

На правах рукописи

Трофимов Александр Геннадьевич

**СИНТЕЗ НЕЙРОСЕТЕВЫХ СТРУКТУР ДЛЯ
МОДЕЛИРОВАНИЯ УПРАВЛЯЕМЫХ
ОБЪЕКТОВ С РАСПРЕДЕЛЕННЫМИ
ПАРАМЕТРАМИ**

Специальность 05.13.01 – системный анализ, управление
и обработка информации (научное обслуживание)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Автор.



Москва – 2008

Работа выполнена в Московском инженерно-физическом институте
(государственном университете)

Научный руководитель кандидат технических наук, доцент
МИШУЛИНА Ольга Александровна

Официальные оппоненты доктор технических наук, профессор
ГАЛУШКИН Александр Иванович, Центр
информационных технологий и систем
органов исполнительной власти

кандидат технических наук, доцент
ТЮМЕНЦЕВ Юрий Владимирович,
Московский авиационный институт
(государственный технический университет)

Ведущая организация Московский государственный технический
университет им Н Э Баумана

Защита состоится «04» июня 2008 г в 16 час 00 мин на заседании
диссертационного совета Д 212 130 03 при Московском инженерно-
физическом институте (государственном университете) по адресу Москва.
Каширское ш , 31 (ауд 408 главного корпуса)

С диссертацией можно ознакомиться в библиотеке Московского
инженерно-физического института (государственного университета)

Отзывы в двух экземплярах, заверенные печатью, просьба
направлять по адресу Каширское ш , 31, диссертационные советы МИФИ
(тел 323-95-26)

Автореферат разослан «29» апреля 2008 г

Ученый секретарь
диссертационного совета,
д т н , профессор



Шумилов Ю Ю

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследований. Распределенные динамические объекты встречаются в различных областях науки и техники физике, экономике, медицине, геологии и пр В отличие от систем с сосредоточенными параметрами, состояние которых в каждый момент времени характеризуется вектором конечной размерности, для распределенных систем состояние описывается функцией одной или нескольких переменных, обычно интерпретируемых как пространственные координаты Проблема моделирования и управления такими объектами связана не только с их распределенностью и нелинейностью, но и с возможным отсутствием априорной информации о внутреннем устройстве моделируемой системы и закономерностях ее функционирования В этом случае практически единственным подходом к построению модели объекта остается использование экспериментальных данных о его функционировании

Активное развитие классической теории идентификации, наблюдавшееся в 70-80-ых годах прошлого века, преимущественно было направлено на моделирование систем с сосредоточенными параметрами, в то время как распределенным системам уделялось существенно меньшее внимание Основными причинами этого являлись сложность распределенной природы моделируемого объекта и трудности применения классических подходов

В последнее время наблюдается рост интереса к применению современных информационных технологий, в частности, нейросетевой технологии, к задачам моделирования распределенных динамических объектов Связано это, по всей видимости, с целым рядом факторов, среди которых можно отметить разнообразие практических приложений, трудности применения стандартных методов идентификации вследствие нелинейности моделей, большого объема данных, неточности их измерений, высокая вычислительная сложность классических методов моделирования распределенных систем

Успех применения нейронных сетей в математическом моделировании в основном связан с задачами обработки статических данных, в то время как нейросетевому моделированию динамических объектов уделяется существенно меньшее внимание Математические модели в классе нейронных сетей обладают рядом замечательных свойств, связанных с их

адаптивностью и способностью к генерализации данных, что говорит об их перспективности в рассматриваемой области

Большинство методов идентификации распределенных объектов предполагают известным вид уравнения с частными производными, описывающего моделируемый объект, а задача идентификации сводится к подбору параметров этого уравнения, при которых удовлетворяются заданные граничные и начальные условия или обеспечивается наилучшее соответствие результатам наблюдений. В зависимости от типа этого уравнения, как правило, применяются различные методы решения. Нейросетевой подход позволяет единым способом решать задачи идентификации объектов, описываемых дифференциальными уравнениями различных типов.

Среди отечественных научных исследований, посвященных проблеме нейросетевого решения уравнений математической физики, особо можно выделить работы Д.А. Тархова, А.Н. Васильева, В.И. Горбаченко. Многие работы в этой области являются узкоспециализированными и посвящены нейросетевому решению уравнений определенных типов либо предполагают нейросетевую реализацию классических методов их численного решения.

В связи с этим создание нейросетевых моделей и методов идентификации систем с распределенными параметрами в отсутствие информации о виде описывающих систему дифференциальных уравнений и конструирование соответствующих нейросетевых алгоритмов представляют актуальную и недостаточно изученную научную проблему.

К управляемым распределенным процессам относится процесс энерговыделения в активной зоне ядерного реактора. Знание трехмерного поля энерговыделения требуется для обеспечения безопасности эксплуатации реактора. Это связано с тем обстоятельством, что только на трехмерном поле энерговыделения можно определить критические теплофизические параметры безопасности.

В настоящее время существует ряд программных моделей поля энерговыделения, используемых на атомных электростанциях (АЭС) и для проведения научных исследований. Эти модели имеют приемлемую погрешность восстановления локального энерговыделения в стационарном или медленно меняющемся состоянии реактора, в то время как в переходных режимах точность расчетов значительно ухудшается. Поскольку физические модели основаны на численном решении уравнений математической физики, повышение требований к точности неизбежно приводит к снижению их быстродействия. Кроме того, для физического расчета необходима

адекватная информация о состоянии и параметрах активной зоны, которая не всегда известна с требуемой точностью

В связи с этим актуальна проблема построения адекватной быстродействующей модели поля энерговыделения в активной зоне реактора, возможное использование которой в качестве советчика оператора позволит повысить эффективность эксплуатации ядерной энергетической установки

Объектом исследования диссертации являются управляемые системы с распределенными параметрами, к которым относится активная зона ядерного реактора типа ВВЭР-1000

Предметом исследования являются методы идентификации управляемых систем с распределенными параметрами на основе экспериментальных данных в классе нейросетевых моделей

Цель и задачи диссертационной работы. Диссертация посвящена созданию нейросетевых моделей и методов идентификации управляемых систем с распределенными параметрами и методических рекомендаций по их применению. Достижение поставленной цели предполагает решение следующих основных задач

1 Формальное описание и теоретическое обоснование класса моделей, в рамках которого предполагается решать задачу идентификации управляемого распределенного объекта

2 Интерпретация моделей в нейросетевых терминах, позволяющая использовать для идентификации методы теории нейронных сетей, и синтез соответствующих нейросетевых моделей

