattern Recognition and Information Processing" PRIP'99 (Минск, 1999);
- на международной конференции "Information Technologies for Education, Science and Business" ITESB'99, (Минск, 1999);
- на международной конференции "IEEE International Workshop on Intelligent Signal Processing" WISP'99, (Будапешт, Венгрия, 1999);
- на международной конференции International Conference on Neural Networks and Artificial Intelligence ICNNAI'99 (Брест, 1999);
- на международной конференции Sixth International Conference "Advanced Computer Systems" ACS'99 (Щецин, Польша);
- на международной конференции Second International ICSC Symposium on Engineering of Intelligent Systems EIS'2000 (Пайслей, Великобритания, 2000);
- на международной конференции Second International ICSC Symposium on Neural Computation NC'2000 (Берлин, Германия, 2000);
- на международной конференции International Joint Conference on Neural Networks ICNN'2000 (Комо, Италия, 2000);
- на международной конференции Euro-International Symposium on Computational Intelligence E-ISCI (Кошице, Словакия, 2000).

Опубликованность результатов. По теме диссертации опубликовано 35 печатных работ, включая 23 статьи в сборниках трудов конференций, 5 статей в журнале, 7 тезисов докладов конференций. Общий объем публикаций – 157 страниц.

Структура и объем диссертации. Диссертация изложена на 143 страницах машинописного текста, содержит 38 рисунков, 9 таблиц и состоит из введения, четырех глав, выводов, списка литературы, включающего 133 названия, и 3 приложений.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Первая глава посвящена обзору современного состояния вопроса, анализу литературы, выбору и обоснованию цели и задач работы.

В главе рассмотрены основные методы и модели статистического и нейросетевого прогнозирования. Анализируются критерии выбора, способы построения и адаптации, а также свойства статистических моделей прогноза. Рассмотрены достоинства и недостатки наиболее распространенных классов моделей авторегрессии, скользящего среднего и их смешанных типов. Показано, что базовые статистические методы прогнозирования временных рядов ориентированы в основном на организацию линейных моделей, что представляет собой серьезное ограничение при исследовании сложных временных процессов. При этом выбор модели прогноза сильно зависит от специфики и размерности исследуемого ряда и представляет собой нетривиальную задачу индивидуального подбора, анализа и оценивания параметров модели.

Выполнен детальный анализ методов прогнозирования с использованием искусственных нейронных сетей. Исследованы применяемые в прогнозировании архитектуры нейронных сетей: характер межнейронных связей, типы функций активации нейронных элементов. Показано, что в нейросетевой технологии моделирования и прогнозирования временных рядов базовыми архитектурами являются гетерогенные многослойные нейронные сети с прямыми и рекуррентными связями с линейной активационной функцией элементов выходного слоя и логистическими активационными функциями скрытых нейронов, обучаемые по методу обратного распространения ошибки.

Показано, что классический алгоритм обратного распространения ошибки обладает рядом существенных недостатков, ограничивающих его эффективное применение в практических задачах. Имеющиеся его модификации, хотя и позволяют улучшить параметры обучения, нуждаются в подборе шага обучения опытным путем индивидуально для каждой выборки. Использование в процедурах обучения более эффективных методов оптимизации (методов сопряженных градиентов и методов второго порядка) приводит к значительному увеличению вычислительной сложности процесса обновления параметров сети в составе итерационной процедуры и сильной зависимости результата обучения от характера целевой функции и погрешностей вычислительного алгоритма.

Исследованы существующие методы структурного синтеза нейронных сетей. Установлено, что существующие методы синтеза нейронных сетей в основном ориентированы на построение моделей с прямыми связями, значительно увеличивают вычислительную сложность и ухудшают обобщающие свойства алгоритма и поэтому не могут быть эффективно использованы для выбора архитектур нейросетевых моделей прогноза.

Во второй главе выполняется анализ архитектур нейронных сетей прогнозирования и различных активационных функций нейроэлементов.

Проведенный анализ свойств различных структур нейронных сетей, применяемых для организации моделей прогноза, позволил сформулировать основные требования к нейросетевым архитектурам. Исследованы свойства сигмоидной функции активации скрытых нейроэлементов, наиболее часто используемой в архитектурах нейронных сетей прогнозирования. На основании данных исследований обоснована возможность использования логарифмической функции активации $g^{[l]}(S_j^{[l]}) = \ln((S_j^{[l]} + \sqrt{(S_j^{[l]})^2 + a}) / \sqrt{a})$, $a > 0$ в скрытых элементах j слоя l нейронных сетей прогнозирования. Данная функция, в отличие от сигмоидной, позволяет формировать выходную активность нейрона в неограниченном диапазоне значений. Это позволяет частично снять ограничения по выходу прогнозной модели, характерные для нейронных сетей с логистическими активационными функциями скрытых нейронов.

В главе разрабатываются методы и алгоритмы эффективного обучения гетерогенных нейронных сетей с прямыми и рекуррентными связями. С этой целью определена обобщенная структура слоя l нейронной сети прогнозирования, сочетающая в себе свойства рассмотренных в главе 1 базовых моделей нейронных сетей прогнозирования с прямыми и рекуррентными связями (рис. 1). Для этого определены следующие правила функционирования элементов задержки B_1 и B_2 :

$$B_1 y(\tau) = \begin{cases} y(\tau-1), & \text{если } B_1 = 1 \\ 0, & \text{если } B_1 = 0 \end{cases}, \quad B_2 y(\tau) = \begin{cases} y(\tau-1), & \text{если } B_2 = 1 \\ 0, & \text{если } B_2 = 0 \end{cases}. \quad (1)$$

Тогда общие правила функционирования нейронов слоя l могут быть представлены в виде:

$$\begin{cases} S_j^{[l]}(\tau) = \sum_{i=1}^{N^{[l-1]}} w_{ij}^{[l]} y_i^{[l-1]}(\tau) + \sum_{k=1}^{N^{[l]}} w_{kj}^{[l]} B_1 y_k^{[l]}(\tau) + \sum_{m=1}^{N^{[l+1]}} w_{mj}^{[l]} B_2 y_m^{[l+1]}(\tau) - w_{bj}^{[l]}, \\ y_j^{[l]}(\tau) = g^{[l]}(S_j^{[l]}(\tau)), \quad j = 1, \dots, N^{[l]}. \end{cases} \quad (2)$$

Очевидно, что различные режимы работы элементов задержки B_1 , B_2 позволят организовывать следующие архитектуры слоя l нейронной сети прогнозирования:

- 1) $B_1 = 0$, $B_2 = 0$ - нейронная сеть с прямыми связями;
- 2) $B_1 = 0$, $B_2 = 1$ - нейронная сеть Джордана;
- 3) $B_1 = 1$, $B_2 = 0$ - нейронная сеть Элмана;
- 4) $B_1 = 1$, $B_2 = 1$ - комбинированная рекуррентная сеть Джордана-Элмана.

Применение правил функционирования сети (1), (2) позволяет разработать общую методику адаптивного обучения для всех вышеуказанных вариантов прогнозирующих нейронных сетей. Суть ее состоит в следующем. Согласно методу градиентного спуска, итерационное изменение весовых коэффициентов и порогов элементов для каждого слоя нейронной сети в алгоритме обратного распространения ошибки выполняется по следующим правилам:

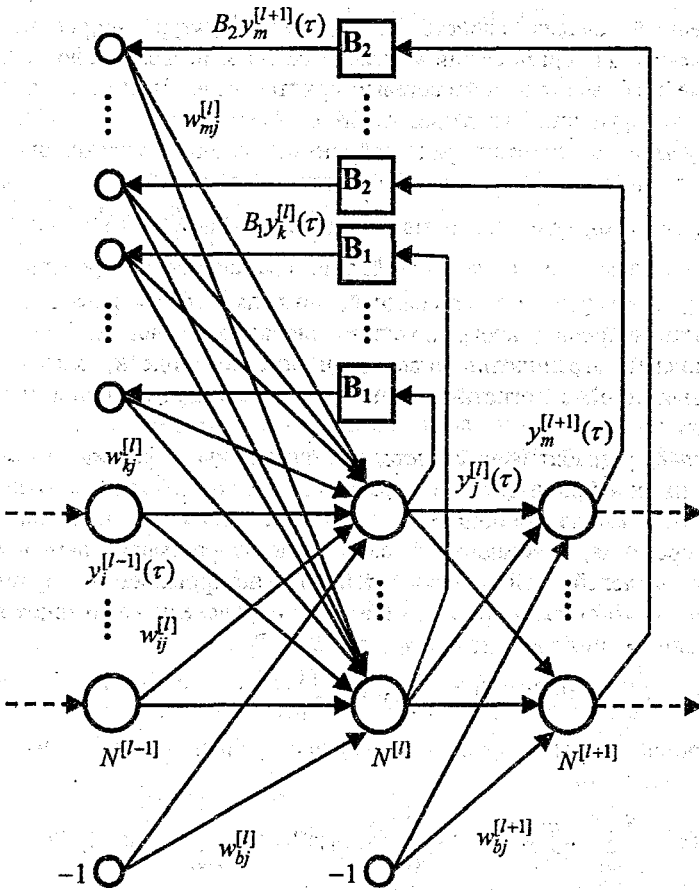


Рис. 1. Фрагмент нейронной сети с различными типами связей

$$\begin{cases} \Delta w_{ij}^{[l]}(t) = -\alpha \frac{\partial E^p(t)}{\partial w_{ij}^{[l]}(t)} = -\alpha \gamma_j^{[l],p}(t) (g^{[l]})'(S_j^{[l],p}(t)) y_i^{[l-1],p}(t), \\ \Delta w_{bj}^{[l]}(t) = -\alpha \frac{\partial E^p(t)}{\partial w_{bj}^{[l]}(t)} = \alpha \gamma_j^{[l],p}(t) (g^{[l]})'(S_j^{[l],p}(t)), \end{cases}$$

где $\partial E^p(t) / \partial w_{ij}^{[l]}(t)$, $\partial E^p(t) / \partial w_{bj}^{[l]}(t)$ - частные производные функции ошибки нейронных связей, вычисляемые на каждой итерации обучения для каждого эталона p , $p = \{1, \dots, P\}$;

$\gamma_j^{[l],p}(t)$ - ошибка нейроэлемента j , определяемая как:

$$\gamma_j^{[L],p}(t) = y_j^{[L],p}(t) - D_j^p \text{ - для нейронов выходного слоя;}$$

$$\gamma_i^{[l-1],p}(t) = \sum_{j=1}^{N^{[l]}} \gamma_j^{[l],p}(t) (g_j^{[l]})' (S_j^{[l],p}(t)) w_{ij}^{[l]}(t) \quad \text{— для нейронов скрытых}$$

слоев сети;

α - шаг модификации весовых коэффициентов;

E^p - среднеквадратичная ошибка нейронной сети для эталона p :

$$E^p(t) = \frac{1}{2} \sum_{j=1}^{M^{[L]}} (\gamma_j^{[L],p}(t) - D_j^p)^2$$

D_j^p — эталонное выходное значение j -го нейрона.

Как было показано в главе 1, ключевым моментом задачи проектирования нейронных систем прогнозирования является проблема эффективного обучения нейронных сетей. При этом используемые для обучения алгоритм обратного распространения ошибки и его модификации обладают принципиальным недостатком, связанным с необходимостью выбора констант обучения. Поэтому первой задачей, которую необходимо решить при создании нейронных систем прогнозирования, является разработка методов определения адаптивного шага обучения с целью минимизации целевой функции обучения. С этой целью в работе сформулирована следующая теорема, позволяющая определять адаптивный шаг обучения на каждой обучающей итерации для каждого эталона с целью минимизации ошибки нейроэлементов сети.