3 Разработка методов и алгоритмов настройки параметров нейросетевых моделей в соответствии с выбранным критерием оптимальности

4 Формирование набора показателей для оценки точности обученных нейросетевых моделей управляемых распределенных объектов и предложений по их использованию

5 Разработка специализированной программной системы для проведения исследований нейросетевых моделей рассматриваемого типа

6 Применение предложенных нейросетевых моделей и алгоритмов для решения прикладной задачи идентификации распределенных процессов энерговыделения в активной зоне ядерного реактора при совершении маневра мощности

7 Экспериментальные исследования построенных моделей энерговыделения средствами разработанного программного обеспечения и оценка их точности и обобщающей способности

Методы исследования. Основой для создания нейросетевых моделей и алгоритмов и проведения экспериментальных исследований являются теория идентификации динамических систем, теория нейронных сетей, теория дифференциальных уравнений в частных производных и обыкновенных дифференциальных уравнений, численные методы решения дифференциальных уравнений, методы оптимизации и статистические методы анализа данных

Научная новизна. В диссертации получены следующие результаты, характеризующиеся научной новизной

1 Предложены и теоретически обоснованы два новых класса моделей управляемых объектов с распределенными параметрами, в рамках которых возможна декомпозиция задачи идентификации на отдельные подзадачи

2 Предложена нейросетевая интерпретация моделей указанных классов в виде нейросетевых структур – совокупности связанных нейросетевых модулей различных архитектур. Применение нейроструктурных моделей является новым подходом в теории идентификации распределенных систем

3 Разработаны два подхода к обучению нейросетевых структур рассматриваемых классов: последовательное и совместное обучение. Предложенный алгоритм совместного обучения обобщает классический метод обратного распространения ошибки во времени на случай нейросетевых структур и является новым в теории нейронных сетей

4 Нейроструктурный подход применен для решения задач моделирования распределенных процессов энерговыделения в активной зоне ядерного реактора. Построенные нейроструктурные модели являются новыми и не применялись ранее в математическом моделировании физических процессов в реакторной установке

5 Проведены серии компьютерных экспериментов, подтверждающие точность и обобщающую способность построенных нейроструктурных моделей процессов энерговыделения в активной зоне реактора

Обоснованность и достоверность результатов обеспечиваются корректностью применения математического аппарата и сопоставлением результатов компьютерных экспериментов созданных нейроструктурных моделей с данными сертифицированной программы NOSTRA, реализующей численный расчет соответствующих физических моделей

Практическая значимость работы.

1 Нейроструктурные модели энерговыделения, обученные на реальных данных, полученных с помощью датчиков системы внутриреакторного контроля (СВРК) АЭС, могут быть использованы в качестве советчика оператора. Использование такой быстродействующей модели совместно с аппаратурой СВРК позволит повысить безопасность эксплуатации ядерной энергетической установки в маневренном режиме.

2 Построенная динамическая модель управляемого аксиального профиля энерговыделения в активной зоне может быть использована для коррекции показаний датчиков аппаратуры контроля нейтронного потока, что позволит получить более точную информацию о состоянии активной зоны и, как следствие, повысить эффективность эксплуатации ядерной энергетической установки в маневренном режиме.

3 Разработанное в диссертации методическое, алгоритмическое и программное обеспечение может быть использовано в учебных курсах по современным методам математического моделирования и нейроинформатике, что позволит повысить уровень подготовки специалистов.

4 Предложенные модели и методы идентификации динамических систем в классе нейросетевых структур могут быть применены для решения прикладных задач в других проблемных областях. В работе описаны две подобные задачи: моделирование процессов гибели опухолевых клеток в живых тканях в условиях гипертермического воздействия и траекторное управление самолетом, который совершает оборонительный маневр от атакующей его ракеты.

Реализация результатов работы. Математические модели, методы, алгоритмы и научно-практические рекомендации диссертации применены при решении задач моделирования физических процессов в активной зоне реактора ВВЭР-1000 и могут быть использованы при решении задач управления.

Полученные в диссертационной работе результаты используются в учебном процессе кафедр "Кибернетика" и "Теплофизика" МИФИ (Москва).

Созданное методическое и алгоритмическое обеспечение, а также практические рекомендации по обучению нейросетевых структур были применены при разработке системы распознавания символьной информации (ООО "АСТ", Москва).

Использование результатов диссертационной работы и достигнутый при этом эффект подтверждены соответствующими актами

Апробация работы. Основные положения диссертации докладывались на следующих всероссийских и международных семинарах и конференциях

- Международная научно-техническая конференция "Интеллектуальные и многопроцессорные системы" (Геленджик, 2003)
- Всероссийская научно-техническая конференция "Нейроинформатика" (Москва, МИФИ, 2005, 2006, 2008)
- Всероссийская конференция "Нейросетевые алгоритмы решения задач математической физики" (Москва, НИИАА, 2007).
- Междисциплинарный научный семинар "Экобионика" (Москва, МГТУ им Н Э Баумана, 2008)
- Annual Meeting of the European Society for Hyperthermic Oncology (Berlin 2006, Prague 2007)
- World Congress on Medical Physics and Biomedical Engineering (Seoul 2006)
- Научная сессия Российского научно-технического общества радиотехники, электроники и связи им А С Попова (Москва, 2007)
- Научная сессия МИФИ (Москва, МИФИ, 2005, 2006, 2007, 2008)
- Международный научно-технический семинар "Современные технологии в задачах управления, автоматизации и обработки информации" (Алушта, МАИ, 2005, 2006, 2007, 2008)

Результаты диссертации использованы в отчетах по проектам РФФИ (№ 05-08-01421) и МНТЦ (#2221)

Публикации результатов. Основные результаты диссертации опубликованы в 26 печатных научных работах, среди которых 2 статьи – в периодических научных изданиях, рекомендованных ВАК России, 8 работ – в статьях и материалах конференций

На защиту выносятся:

1 Новый тип нейросетевых математических моделей управляемых объектов с распределенными параметрами – нейросетевые структуры – в рамках которых предлагается решать задачу идентификации

2 Два подхода к обучению нейроструктурных моделей последовательное, предполагающее обучение отдельных модулей структуры в определенной последовательности, и совместное, основанное на алгоритме обратного распространения ошибки во времени Оба подхода допускают обобщение на более широкий класс нейросетевых структур