Теорема 1. Для нейронной сети, функционирующей по правилам (1), (2) и обучаемой по методу обратного распространения ошибки, адаптивный шаг обучения нейроэлементов слоя l с целью минимизации ошибки обучения сети на каждой последующей итерации обучения может быть определен согласно выражению:

$$\alpha^{[l],p}(t) = \frac{\sum_{j=1}^{N^{[l]}} (\gamma_j^{[l],p}(t))^2 g_j^{[l]'}(S_j^{[l],p}(t))}{g_j^{[l]'}(0) \sum_{j=1}^{N^{[l]}} (\gamma_j^{[l],p}(t))^2 (g_j^{[l]'}(S_j^{[l],p}(t)))^2 + \sum_{i=1}^{N^{[l-1]}} (\gamma_i^{[l-1],p}(t))^2 + \sum_{k=1}^{N^{[l]}} (B_k \gamma_k^{[l],p}(t))^2 + \sum_{m=1}^{N^{[l+1]}} (B_m \gamma_m^{[l+1],p}(t))^2} \quad (3)$$

На основании данной теоремы получены аналитические выражения для вычисления шага обучения для линейной, сигмоидной и логарифмической активационных функций нейроэлементов, применяемых в архитектурах нейронных сетей прогнозирования.

Установлено, что точность вычисления адаптивного шага по вышеприведенной методике обуславливается погрешностью разложения нелинейной функции в ряд Тейлора, которое используется в выражении (3). Поэтому необходимым условием эффективного применения адаптивного шага в практических задачах обучения нейронных сетей является определение допустимых границ шага обучения. С этой целью исследованы условия протекания процесса обучения для типичного нейроэлемента сети, обладающего сигмоидной функцией активации. Размер шага обучения совместно с градиентами ошибок определяет величину модификации

весовых коэффициентов нейрона. Если выбранное значение шага слишком велико, то это может привести к формированию выходной активности элемента близкой к 0 или 1. Это в свою очередь приведет к значительному уменьшению градиентов ошибок, так как в этом случае производная сигмоидной функции будет стремиться к 0. В результате, несмотря на возможно высокую ошибку элемента, обучающая способность рассматриваемого нейрона будет снижена вследствие малого размера градиентов, в которых в качестве сомножителей фигурирует производная активационной функции. Поэтому в качестве критерия для оценивания верхней границы шага обучения целесообразно использовать минимально допустимый размер производной активационной функции, обозначаемый в рамках следующих теорем как ε .

Теорема 2. Для сигмоидной активационной функции нейроэлементов скрытого слоя нейронной сети, функционирующей по правилам (1), (2), верхняя граница шага обучения определяется в соответствии с выражениями:

$$\alpha^{p,[l]}(t) < \frac{S_j^{p,[l]}(t) - \ln \frac{1 + \sqrt{1 - 4\varepsilon}}{1 - \sqrt{1 - 4\varepsilon}}}{\Delta}, \text{ если } \Delta < 0, \quad (4)$$

$$\alpha^{p,[l]}(t) < \frac{S_j^{p,[l]}(t) - \ln \frac{1 - \sqrt{1 - 4\varepsilon}}{1 + \sqrt{1 - 4\varepsilon}}}{\Delta}, \text{ если } \Delta > 0, \quad (5)$$

где

$$\Delta = \sum_{i=1}^{N^{[l-1]}} y_i^{p,[l-1]} \frac{\partial E^p}{\partial w_{ij}^{[l]}} + \sum_{k=1}^{N^{[l]}} B_1 y_k^{p,[l-1]} \frac{\partial E^p}{\partial w_{kj}^{[l]}} + \sum_{m=1}^{N^{[l+1]}} B_2 y_m^{p,[l-1]} \frac{\partial E^p}{\partial w_{mj}^{[l]}} - \frac{\partial E^p}{\partial w_{bj}^{[l]}} \quad (6)$$

Теорема 3. Для логарифмической активационной функции нейроэлементов скрытого слоя нейронной сети, функционирующей по правилам (1), (2), верхняя граница шага обучения определяются в соответствии с выражениями:

$$\alpha^{p,[l]}(t) < \frac{S_j^{p,[l]}(t) + \sqrt{\frac{1}{\varepsilon^2} - a}}{\Delta}, \text{ если } \Delta > 0 \quad (7)$$

$$\alpha^{p,[l]}(t) < \frac{S_j^{p,[l]}(t) - \sqrt{\frac{1}{\varepsilon^2} - a}}{\Delta}, \text{ если } \Delta < 0, \quad (8)$$

где Δ определяется выражением (6).

Таким образом, теоремы 2, 3 позволяют определять допустимые границы шага обучения для рассмотренных в данной диссертационной работе функций активации нейроэлементов и могут быть использованы для корректировки адаптивного шага обучения с целью сохранения состояния активности обучения нейронов сети.

В главе исследуется задача выбора начальных значений синаптических связей и порогов нейроэлементов при инициализации сети. От распределения этих значений сильно зависит исход процедуры обучения, а следовательно, и возможность получения требуемой модели в целом. Так, при неудачном выборе начальных параметров нейроэлементов обучение проходит крайне медленно или же вообще прекращается. Это связано с попаданием целевой функции обучения на начальной стадии процесса обучения в локальные минимумы. Это является следствием описанной выше ситуации, когда в процессе инициализации сети, несмотря на возможно большую ошибку нейроэлементов, производная активационной функции близка к 0, и поэтому нейроэлемент оказывается заблокированным от обучения (состояние пассивности при обучении). Как правило, начальные параметры нейронной сети выбираются из некоторого диапазона случайных равномерно распределенных чисел, границы которого определяются эмпирически. Вероятность перехода нейроэлемента j слоя l в пассивное состояние обусловлена значением взвешенной суммы его входной активности $S_j^{[l]}$, которая определяется следующими параметрами элемента:

- 1) Уровнем входной активности нейроэлемента;
- 2) Количеством синаптических связей нейроэлемента;
- 3) Верхней и нижней границей диапазона инициализации весовых коэффициентов.

Таким образом, существует проблема адаптивного выбора параметров инициализации сети с учетом вышеперечисленных факторов. Для решения этой задачи в главе сформулирована следующая теорема.