3 Построенные в классе нейросетевых структур модели управляемых распределенных процессов энерговыделения в активной зоне ядерного реактора типа ВВЭР-1000 модель энерговыделения в фиксированном узле активной зоны, модель аксиального профиля энерговыделения тепловыделяющей сборки, модель поля энерговыделения активной зоны

4 Результаты компьютерных экспериментов построенных нейроструктурных моделей энерговыделения, демонстрирующих точность и обобщающие способности моделей в ограниченной области типовых управляющих воздействий, а также практические рекомендации по их обучению

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения, библиографического списка и четырех приложений Основная часть работы содержит 148 страниц машинописного текста, включая 14 таблиц, 58 рисунков и перечень использованной литературы из 128 наименований

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы исследования, формулируются цель и решаемые в работе задачи, отмечаются научная новизна, практическая значимость работы, приводятся основные положения, выносимые на защиту, кратко излагается содержание диссертации

В первой главе дается формулировка задачи идентификации управляемого динамического объекта с распределенными параметрами Определяется исследуемый класс математических моделей распределенных систем

$$\begin{cases} \frac{ds(x,t)}{dt} = \varphi(x, s(x,t), s'_x(x,t), s_x^{(n)}(x,t), u(t), u'(t), u^{(m)}(t), \gamma(x)), \\ y(x,t) = \psi(x, s(x,t), \gamma(x)), \quad x \in \Omega, \quad t \in \mathcal{T}, \end{cases} \quad (1)$$

где t, x – время функционирования и пространственные координаты соответственно, $u(t)$ – вектор управляющих воздействий в момент времени t , $s(x, t)$, $y(x, t)$, $\gamma(x)$ – векторы состояния, выходов и параметров системы соответственно, φ – оператор эволюции системы, ψ – функциональное преобразование устройства наблюдения

В главе рассматриваются существующие подходы к идентификации систем с распределенными параметрами (СРП) и известные классические

методы идентификации В результате анализа литературы установлено, что все методы идентификации на основе данных наблюдений над объектом типа "вход-выход" предполагают аппроксимацию бесконечномерной распределенной системы некоторой конечномерной сосредоточенной системой (*редукцию СРП*) В связи с этим рассмотрение методов идентификации распределенных систем невозможно без рассмотрения методов идентификации систем с сосредоточенными параметрами В диссертации обсуждаются проблемы, связанные с использованием классических способов редукции и методов численного моделирования

Применение нейросетевого подхода для идентификации распределенных систем позволяет преодолеть ряд проблем классических методов, среди которых можно выделить их высокую вычислительную сложность

При использовании нейронных сетей для идентификации исследователь сталкивается с двумя основными проблемами определение архитектуры нейросетевой модели и выбор (или разработка) процедуры ее обучения В главе приводится обзор и классификация архитектур нейронных сетей, применяемых для решения задач идентификации динамических объектов В результате анализа литературы сделан вывод, что все нейросетевые архитектуры, применяемые для решения задач идентификации, могут быть разделены на три группы

1) В первую группу входят статические нейронные сети, как правило, включающие линии задержек управляющего и выходного сигналов объекта, с помощью которых в сеть привносится память

2) Вторую группу образуют нейросети с обратными связями – рекуррентные, или динамические, нейросети Такие сети вызывают гораздо больший интерес исследователей при решении задач идентификации динамических систем Выбор динамической сети в качестве модели динамического объекта представляется более обоснованным

3) В третью группу входят специальные архитектуры нейронных сетей, построенные на нейронах высших порядков или нейронах с неклассическими моделями

Проводится сравнительный анализ, указываются преимущества и недостатки рассматриваемых нейросетевых архитектур Учитывая опыт авторов публикаций, для многих из рассмотренных нейросетевых моделей приводятся сложности, с которыми сталкивается исследователь при их построении и использовании, а также рекомендации по их применению

Анализируется состояние проблем нейросетевой идентификации систем с распределенными параметрами. Отмечается, что эта область малоизученна, в ней имеется относительно небольшое число монографий и публикаций, что, по всей видимости, связано со сложностью решаемых задач.

Во второй главе представлены методические и алгоритмические основы идентификации управляемых систем с распределенными параметрами в классе нейросетевых моделей. Предложены два подхода к редукции модели распределенной системы (1). Первый подход – *редукция в пространственной области* – предполагает аппроксимацию бесконечномерной модели СРП (1) совокупностью систем с сосредоточенными параметрами в конечном числе пространственных точек с последующей дискретизацией во времени. В результате получена система разностных уравнений, в которых пространственная координата x рассматривается как параметр. Основываясь на теории вложения, для этой системы записана нелинейная авторегрессионная модель (*Nonlinear AutoRegressive model with exogenous inputs, NARX-модель*) с вектором параметров w , зависящим от x

$$\begin{cases} y(x,t) = F(w, y(x,t-1), \dots, y(x,t-n), u(t), u(t-1), \dots, u(t-m)), \\ w = G(x, \gamma), \quad x \in \Omega, \quad \gamma \in \Gamma, \end{cases} \quad (2)$$

где t – дискретное время, $t = \overline{0, T}$, m, n – число задержек в линии управления и цепи обратной связи соответственно, F – функция авторегрессии.

Задача идентификации распределенной системы, таким образом, сводится к задаче аппроксимации функций F и G в модели (2). Учитывая универсальные аппроксимирующие способности нейронных сетей, для аппроксимации функции авторегрессии F предложено использовать *NARX-нейросеть* (рис 1), для аппроксимации функции G – многослойный персептрон *MLP*. Схема полученной нейросетевой модели приведена на рис 2.

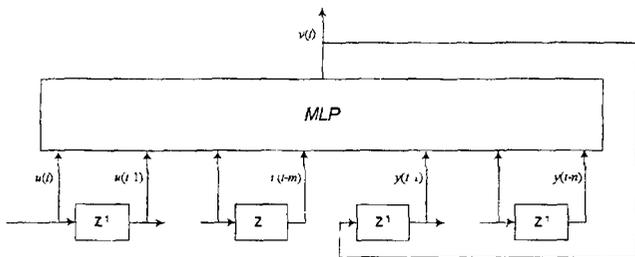


Рис 1 Нейросетевая NARX-модель

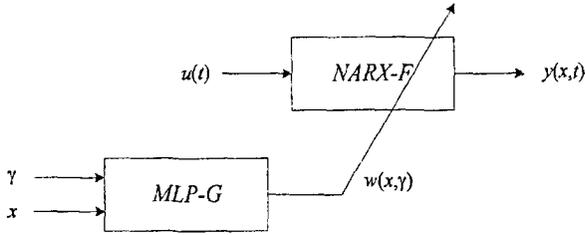


Рис 2 Схема модели управляемой распределенной системы, основанной на редукции в пространственной области

Введено понятие нейросетевой структуры (НСС) как вычислительной модели, включающей совокупность взаимосвязанных нейросетевых модулей, в общем случае, различных нейросетевых архитектур. Пример нейросетевой структуры изображен на рис 2.