Теорема 4. Для нейроэлемента j слоя l , функционирующего по правилам (1), (2), обладающего активационной функцией $g^{[l]}$, его весовые коэффициенты относительно известных входного $y_s^{[l-1]}$ и желаемого выходного $d_j^{[l]}$ значений определяются следующим выражением:

$$w_{sj}^{[l]} = \frac{1}{1 + \sum_{n=1}^{N^{[l-1]}} (y_n^{[l-1]})^2 + \sum_{k=1}^{N^{[l]}} (B_1 y_k^{[l]})^2 + \sum_{m=1}^{N^{[l+1]}} (B_2 y_m^{[l+1]})^2} y_s^{[l-1]} G^{[l]}(d_j^{[l]}), \quad (9)$$

где $G^{[l]}$ - функция, обратная активационной функции $g^{[l]}$;

в $y_s^{[l-1]}$ индекс s характеризует источник входной активности для скрытого нейроэлемента слоя l .

На основании теоремы 4 предложены аналитические выражения, определяющие правила инициализации весовых коэффициентов скрытых и выходных нейронов относительно уровней их входной активности, желаемого выхода и количества входов. Данные правила позволяют на начальной стадии процесса обучения как сохранить активность обучения нейроэлементов, так и уменьшить начальную погрешность моделей относительно обучающего множества.

Исследована эффективность алгоритмов обучения сетей с разнотипными активационными функциями нейроэлементов. На основании анализа разработан обобщенный алгоритм последовательного послойного обучения гетерогенных нейронных сетей с прямыми и рекуррентными связями, которым интегрированы предложенные в главе методы адаптивной инициализации и адаптивного обучения. Алгоритм исключает эмпирический подбор констант обучения и балансировки процесса обучения разнотипных слоев нейронной сети.

В третьей главе анализируются обобщающие свойства нейронных структур с различным количеством скрытых нейроэлементов. На основе данного анализа и рассмотренных в предыдущей главе методов адаптивного обучения предложен алгоритм проектирования скрытого слоя нейронной сети для организации нейросетевых моделей прогнозирования временных последовательностей. Предлагаемый алгоритм ориентирован на определение оценки необходимого количества скрытых нейронов в трехслойных гетерогенных нейронных сетях с прямыми связями и рекуррентными связями, для которых разработаны методы адаптивного обучения и адаптивной инициализации. В качестве функции нелинейного преобразования нейронов скрытого слоя могут быть применены сигмоидная или логарифмическая функции, а в выходном слое — линейная функция. Минимизируемой функцией алгоритма синтеза является:

$$Q_{NN} = \frac{MSE_{test}[T_{test}]}{MSE_{train}[T_{train}] + MSE_{test}[T_{test}]}, \quad (10)$$

где $MSE_{train}[T_{train}]$, $MSE_{test}[T_{test}]$ - соответственно среднеквадратичные ошибки обучения и тестирования подмножествах T_{train} , T_{test} , формируемых из элементов исходной временной последовательности.

Результатом успешного выполнения любой стадии процесса синтеза является достижение меньшего уровня погрешности Q_{NN} по сравнению с предыдущим этапом алгоритма.

Первой стадией выполнения алгоритма синтеза является определение и обучение стартовой архитектуры нейронной сети. С целью избежания структурной избыточности на начальных этапах процесса синтеза в качестве стартовой определяется модель с малым количеством скрытых нейроэлементов минимальной архитектуры: $NN = (N^{[1]}, N^{[2], \min}, 1)$, которая соответствует трехслойной нелинейной модели с $N^{[1]}$ количеством входных нейроэлементов, $N^{[2], \min}$ нелинейных нейронов скрытого слоя с сигмоидной или логарифмической функцией активации и одним линейным нейроэлементом скрытого слоя. Процесс обучения завершается при достижении заданной минимальной скорости убывания ошибки обучения $V_{E, \min}$.

Этап добавления нелинейного элемента в скрытый слой сети позволяет увеличивать нелинейные свойства сети, улучшая тем самым

аппроксимирующие способности синтезируемой модели прогноза. Очевидно, что непосредственное добавление необученного элемента является неэффективным шагом процесса синтеза, так как приведет к значительному перераспределению пространства весовых коэффициентов всех элементов сети, что эквивалентно полному переобучению всего синтезированного фрагмента сетевой архитектуры. Поэтому в главе предлагается метод последовательного добавления элемента и соответствующих синаптических связей, предполагающий его независимое обучение с целью уменьшения достигнутой на предыдущих этапах синтеза ошибки обучения.

Утверждение. Добавление в скрытый слой гетерогенной сети архитектуры $NN = (N^{[1]}, N^{[2]}, 1)$, $N^{[2]} \geq 1$, функционирующей по правилам (1), (2), нелинейного элемента, обучаемого независимо в составе сформированной архитектуры сети, уменьшает общую ошибку обучения сети в совокупности с добавленным нейроном.

После добавления нейроэлемента и соответствующих весовых связей в составе общей нейронной структуры необходимо модифицировать весовые коэффициенты таким образом, чтобы уменьшить уже достигнутую на предыдущих этапах синтеза погрешность обучения модели. Для этого фиксируются параметры (значения весов и порогов нейроэлементов) уже синтезированной сети и выполняется обучение добавленного элемента согласно общих правил адаптивного обучения, рассмотренных в главе 2. При этом на этапе вычисления выходной активности сети (которая необходима для выполнения итерации обучения) участвуют все нейроэлементы сети (их синаптические связи и пороги), и адаптации подлежат только добавленные в результате реконфигурации межнейронные связи. Следовательно, модификация параметров добавленного фрагмента будет минимизировать только ту часть погрешности, которая не смогла быть скомпенсирована моделью предшествовавшей архитектуры. Процесс обучения завершается при достижении минимальной скорости убывания погрешности обучения. Таким образом, нейроэлемент, обученный в соответствии с предложенным выше алгоритмом, после внедрения в сеть уменьшает общую ошибку обучения, что позволяет сохранить адаптированные параметры исходного фрагмента архитектуры в процессе дообучения реконфигурированной модели. Окончательной фазой добавления нового нейрона является дообучение всех нейроэлементов реконфигурированной архитектуры (до достижения параметра $V_{E,\min}$ заданного значения) с последующим тестированием модели и вычислением Q_{NN} .

Таким образом, в результате последовательного выполнения процедуры адаптивного добавления нейроэлементов в синтезируемую модель формируется гетерогенная архитектура $NN = (N^{[1]}, \hat{N}^{[2]}, 1)$, где оценка $\hat{N}^{[2]}$ определяется с использованием вышеописанных правил, исходя из минимизации Q_{NN} на каждой последующей стадии добавления

нейроэлементов. Процесс последовательного добавления элементов заканчивается при увеличении Q_{NN} относительно предыдущей архитектуры.

Необходимой завершающей стадией предлагаемого алгоритма является обучение сформированной модели до минимально возможной погрешности. Для этого используется вся совокупность обучающих эталонов, сформированных из элементов моделируемого временного ряда.

В четвертой главе выполняется анализ разработанных методов структурного проектирования и адаптивного обучения нейронных сетей в практических задачах моделирования и прогнозирования временных последовательностей и нелинейных систем. Рассматриваются задачи построения моделей прогнозирования сложных временных рядов. Выполняется сравнительный анализ параметров статистических моделей (наилучших вариантов моделей АРИСС, полученных для рассматриваемых временных последовательностей) и нейросетевых моделей, построенных на основе предлагаемых в работе методов обучения и синтеза. Показано, что в отдельных случаях прогнозирования трендовых последовательностей использование в архитектурах сетей логарифмической активационной функции скрытых нейронов позволяет улучшить как качество прогноза, так и параметры обучающего алгоритма.

Исследованы вопросы организации и применения нейросетевых моделей для прогнозирования хаотических процессов. Рассмотрены аспекты применения предложенных адаптивных алгоритмов синтеза и обучения нейронных сетей для организации нейросетевых моделей идентификации поведения нелинейных динамических систем на основе построения странных аттракторов. Поведение странного аттрактора в псевдофазовой плоскости может быть описано следующими рекуррентными выражениями:

$$X(t+1) = 1 - \alpha X(t)^2 + Y(t), \quad Y(t+1) = \beta X(t),$$

называющимися отображением Энона с фрактальной размерностью 1.26. Показано, что нейросетевая модель весьма точно восстанавливает траекторию оригинального аттрактора для большого количества итераций прогноза, демонстрируя тем самым высокую корректность обобщения свойств нелинейного динамического процесса при использовании выбранных архитектур и методов обучения нейронных сетей. Рассмотрена задача моделирования и итерационного прогнозирования последовательности, построенной на основе дифференциального уравнения с задержкой Маккей-Гласса, определяемого соотношением:

$$\frac{\partial X(t)}{\partial t} = -0.1X(t) + \frac{0.2X(t-\tau)}{1 + X(t-\tau)^{10}},$$

с использованием предложенных в работе методов обучения и синтеза нейронных сетей.

Предложена архитектура нейронной системы обработки погрешности дрейфа сенсорных устройств, ориентированная на выполнение функций организации индивидуальных моделей прогнозирования для последующего их использования при компенсации погрешности дрейфа. Система позволяет

выполнять прогнозы ряда погрешностей с шагом, задаваемым необходимым периодом измерения физических величин. Нейропостановка задачи прогнозирования погрешностей дрейфа с учетом специфики их измерений, заключается в следующем. Пусть $\delta(\tau)$ характеризует погрешность дрейфа сенсорного устройства, определяемую в момент τ его калибровки, $\tau \in [0, T]$, где T - период эксплуатации датчика. При этом вследствие нерегулярности моментов измерений погрешностей дрейфа интервалы времени между соседними измерениями носят произвольный характер. Совокупность значений:

$$\delta(\tau_1), \delta(\tau_2), \dots, \delta(\tau_R) \quad (11)$$

представляет собой ряд из R последовательных наблюдений погрешностей, выполненных в течение периода T . Задача нейронной системы состоит в определении уровня дрейфа $\delta(i)$ в фиксированные равноотстоящие моменты времени i , являющиеся точками измерения наблюдаемых величин, $i \in [0, T]$. Исходя из заданных условий, необходимо выполнить следующие этапы обработки:

1) На основе ряда погрешностей (11) построить аналитическую модель дрейфа:

$$\delta(\tau) = F(\tau) + \varepsilon(\tau). \quad (12)$$

2) На основе модели погрешности (12) определить значения погрешностей $\delta(i)$ в точках $i = 1, i + \Delta, i + 2\Delta, \dots$, входящих в диапазон аппроксимации $[0, T]$, Δ - шаг, соответствующий периоду измерения сигнала сенсора. В результате выполнения этого этапа получим преобразованный ряд погрешностей дрейфа:

$$\delta'(1), \delta'(\Delta), \delta'(2\Delta), \dots, \delta'(A\Delta), \quad (13)$$

который может быть использован для построения модели прогнозирования с шагом, соответствующим шагу измерений.

3) Используя ряд (13), построить аналитическую модель прогноза дрейфа.

4) Вычислить прогнозы дрейфа:

$$\hat{\delta}((A+1)\Delta), \hat{\delta}((A+2)\Delta), \dots$$

для соответствующих моментов измерения физической величины.

Описываются разработанные программные средства организации нейронных систем моделирования и прогнозирования временных последовательностей, позволяющие выполнять этапы инициализации, проектирования, обучения архитектур нейронных сетей прогнозирования и осуществлять прогнозы временных последовательностей.

Заключение содержит основные результаты исследований и выводы, сделанные на их основе.

В приложениях приведены исходные данные моделируемых временных рядов; параметры статистических и нейросетевых моделей прогнозирования; назначения команд программной системы и определяемые ими функции; документы о внедрении результатов разработок.