Второй подход – *редукция во временной области* – основан на представлении выхода распределенной системы в каждый момент времени t , $t \in \overline{0, T}$, в виде некоторой параметрической функции

$$y(x, t) = G^*(x, p(t, \gamma)), \quad x \in \Omega, \quad \gamma \in \Gamma,$$

где $p(t, \gamma)$ – вектор параметров функции G^* в момент времени t . В этом случае динамика пространственного распределения $y(x, t)$ связана с динамикой управляемого вектора $p(t, \gamma)$. Предполагая, что вектор $p(t, \gamma)$ удовлетворяет разностному уравнению, для которого может быть записана NARX-модель, предложен еще один класс моделей для идентификации

$$\begin{cases} y(x, t) = G^*(x, p(t, \gamma)), \quad x \in \Omega, \quad \gamma \in \Gamma, \\ p(t) = F^*(w^*, p(t-1), \dots, p(t-n), u(t), u(t-1), \dots, u(t-m)), \\ w^* = H^*(\gamma), \quad \gamma \in \Gamma, \end{cases} \quad (3)$$

где m, n – число задержек в линии управления и цепи обратной связи соответственно, F^* – функция авторегрессии, w^* – вектор параметров.

Модель (3) также может быть реализована в виде нейросетевой структуры (рис 3). В качестве нейросетевой модели функции F^* выбрана NARX-нейросеть, в качестве аппроксиматоров функций G^* и H^* – многослойные перцептроны.

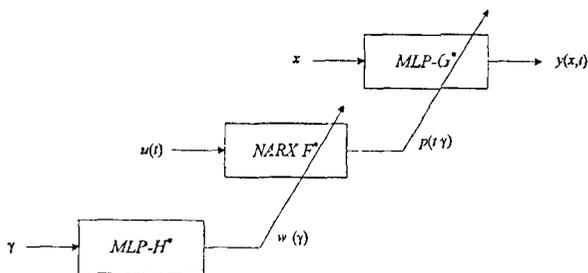


Рис 3 Схема модели управляемой распределенной системы, основанной на редукции во временной области

В диссертации рассмотрены математическая модель *NARX*-нейросети и алгоритм ее обучения, основанный на модификации метода обратного распространения ошибки во времени (*Back Propagation Through Time, BPTT*)

Для обучения предлагаемых нейросетевых структур используется совокупность результатов наблюдений $\Sigma = \{\pi^{(p)}, p = \overline{1, P}\}$ распределенного объекта в конечном числе пространственных точек x^r , $x^r \in \Omega$, $r = \overline{1, R}$, проведенных в дискретные моменты времени t , $t = \overline{0, T}$ Результат p -го наблюдения $\pi^{(p)}$ определен как совокупность

$$\pi^{(p)} = \left\langle c^{(p)}, \left(u^{(p)}(t), t = \overline{0, T} \right), \left(\sigma^{(p)}(t), t = \overline{0, T} \right), \gamma^{(p)} \right\rangle,$$

где $c^{(p)}$ – идентификатор начального состояния $s_0^{(p)}(x)$, в котором находилась СРП при проведении p -го эксперимента, T – конечное число временных тактов, $u^{(p)}(t)$ – вектор управляющих воздействий на такте времени t , $\sigma^{(p)}(t) = \left(\sigma^{(p)}(x_1, t), \dots, \sigma^{(p)}(x_R, t) \right)^T = \left(\sigma^{1(p)}(t), \dots, \sigma^{R(p)}(t) \right)^T$ – вектор наблюдаемых выходов системы на такте времени t в пространственных точках x^1, \dots, x^R , $x^r \in \Omega$, $r = \overline{1, R}$, $\gamma^{(p)}$ – вектор параметров СРП, при которых проводился эксперимент

В диссертации предложены два подхода к обучению нейросетевых структур Первый подход основан на последовательном обучении нейросетевых модулей структуры, используя многошаговую процедуру Второй – на вычислении частных производных критерия обучения по всем настраиваемым параметрам НСС Эти подходы названы *последовательным* и *совместным (параллельным)* обучением соответственно

Суть последовательной процедуры обучения НСС, полученной в результате редукции в пространственной области (рис 2), состоит в следующем. На основе результатов наблюдений $\pi^{(p)}$ для каждой фиксированной пространственной точки x^r , $r = \overline{1, R}$, строится отдельная нейросетевая *NARX*-модель, характеризующаяся вектором параметров $w(x^r)$. В результате многократного обучения будет получена конечная совокупность *NARX*-нейросетей, соответствующих отдельным пространственным точкам. Далее проводится аппроксимация векторов $w(x^r)$ в пространственной области Ω с помощью многослойного персептрона *MLP-G*. В диссертации приводятся особенности, преимущества и недостатки последовательной схемы обучения, одним из которых является необходимость формирования и хранения вспомогательных выборок данных, используемых для обучения многослойного персептрона *MLP-G*.

Совместное обучение модулей НСС предполагает оптимизацию критерия обучения в пространстве всех настраиваемых параметров структуры

$$I = \frac{1}{2} \sum_{p=1}^P \sum_{r=1}^R \sum_{t=0}^T \alpha(t) \left(y^{(p)}(x^r, t) - \sigma^{(p)}(x^r, t) \right)^2 \rightarrow \min_{\nu} \quad (4)$$

В связи с тем, что вектор выходов w сети *MLP-G* представляет собой вектор параметров *NARX*-нейросети, то рассматриваемая НСС характеризуется только вектором ν синаптических коэффициентов сети *MLP-G*, в пространстве которых решается задача оптимизации критерия (4).

В диссертации предлагается метод расчета градиента критерия (4) при совместном обучении модулей НСС, обобщающий метод обратного распространения ошибки во времени. В отличие от классического метода *BPIT*, обратное распространение в структуре происходит не через нейроны сети, а через отдельные нейросетевые модули, рассматриваемые как элементарные объекты сети, в общем случае имеющие различные математические модели. Предлагаемый метод обучения может быть применен для более широкого класса нейросетевых структур.