ЗАКЛЮЧЕНИЕ

Основные результаты диссертационной работы можно сформулировать следующим образом:

1. Разработана методика синтеза адаптивных алгоритмов обучения базовых архитектур нейронных сетей прогнозирования для различных функций выходной активности нейроэлементов. Данная методика обеспечивает более высокую сходимость, устойчивость и точность обучения по сравнению с классическим алгоритмом обратного распространения ошибки [1, 3-12, 14-17, 23-25, 27, 29-34].
2. Предложен адаптивный алгоритм последовательного послойного обучения гетерогенных архитектур многослойных нейронных сетей с прямыми и обратными связями, применяемых в прогнозирующих нейронных системах. Данный алгоритм обладает высокими параметрами обучения и исключает эмпирический подбор констант обучения и балансировки процесса обучения разнотипных слоев нейронной сети [10, 11, 15, 16, 19, 20, 22, 24, 35].
3. Разработана методика адаптивной инициализации параметров нейронной сети, которая обеспечивает снижение вероятности попадания целевой функции в локальные экстремумы на начальных стадиях процесса обучения, что повышает точность и стабильность обучения [15, 16, 27, 35].
4. На основании результатов исследований различных функций активации нейроэлементов сформулированы требования к архитектуре нейронных сетей прогнозирования. Предложен новый вариант функции выходной активности нейронных элементов, позволяющей как оптимизировать процедуры обучения, так и снизить ограничения традиционных нейросетевых моделей при прогнозировании различных классов временных последовательностей [11, 15, 24, 27, 32].
5. Предложен метод синтеза скрытого слоя нейронных сетей прогнозирования, позволяющий уменьшить структурную избыточность сети и исключить произвольный выбор архитектуры, обеспечивая при этом необходимую точность обучения и прогнозирования. Метод показал высокую сходимость процесса адаптации модели и высокую точность прогнозирования для разных типов временных рядов и хаотических процессов [26, 35].
6. Предложена нейронная система прогнозирования погрешностей дрейфа сенсорных устройств, учитывающая особенности характеристик сенсорной информации и калибровки сенсорных устройств. Система позволяет строить индивидуальные модели дрейфов сенсоров и выполнять прогнозы ряда погрешностей с шагом, задаваемым необходимым периодом измерения физических величин [13, 18, 28].

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в сборниках и журналах

1. Golovko V., Savitsky Y., Gladischuk V. Neural Nets Training Algorithms for Pattern Recognition and Prediction // Proc. Of the Third International Conference "Pattern Recognition and Information Analysis", Sept. 19-21, Minsk. - 1995. - P. 167-171.
2. Vladimir A. Golovko, Yury V. Savitski, Vitaly V. Gladischuk. Predicting Neural Net // Proc. of the Int. Conf. CMNDT-95, Minsk, Belarus. - 1995. - P. 348-353.
3. Golovko V., Savitsky Y., Gladischuk V. Neural Nets Training Algorithms for Pattern Recognition and Prediction // Proc. of Int. Conf. on Design Methodologies for Signal Processing, Szczecin, Technical University, Poland. - 1995. - P. 167-171.
4. Golovko V., Savitsky Y., Gladischuk V. Neural Networks for Autonomous Mobile Robot Control // Proc. of Int. Conf. NITE'96, Szczecin, Technical University. - 1996. - P. 234-238.
5. Golovko V., Savitsky Y., Gladischuk V. Neural System for Intelligent Robot Navigation // Proc. of Int. Conf. on Technical Informatics, Timisoara, Technical University, Romania. - 1996. - P. 63-70.
6. Golovko V., Savitsky Y., Gladischuk V. A Neural Net for Prediction Problems // Proceedings Of Int. Conf. on technical information, Timisoara, University of Timisoara, Romania. - 1996. - P. 49-52.
7. Golovko V., Syhodolsky O., Dimakov V., Savitsky Y., Gladischuk V. Neural Networks for Optimization Problems // Proc. of Int. Conf. on Design Methodologies for Signal Processing, Szczecin, Technical University, Poland. - 1996. - P. 38-41.
8. Савицкий Ю.В. Архитектура и алгоритм обучения нейронной сети прогнозирования // Материалы X научно-технической конференции "Новые технологии в машиностроении и вычислительной технике", ч. 2. - Брест, БПИ, 1998. - С. 41-48.
9. Савицкий Ю.В., Дунец А.П. Нейронные системы в задачах прогнозирования // Мат. междунар. конф. "Управление промышленным комплексом в условиях новой промышленной экономики УПК 98". - Брест. - 1998. - С. 170-174.
10. Savitsky Y. Neural Networks Training Algorithms and Architecture Selection for Prediction // Proc. of Second Int. Conf. CMNDT'98, 20-23 Oct., Minsk, Belarus. - 1998. - P. 363-367.
11. Theodore Laopoulos, Vladimir Golovko, Yury Savitsky, Anatoly Sachenko, Lucio Grandinetti. Architecture and Adaptive Training Algorithms of the Recurrent Neural Network. // Proc. of Int. Conf. "Information Technologies for Education, Science and Business" ITESB'99, Minsk, Belarus. - 1999. - P. 215-219.
12. Vladimir Golovko, Andrew Dunets, Yury Savitsky. The Training of Feed-Forward Neural Networks. // Proc. of the Int. Conf. on Neural Networks and Artificial Intelligence ICNNAI'99, Brest, Belarus. - 1999. - P. 36-39.

13. Vladimir Golovko, Yury Savitsky, Anatoly Sachenko, Volodimir Kochan, Volodimir Turchenko, Theodore Laopoulos, Lucio Grandinetti. Intelligent System for Prediction of Sensor Drift. // Proc. of the Int. Conf. on Neural Networks and Artificial Intelligence ICNNAI'99, Brest, Belarus. – 1999. – P. 126-135.
14. Vladimir Golovko, Yury Savitsky, Andrew Dunets, Anatoly Sachenko, Theodore Laopoulos, Lucio Grandinetti. Neural Network System for Time Series Prediction. // Proc. of the Sixth International Conference “Advanced Computer Systems” ACS'99, Szczecin, Poland. – 1999. – P. 223-227.
15. Vladimir Golovko, Yury Savitsky, Andrey Dunets, Anatoly Sachenko, Theodore Laopoulos, Lucio Grandinetti. Modifications of the Architecture and Training Algorithms of the Recurrent Neural Network for Time Series Prediction. // IEEE International Workshop on Intelligent Signal Processing WISP'99, Budapest, Hungary. – 1999. – P. 141-145.
16. Vladimir Golovko, Yury Savitsky. New Approach of the Recurrent Neural Network Training. // Proc. of the Int. Conf. on Neural Networks and Artificial Intelligence ICNNAI'99, Brest, Belarus. – 1999. – P. 32-35.
17. Yury Savitsky, Vladimir Golovko. Training of the Recurrent Neural Networks for Prediction. // Proc. of Fifth Int. Conf. “Pattern Recognition and Information Processing” PRIP'99, Minsk, Belarus. – 1999. – P. 248-253.
18. Sachenko A., Kochan V., Turchenko V., Golovko V., Savitsky Y, Dunets A. Laopoulos T. Sensor Errors Prediction Using Neural Networks // Proc. of Int. Joint Conf. on Neural Networks IJCNN'2000, Como, Italy. Vol. IV. – 2000. – P. 441-447.
19. Vladimir Golovko, Yury Savitsky, Theodore Laopoulos, Anatoly Sachenko, Lucio Grandinetti. Technique of Learning Rate Estimation for Efficient Training of MLP // Proc. of Int. Joint Conf. on Neural Networks IJCNN'2000, Como, Italy. Vol. 1. – 2000. – P. 323-329.
20. Vladimir Golovko, Yury Savitsky, Theodore Laopoulos, Anatoly Sachenko, Lucio Grandinetti. Technique for Efficient Training of MLP with Adaptive Training Step Rate Estimation. // Proc. of Second Int. ICSC Symposium on Engineering of Intelligent Systems EIS'2000, University of Paisley, Scotland, U.K. – 2000. – P. 145-148.
21. Vladimir Golovko, Oleg Ignatiuk, Yury Savitsky, Theodore Laopoulos, Anatoly Sachenko, Lucio Grandinetti. Unsupervised learning for dimensionality reduction. // Proc. of Second Int. ICSC Symposium on Engineering of Intelligent Systems EIS'2000, University of Paisley, Scotland, U.K. – 2000. – P. 140-144.
22. Vladimir Golovko, Yury Savitsky, Theodore Laopoulos, Anatoly Sachenko, Lucio Grandinetti. Unsupervised Neural Network Training for data Compression. // Proc. of Second Int. ICSC Symposium on Neural Computation NC'2000, Technical University of Berlin, Germany. – 2000. – P. 205-209.
23. Vladimir Golovko, Yury Savitsky, Theodore Laopoulos, Anatoly Sachenko, Lucio Grandinetti. Efficient Training of MLP with Training Step Rate Estimation. // Proc. of Euro-International Symposium on Computational Intelligence E-ISCI, Kosice, Slovakia. – 2000. – P. 21-26.

24. Головки В.А., Савицкий Ю.В. Адаптивные методы обучения градиентных нейронных сетей Вестник Брестского государственного технического университета. – 2000. – №4. – С. 68-75.
25. Головки В.А., Савицкий Ю.В. Метод обучения рекуррентных нейронных сетей // Вестник Брестского государственного технического университета. – 2000. – №4. – С. 53-56.
26. Головки В.А., Савицкий Ю.В., Фоменкова Н.Е. Применение нейронных сетей для прогнозирования и моделирования нелинейных систем // Вестник Брестского государственного технического университета. – 2000. – №4. – С. 56-59.
27. Савицкий Ю.В., Головки В.А. Метод адаптивной инициализации нейроэлементов в алгоритмах обучения градиентных нейронных сетей // Вестник Брестского государственного технического университета. – 2000. – №4. – С. 75-79.
28. Саченко А., Кочан В., Турченко В., Головки В., Савицкий Ю., Дунец А. Прогнозирование ошибок сенсорных устройств с использованием нейронных сетей // Вестник Брестского государственного технического университета. – 2000. – №4. – С. 59-63.

Тезисы докладов и материалы конференций

29. Головки В.А., Савицкий Ю.В. Отображение нейронных сетей на систолические архитектуры // Современные проблемы радиотехники, электроники и связи: Тез. докл. науч. техн. конф. – Минск, БГУИР, 1995. – С. 330.
30. Савицкий Ю.В., Головки В.А., Лазарчук А.П. Распознавание образов на нейронных сетях // Современные проблемы радиотехники, электроники и связи: Тез. докл. науч. техн. конф. – Минск, БГУИР, 1995. – С. 330-331.
31. Головки В.А., Савицкий Ю.В., Гладышук В.Б. Прогнозирующая нейронная сеть. // Современные проблемы радиотехники, электроники и связи: Тез. докл. науч. техн. конф. – Минск, БГУИР, 1995. – С. 331-332.
32. Савицкий Ю.В. Нейронные сети для решения задач прогнозирования. // Актуальные проблемы информатики: математическое, программное и информационное обеспечение: Тез. докл. научн. конф. – Минск, БГУ. – 1996. – С. 80.
33. Головки В.А., Савицкий Ю.В. Нейронные системы прогнозирования // Матер. науч. техн. конф., посвященной 30-летию института: Тезисы докладов конференции. – Брест, 1996. – С. 87.
34. Савицкий Ю.В., Ваткин М.Е., Поляков М.М. Нейросетевые технологии в задачах прогнозирования // Материалы X научно-технической конференции "Новые технологии в машиностроении и вычислительной технике", ч. 2. – Брест, БПИ, 1998. – С. 48-49.
35. Theodore Laopoulos, Vladimir Golovko, Yury Savitsky, Anatoly Sachenko, Lucio Grandinetti. The Recurrent Neural Networks Architectures and Adaptive Training Algorithms for Time Series Prediction. // Proc. of the 9-th World Conference on Titanium, St. Petersburg, Russia. – 1999. – P. 13.

РЕЗЮМЕ

Савіцкі Юрый Віктаравіч

Метады і сродкі генерацыі штучных нейронных сетак
для рашэння задач прадказання

Ключавыя словы: нейрасеткавыя маделі прадказання, шматслойныя нейронныя сеткі з рэкурэнтнымі сувязямі, гетэрагенныя нейронныя сеткі, функцыя актывацыі, алгарытм зваротнага распаўсюджвання памылкі, адаптыўны інтэрвал трэнавання, адаптыўная ініцыялізацыя міжнейронных сувязей, сінтэз нейроннай сеткі.

У дысертацыйнай рабоце даследваюцца метады будавання нейронных сістэм прадказання часавых паслядоунасцей. Мэта работы з'яўляецца распрацоўка метадаў і сродкаў сінтэза прагназавальных нейронных сетак і эфектыўных алгарытмаў іх трэнавання для павышэння дакладнасці прадказання.