Для сокращения размерности пространства параметров обучаемых *NARX*-нейросетей и повышения эффективности обучения в диссертации предлагается модификация нейросетевой *NARX*-модели, основанная на параметризации синаптических коэффициентов нейронов первого слоя многослойного ядра *NARX*-сети, и рассматривается алгоритм ее обучения. Введение такой параметризации для ряда динамических систем соответствует физическим представлениям о закономерностях их функционирования.

Для оценки точности нейросетевых моделей объектов с распределенными параметрами приводится набор показателей, которые используются при обучении, валидации и тестировании сети, а также в критериях останова итерационной обучающей процедуры. Формулируются практические рекомендации по организации процесса обучения НСС.

Третья глава посвящена применению предложенных во второй главе моделей и методов для решения задачи идентификации управляемого процесса энерговыделения (ЭВ) в активной зоне ядерного реактора типа ВВЭР-1000 при совершении маневра мощности. В этой задаче выделены три подзадачи:

- идентификация процесса энерговыделения в фиксированном узле активной зоны,
- идентификация аксиального профиля энерговыделения тепловыделяющей сборки (ТВС),
- идентификация пространственного поля энерговыделения в активной зоне.

Приводится краткое описание проблемной области и исходных данных. В качестве модели энерговыделения в фиксированном узле активной зоны предложено использовать нейросетевую *NARX*-модель. Управляющим воздействием $u(t)$ нейросети является положение органов регулирования системы управления и защиты (ОР СУЗ). Желаемым выходным процессом – соответствующий этому управлению процесс ЭВ в фиксированной точке активной зоны. Вектор параметров γ активной зоны предполагается фиксированным для всех рассматриваемых режимов управления.

Для построения нейросетевой модели распределенного вдоль ТВС процесса энерговыделения проведена редукция модели СРП во временной области. Для аппроксимации профиля ЭВ в каждый момент времени t , $t = \overline{0, T}$, маневра мощности реактора выбран полином степени k

$$y(z, t) = p_k(t)z^k + p_{k-1}(t)z^{k-1} + \dots + p_1(t)z + p_0(t), \quad (5)$$

где z – масштабированная координата точки на ТВС.

Задача моделирования управляемого профиля ЭВ, таким образом, сведена к задаче моделирования коэффициентов аппроксимирующего полинома (5). Сформулировано предположение, что в динамике элементов вектора $p(t)$ присутствует небольшое число главных компонентов, $\tilde{p}_i(t)$, $i = \overline{0, \tilde{k}}$, $\tilde{k} < k$, образующих вектор $\tilde{p}(t)$, инвариантных относительно рассматриваемых режимов управлений. Рассмотрена схема перехода от

исходного пространства коэффициентов к пространству главных компонентов и обратно. Для моделирования процессов $\bar{p}_i(t)$, $i = \overline{0, k}$, использованы *NARX*-нейросети. На рис 4 изображена полученная нейроструктурная модель аксиального профиля ЭВ ТВС.

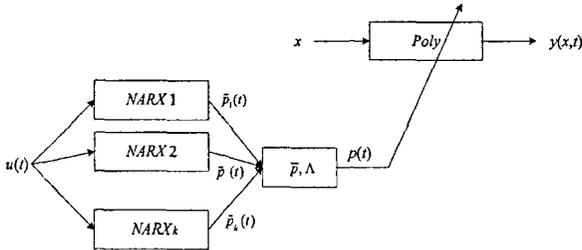


Рис 4 Схема нейросетевой модели аксиального профиля ЭВ ТВС

При построении модели управляемого пространственного поля энерговыделения в активной зоне реактора ВВЭР-1000 также проведена редукция модели СРП во временной области. Реализация модели в виде нейросетевой структуры приведена на рис 5. Для аппроксимации поля ЭВ в каждый момент времени t , $t = \overline{0, T}$, маневра мощности реактора использован многослойный перцептрон *MLP-G*, характеризующийся вектором синаптических коэффициентов $p(t)$. Как и в предыдущей задаче, среди элементов этого вектора выделены главные независимые компоненты, для моделирования которых использованы *NARX*-нейросети. Многослойный

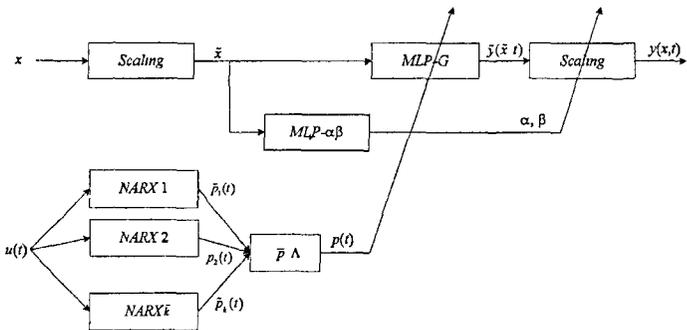


Рис 5 Схема нейросетевой модели пространственного поля энерговыделения в активной зоне реактора

персептрон $MLP-\alpha\beta$, включённый в структуру, решает вспомогательную задачу масштабирования энерговыделений. В диссертации описана схема функционирования НСС и рассмотрен метод её обучения.

Для выполнения экспериментальных исследований предложенных нейроструктурных моделей разработана специализированная программная система. В диссертации приводится краткое описание архитектуры и функциональных возможностей системы.

В четвёртой главе приводятся результаты экспериментальных исследований нейросетевых моделей энерговыделения в активной зоне ядерного реактора типа ВВЭР-1000. Описываются эксперименты, выполненные с использованием пакета расширения MATLAB для идентификации линейных динамических систем (*System Identification Toolbox*), подтверждающие существенную нелинейность моделируемого объекта.

Приводятся результаты экспериментальных исследований нейросетевых $NARX$ -моделей процесса ЭВ в фиксированном узле активной зоны. Исходные данные для построения моделей изображены на рис. 6.

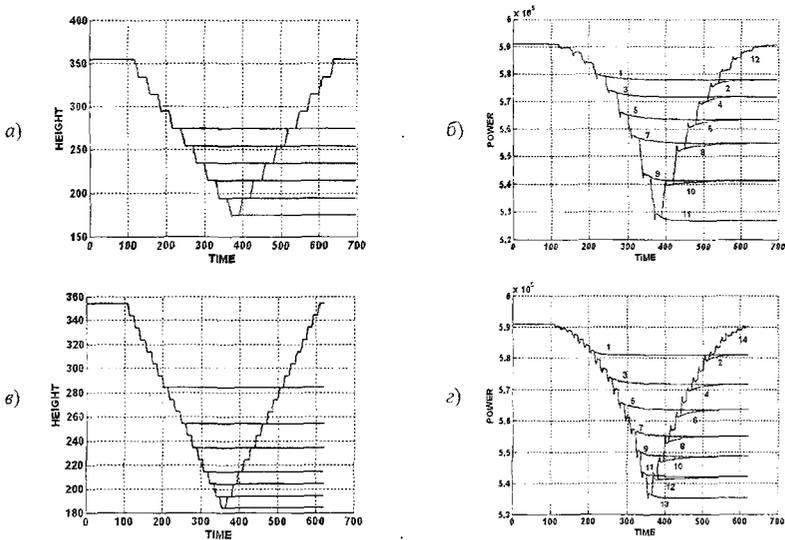


Рис. 6. Процессы энерговыделения, наблюдаемые в узле с пространственным положением $x' = (1, 1)$ (б, г), и соответствующие им процессы движения ОР СУЗ (а, в). Графики а и в соответствуют различным режимам управления. Положение ОР СУЗ указано в см, считая от низа активной зоны, единица измерения ЭВ – $MВт$, время указано в с

Для обучения использовался один обучающий пример (процесс 10 на рис. 6б и соответствующий ему управляющий процесс), остальные процессы использованы для тестирования. Каждый управляющий процесс представляет собой кусочно-линейную функцию с выходом на постоянный уровень.

Для оценки точности нейросетевых моделей введён ряд показателей, среди которых MSE – среднеквадратическая ошибка, $MAPE$, $MAXPE$ – средняя и максимальная относительные ошибки, $MAXE$ – максимальная абсолютная ошибка, $MEPP$, $MAXPP$ – средняя и максимальная ошибки, отнесённые к размаху сигнала.

В результате исследований различных архитектур $NARX$ -моделей установлено, что сеть с архитектурой $m = 5$, $n = 5$, $K = 2$, $N_1 = 5$, $N_2 = 1$ (K – число слоёв, N_k – число нейронов в k -ом слое) обладает требуемыми точностными характеристиками (относительные показатели менее 5%). Показатели точности нейросетевой модели на обучающей и тестовых реализациях приведены в таблице.

	MSE	$MAPE$	$MAXE$	$MEPP$	$MAXPP$
обуч.	5.71e+5	0.094%	2.62e+3	0.83%	4.13%
тест.	1.38e+6	0.15%	3.32e+3	1.36%	5.23%

Результаты воспроизведения сетью обучающих и тестовых процессов ЭВ приведены на рис. 7. Из рисунка видно, что желаемые и фактические выходы сети практически совпадают.

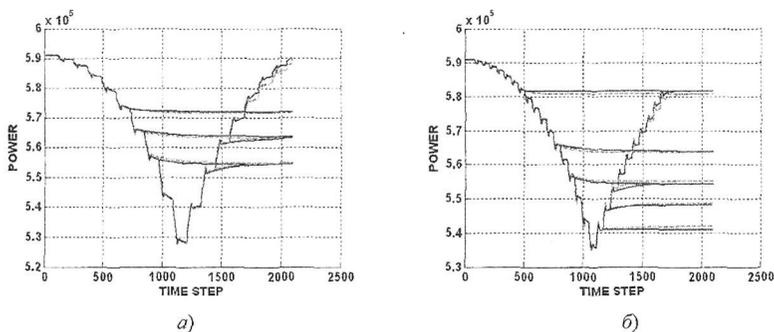


Рис. 7. Результаты воспроизведения обученной $NARX$ -нейросетью обучающей и тестовых реализаций. Пунктирной линией показаны исходные реализации, сплошной – выходы модели. Архитектура сети: $m = 5$, $n = 5$, $K = 2$, $N_1 = 5$, $N_2 = 1$. Активационные характеристики нейронов первого слоя – сигмиды, выходного нейрона – линейная.

Сделан ряд выводов о влиянии числа нейронов и числа задержек m и n в цепях управления и обратной связи $NARX$ -нейросети на её динамические свойства и способности к обобщению в пространстве управляющих воздействий. Проведено исследование зависимости точности и обобщающей способности сети от начальных значений синаптических коэффициентов при обучении градиентными методами. В диссертации обсуждаются особенности применения построенных нейросетевых моделей энерговыделения.

Исходные данные для построения нейросетевой модели аксиального профиля энерговыделения ТВС и результаты его полиномиальной аппроксимации показаны на рис. 8. В качестве модели профиля использована нейросетевая структура, изображённая на рис. 4.

В результате исследований $NARX$ -нейросетей, входящих в структуру, средствами разработанного программного обеспечения показано, что сети с архитектурными параметрами $m = 3, n = 3, K = 2, N_1 = 5, N_2 = 1$ воспроизводят обучающие и тестовые реализации факторных очков $\tilde{p}_i(t), i = \overline{0, \tilde{k}}$, с требуемой точностью (показатели $MEPP < 5\%$).

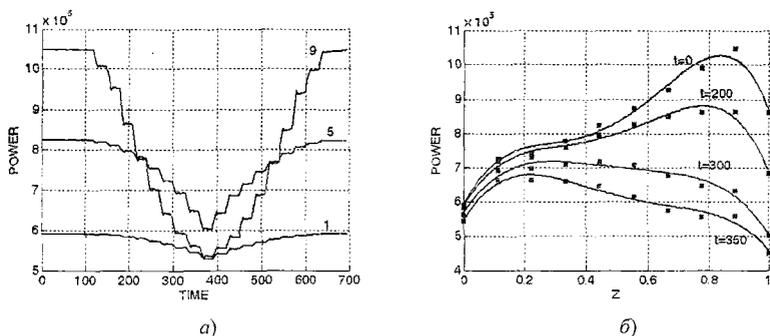


Рис. 8. Процессы энерговыделения, наблюдаемые в различных узлах ТВС и используемые при построении нейросетевой модели (а) и аксиальные профили энерговыделения ТВС в различные моменты времени манёвра мощности реактора (б). Порядок аппроксимирующего полинома $k = 4$. Соответствующий режим управления приведён на рис. 6а

В диссертации исследована зависимость точности модели от порядка k аппроксимирующего полинома и числа \tilde{k} выделяемых главных компонент. Показано, что главное подпространство практически

инвариантно к рассматриваемым режимам ступенчатого управления. На рис. 9 приведены результаты работы построенной модели.