Прапанавана метадыка і алгарытмы трэнавання базавых архітэктур нейронных сетак прадказання для розных функцый выходнай актывнасці нейраэлементаў. Гэтая метадыка забяспечвае больш высокую збегнасць, устойлівасць і дакладнасць трэнавання у параўнанні з класічным алгарытмам зваротнага распаўсюджвання памылкі. Распрацаваны алгарытм трэнавання гетэрагенных архітэктур шматслойных нейронных сетак з прамымі і зваротнымі сувязямі. Прапанавана метадыка адаптыўнай ініцыялізацыі параметраў нейроннай сеткі, якая забяспечвае памяншэнне пачатковай хібнасці маделі і зніжэнне верагоднасці пападання мэтай функцыі у лакальныя экстрэмы на пачатковых стадыях працэса трэнавання. На падставе вынікаў даследвання розных функцый актывацыі нейраэлементаў сфармуляваны патрабаванні да архітэктур нейронных сетак прадказання; прапанаваны новы варыянт функцыі актывацыі нейраэлементаў, які дазваляе аптымізіраваць працэдурны трэнавання, і таксама паменшыць абмежаванні традыцыйных нейрасеткавых маделей прадказання. Распрацаваны алгарытм сінтэза схаванага слою нейронных сетак з прамымі і рэкурэнтнымі сувязямі, які дазваляе паменшыць структурную збытковасць сеткі і выключыць давольны выбар архітэктур, забяспечваючы пры гэтым неабходную дакладнасць трэнавання і прадказання. Прапанавана архітэктур нейроннай сістэмы для прадказання часавых паслядоунасцей з патрэбным інтэрвалам.

Атрыманыя вынікі могуць ужывацца для будавання нелінейных маделей прадказання розных класаў складаных часавых працэсаў.

РЕЗЮМЕ

Савицкий Юрий Викторович

Методы и средства генерации искусственных нейронных сетей для решения задач прогнозирования

Ключевые слова: нейросетевые модели прогнозирования, многослойные нейронные сети с рекуррентными связями, гетерогенные нейронные сети, функция активации, алгоритм обратного распространения ошибки, адаптивный шаг обучения, адаптивная инициализация межнейронных связей, синтез нейронной сети.

В диссертационной работе исследуются методы построения нейронных систем прогнозирования временных последовательностей. Целью работы является разработка методов и средств синтеза прогнозирующих нейронных сетей и эффективных алгоритмов их обучения для повышения точности прогноза.

Предложена методика и алгоритмы обучения базовых архитектур нейронных сетей прогнозирования для различных функций выходной активности нейроэлементов. Данная методика обеспечивает более высокую сходимость, устойчивость и точность процесса обучения по сравнению с классическим алгоритмом обратного распространения ошибки. Разработан алгоритм обучения гетерогенных архитектур нейронных сетей с прямыми и обратными связями. Предложена методика адаптивной инициализации параметров нейронной сети, обеспечивающая уменьшение начальной погрешности модели и снижение вероятности попадания целевой функции в локальные экстремумы на начальных стадиях обучения. На основании результатов исследований различных функций активации нейроэлементов сформулированы требования к архитектуре нейронных сетей прогнозирования; предложен новый вариант функции активации нейроэлементов, позволяющей как оптимизировать процедуры обучения, так и снизить ограничения традиционных нейросетевых моделей прогнозирования. Разработан алгоритм синтеза скрытого слоя нейронных сетей с прямыми и рекуррентными связями, позволяющий уменьшить структурную избыточность сети и исключить произвольный выбор архитектуры, обеспечивая при этом необходимую точность обучения и прогнозирования. Предложена архитектура нейронной системы для прогнозирования временных рядов с требуемым шагом.

Полученные результаты могут быть применены для организации нелинейных моделей прогнозирования широких классов сложных временных процессов.

SUMMARY

Savitsky Yury Victorovich

Methods and Means of Artificial Neural Networks Development
for Solving of Prediction Tasks

Key words: neural model of prediction, recurrent multilayer neural networks, heterogeneous neural networks, activation function, backpropagation algorithm, adaptive training step, adaptive initializing of interconnection weights, synthesis of the neural network.

In this dissertation operation the methods of construction of neural systems for time series prediction are researched. The purpose of dissertation is the development of methods and means of synthesis of forecasting neural networks and effective training algorithms for rise of accuracy of the prediction.

The technique and algorithms of training of base architectures of predicting neural networks for different functions of output activity of neurons is offered. The given technique provides higher convergence, stability and accuracy of learning process in comparison with classic backpropagation algorithm. The algorithm of training of heterogeneous architectures of neural networks with forward and feedbacks is developed. The technique of adaptive initializing of parameters of the neural network providing decrease of an initial error of model is offered. It allows to lower a hit probability of the goal function in local extremums on initial states of training. On the basis of results of researches of different activation functions the requirements to the architecture of neural networks of prediction are formulated; The new variant of the activation function of neurons is offered permitting to optimize the process of training and to reduce the limitations of traditional neural models of prediction. The algorithm of synthesis of the hidden layer of neural networks with forward and recurrent links permitting to reduce structural network overhead and to eliminate arbitrary choice of the architecture is developed. It provides thus necessary accuracy of training and prediction. The architecture of the neural system for time series prediction with required step is offered.

The obtained results can be applied for organization of nonlinear models of prediction of various types of complex temporal processes.

Савицкий Юрий Викторович

МЕТОДЫ И СРЕДСТВА ГЕНЕРАЦИИ ИСКУССТВЕННЫХ
НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ПРОГНОЗИРОВАНИЯ

Специальность 05.13.15. – вычислительные машины и системы

Автореферат диссертации
на соискание ученой степени кандидата технических наук

Подписано в печать 13.11. 2000:

Формат 60×84 1/16.

Бумага офсетная. Печать ризографическая.

Усл. печ. л. 1,63.

Уч.-изд. л. 1,0. Тираж 90 экз.

Заказ 467.

Белорусский государственный университет информатики и
радиоэлектроники. Отпечатано в БГУИР. Лицензия ЛП N 156,
220013, Минск, ул. П. Бровки, 6.