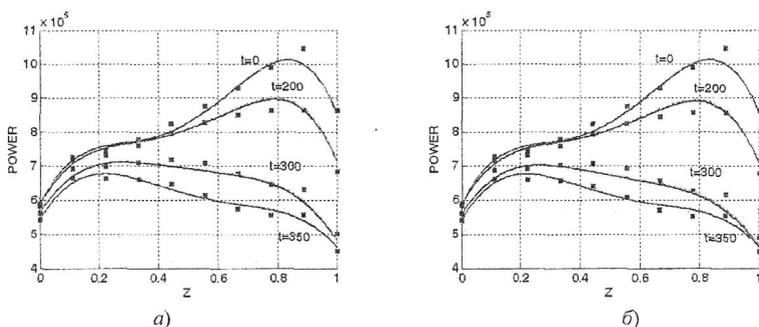


Рис. 9. Аксиальные профили энерговыделения ТВС в различные моменты времени манёвра мощности. Время указано в секундах, считая от начала манёвра. Сплошной линией показаны выходы нейросетевой модели, пунктирной – профили, аппроксимированные полиномом порядка $k=4$. Рисунки *a* и *б* соответствуют различным режимам управления (см. рис. 6а, в)

Для построенной нейросетевой модели на протяжении всего времени манёвра мощности реактора значение показателей точности $MAXPP < 7\%$ ($MAXE < 3\%$), что не уступает аналогичным показателям физической модели NOSTRA. Указанные значения говорят о способности модели к обобщению в ограниченном пространстве рассматриваемых режимов ступенчатого управления.

Для моделирования пространственного поля энерговыделения в активной зоне реактора использована нейросетевая структура, изображённая на рис. 5. В результате исследований средствами разработанной программной системы многослойных персептронов $MLP-\alpha\beta$ и $MLP-G$, входящих в НСС, найдены их архитектурные параметры, при которых обеспечиваются требуемые точностные характеристики на обучающей и тестовой выборках. Отмечена сложность аппроксимации поля ЭВ с помощью многослойной нейросети, связанная с его сильной неравномерностью и многоэкстремальностью.

В результате исследований различных архитектур $NARX$ -моделей построена сеть с архитектурными параметрами $m = 10$, $n = 5$, $K = 2$, $N_1 = 10$, $N_2 = 1$, которая воспроизводит реализации факторных очков $\bar{p}_i(t)$, $i = 0, \bar{k}$, с

точностью $MEPP < 12\%$ на обучающих реализациях и $MEPP < 20\%$ на тестовых. На рис. 10 показаны процессы энерговыделения в отдельных узлах активной зоны, полученные по результатам работы построенной НСС. В диссертации проводится анализ зависимости точности моделирования от архитектурных параметров отдельных модулей НСС.

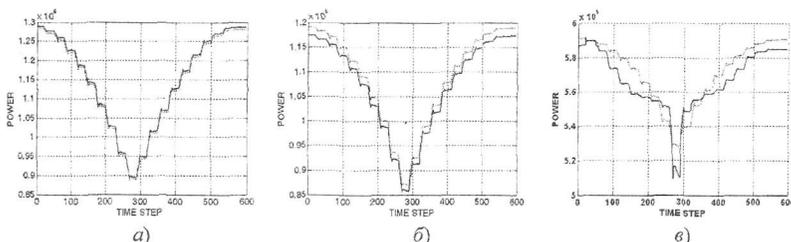


Рис. 10. Процессы энерговыделения в отдельных узлах активной зоны, построенные по результатам работы нейросетевой модели пространственного поля ЭВ (сплошная линия) и соответствующие исходные процессы ЭВ (пунктирная линия). На рис. *а* – обучающий процесс, на рис. *б*, *в* – тестовые.

Показатели точности построенной нейросетевой модели на обучающей и тестовых реализациях приведены в таблице.

	<i>MSE</i>	<i>MAPE</i>	<i>MAXE</i>	<i>MEPP</i>	<i>MAXPP</i>
<i>обуч.</i>	6.47e+7	1.22%	1.78e+4	10.91%	28.06%
<i>обуч.</i>	6.78e+9	6.89%	1.79e+5	8.12%	23.28%
<i>обуч.</i>	8.81e+6	0.52%	1.28e+4	2.68%	16.85%
<i>тест.</i>	3.52e+9	4.43%	1.35e+5	8.92%	25.88%
<i>тест.</i>	8.88e+9	10.52%	1.71e+5	14.18%	28.73%

В диссертации приводится обсуждение причин возникновения погрешностей моделирования и предлагаются пути дальнейшего повышения точности. По показателям точности построенная нейросетевая модель сравнима с математической моделью, реализованной в программе NOSTRA, при этом существенно превышая её в быстродействии. Однако в отличие от математических моделей, предполагающих численное решение уравнений математической физики, область применения построенных нейросетевых моделей ограничивается рассматриваемыми режимами кусочно-линейного управления.

В приложения вынесены вывод уравнений обучения нейросети нелинейной авторегрессии, описание применения предлагаемого нейроструктурного подхода для решения актуальной задачи в радиобиологии человека и животных, описание функциональных возможностей разработанной программной системы для экспериментальных исследований нейросетевых структур и акты об использовании результатов диссертационного исследования

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В заключении приведены основные результаты диссертации и сформулированы выводы

Итогом диссертационной работы являются следующие научные и практические результаты

1 Поставлена задача идентификации управляемого объекта с распределенными параметрами на основе результатов наблюдений типа "вход-выход" Введены и теоретически обоснованы два класса моделей для идентификации, предполагающие редукцию математической модели распределенной системы

2 Предложена нейросетевая реализация математических моделей указанных классов в виде нейросетевых структур – моделей нового типа, состоящих из взаимосвязанных нейросетевых модулей в общем случае различных архитектур

3 Предложены два подхода к обучению нейросетевых структур последовательное, включающее многошаговую процедуру обучения отдельных модулей, и совместное, предполагающее одновременную адаптацию всех параметров структуры Разработан алгоритм обучения нейросетевой структуры как целого, обобщающий метод обратного распространения ошибки во времени

4 Для моделирования сосредоточенной управляемой динамики использована управляемая нейросеть нелинейной авторегрессии (*NARX*-нейросеть) Предложена модификация математической модели *NARX*-нейросети, позволяющая сократить число настраиваемых параметров и повысить эффективность процедуры обучения

5 Предложен набор показателей для оценки точности нейроструктурных моделей управляемых распределенных объектов Сформулированы практические рекомендации по организации процесса обучения нейросетевых структур и их применению

6 Предложенные нейроструктурные модели и алгоритмы применены для решения прикладной задачи идентификации распределенных процессов энерговыделения в активной зоне ядерного реактора типа ВВЭР-1000 при совершении маневра мощности Построены модель управляемых процессов энерговыделения в фиксированном узле активной зоны, модель профиля ЭВ ТВС и модель пространственного поля ЭВ в активной зоне реактора

7 Для проведения исследований нейроструктурных моделей управляемых распределенных систем разработана специализированная программная система

8 Средствами разработанной программной системы проведены серии компьютерных экспериментов, демонстрирующих точность построенных нейроструктурных моделей энерговыделения и подтверждающих их обобщающие способности в ограниченном пространстве типовых управляющих воздействий

Результаты проведенных теоретических и экспериментальных исследований подтверждают применимость разработанного методического и алгоритмического обеспечения для решения задач идентификации управляемых систем с распределенными параметрами, а новый предложенный тип нейроструктурных математических моделей может найти применение и в других проблемных областях

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

- 1 Мишулина О А , Трофимов А Г Нейросетевое моделирование распределенных процессов энерговыделения в активной зоне реакторной установки // Нейрокомпьютеры разработка и применение – 2007 – № 9 – С 57-70
- 2 Two neural network models for nuclear engineering applications Two neural network models for nuclear engineering applications / О А Mishulina, V S Kharitonov, A G Trofimov, et al // Optical Memory & Neural Networks (Information Optics) – 2005 – Vol 14, № 1. – P 23-44
- 3 Нейросети в задаче траекторного управления динамическим объектом / О А Мишулина, А Г Трофимов, Б Е Федун и др // Авиакосмическое приборостроение – 2007 – № 1 – С 31-37
- 4 Мишулина О А , Трофимов А Г Нейросетевой прогноз выживаемости клеток при гипертермическом воздействии // Новости искусственного интеллекта – 2007

- 5 Пакет прикладных программ N-FONTO для построения нейросетевых моделей в ядерной энергетике / О А Мишулина, А Г Трофимов, В С Харитонов и др // Сборник научных трудов VII Всероссийской научно-технической конференции "Нейроинформатика-2005" – М МИФИ, 2005 – Т 2 – С 182-190
- 6 Костерев В В, Мишулина О А, Трофимов А Г Нейросетевая модель радиобиологического действия ионизирующих излучений в условиях гипертермии // Сборник научных трудов VIII Всероссийской научно-технической конференции "Нейроинформатика-2006" – М МИФИ, 2006 – Т 3 – С 188-196
- 7 Kosterev V V, Mishulina O A, Trofimov A G Neural network model of cell survival under hyperthermia // IFMBE Proceedings WC 2006 "World Congress on Medical Physics and Biomedical Engineering" – 2006 – Vol 14 – P 3393-3396
- 8 Мишулина О А, Трофимов А Г, Щербинина М В Нейросетевые структуры в задачах моделирования физических переменных // Труды 62-ой Научной сессии РНТОРЭС им А С Попова – 2007 – С 102-105
- 9 Мишулина О А, Трофимов А Г Нейросетевая структура для моделирования управляемых систем с распределенными параметрами // Сборник научных трудов X Всероссийской научно-технической конференции "Нейроинформатика-2008" – М МИФИ, 2008 – Т 2 – С 40-49
- 10 Разработка системы N-FONTO для решения задач ядерной теплоэнергетики / О А Мишулина, А Г Трофимов, М В Щербинина и др // Научная сессия МИФИ-2005 Сборник научных трудов – М МИФИ, 2005 – Т 2 – С 117-118
- 11 Нейросетевая модель прогноза энерговыделений в активной зоне реактора ВВЭР-1000 / О А Мишулина, А Г Трофимов, М В Щербинина и др // 14-й Международный научно-технический семинар "Современные технологии в задачах управления, автоматике и обработки информации" Сборник научных трудов – Самара СГАУ им Академика С П Королева, 2005 – С. 143
- 12 Трофимов А Г Синтез нейросетевой структуры для моделирования динамической системы с распределенными параметрами // 15-й Международный научно-технический семинар "Современные технологии в задачах управления, автоматике и обработки информации" Сборник научных трудов – М МИФИ, 2006 – С 166

- 13 Kosterev V V , Mishulina O A , Trofimov A G Cell survival data processing with artificial neural networks // 23-rd Annual Meeting of the European society for hyperthermic oncology Abstract Book – 2006 – P 114-115
- 14 Трофимов А Г Разработка системы NSTRUCT для моделирования и обучения нейросетевых структур // Научная сессия МИФИ-2007 Сборник научных трудов – М МИФИ, 2007 – Т 2 – С 12-13
- 15 Моделирование нейтронного потока в активной зоне ВВЭР с помощью нейросетевых технологий / О А Мишулина, А Г Трофимов, Б Е Шумский и др // Научная сессия МИФИ-2007 Сборник научных трудов – М МИФИ, 2007 – Т 8 – С 160-162
- 16 Neural model of cells equisurvival lines under hyperthermia / V V Kosterev, O A Mishulina, A G Trofimov, et al // 24-rd Annual Meeting of the ESHO Abstract Book – 2007 – P 89-90
- 17 Мишулина О А , Трофимов А Г Нейросетевая структура с форсирующими звеньями // 16-й Международный научно-технический семинар "Современные технологии в задачах управления автоматике и обработки информации" Сборник научных трудов – 2007 – С 144
- 18 Трофимов А Г , Скругин В И Разработка программной системы NeuroDynamics для нейросетевого моделирования динамических объектов // Научная сессия МИФИ-2008 Сборник научных трудов – М МИФИ, 2008 – Т 11 – С 109-110
- 19 Параллельная реализация алгоритмов обучения модульной нейросети и ее работы / В А Пересыпкин, А Г Трофимов, Е В Чешин и др // Научная сессия МИФИ-2008 Сборник научных трудов – М МИФИ, 2008 – Т 12 – С 121-122
- 20 Применение нейросетевой динамической модели аксиального профиля энерговыделения активной зоны ВВЭР-1000 для коррекции показаний ионизационных камер / О А Мишулина, А Г Трофимов, В С Харитонов и др // Научная сессия МИФИ-2008 Сборник научных трудов – М МИФИ, 2008 – Т 1 – С 113-115

Подписано в печать 28 04 2008 г
Печать трафаретная

Заказ № 348
Тираж 100 экз

Типография «11-й ФОРМАТ»
ИНН 7726330900
115230, Москва, Варшавское ш , 36
(495) 975-78-56, (499) 788-78-56
www.autoreferat.ru