

Федеральное государственное бюджетное учреждение науки
«Санкт-Петербургский Федеральный исследовательский центр
Российской академии наук»
(СПб ФИЦ РАН)

На правах рукописи



Милосердов Дмитрий Игоревич

**Модели, методы и архитектуры программных систем нейросетевого
прогнозирования трудноформализуемых событий
с непрерывным обучением**

Специальность 2.3.5 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель
д.т.н., профессор
Осипов В.Ю.

Санкт-Петербург – 2022

ОГЛАВЛЕНИЕ

Введение.....	4
1 Анализ процесса прогнозирования трудноформализуемых событий.....	14
1.1 Цели и условия прогнозирования трудноформализуемых событий.....	14
1.2 Анализ известных методов и средств прогнозирования трудноформализуемых событий.....	21
1.3 Логическая схема исследования.....	35
Выводы по главе 1.....	36
2 Модели и методы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.....	38
2.1 Обобщенная модель нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.....	38
2.2 Математические формулировки задач поиска целесообразных способов нейросетевого прогнозирования трудноформализуемых событий...42	
2.3 Метод нейросетевого прогнозирования с непрерывным обучением системы и с временными сдвигами сигналов.....	47
2.4 Метод нейросетевого прогнозирования без временных сдвигов с управлением направленностью вызова сигналов из ассоциативной памяти.....	55
Выводы по главе 2.....	61
3 Архитектуры программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.....	63
3.1 Архитектура программной системы нейросетевого прогнозирования с непрерывным обучением.....	63
3.1.1 Требования, предъявляемые к программной архитектуре.....	63
3.1.2 Структура, состав и предназначение элементов архитектуры.....	65
3.1.3 Правила функционирования программной архитектуры.....	70
3.2 Алгоритм рационального использования памяти для систем нейросетевого прогнозирования с непрерывным обучением.....	74

3.3 Способ ускоренного расчета синаптических весов в программных системах нейросетевого прогнозирования.....	80
Выводы по главе 3.....	82
4 Результаты моделирования и предложения по совершенствованию нейросетевого прогнозирования трудноформализуемых событий	84
4.1 Результаты и рекомендации по нейросетевому прогнозированию на примере предсказания событий, связанных с дорожным движением.....	84
4.2 Результаты и рекомендации по нейросетевому прогнозированию лексического содержания новостных лент.....	97
4.3 Результаты оценивания эффективности архитектурных решений по нейросетевому прогнозированию.....	101
4.4 Результаты использования разработанных методов и архитектур в исследованиях по обнаружению БПЛА.....	103
4.5 Программная система нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.....	105
4.6 Рекомендации по взаимодействию программной системы нейросетевого прогнозирования с другими системами.....	110
Выводы по главе 4.....	112
Заключение.....	114
Список сокращений и условных обозначений.....	118
Список литературы.....	119
Список иллюстративного материала.....	135
Приложение А. Список публикаций соискателя по теме диссертации.....	139
Приложение Б. Полученные свидетельства об интеллектуальной собственности.....	141
Приложение В. Акты внедрения результатов диссертационной работы.....	143

ВВЕДЕНИЕ

Актуальность темы исследования. Прогнозирование событий – актуальная научно-практическая задача, представляющая интерес во многих областях [1, 2]. Повышение сложности анализируемых процессов и свойственным им событий требуют все более совершенных инструментов прогнозирования. Задачи по управлению транспортом, экономикой, финансами, социальной сферой, сложными техническими объектами невозможно эффективно решать без получения точной и своевременной информации о ближайшем будущем, что определяет **важность** и **значимость** решаемой научной задачи. Потребителями прогнозов могут быть государственные органы, коммунальные и промышленные предприятия, средства массовой информации, веб-сервисы, автономные технические средства, классические и мобильные приложения, отдельные люди. Независимо от целей и задач прогнозирования, оно выполняется в условиях неопределенности ситуации, когда на интересующий параметр влияют прямые и косвенные факторы, изменяющиеся во времени. Информация об этих факторах зачастую не может быть представлена в виде набора простых аналитических моделей: она закодирована в большом массиве данных, а ее извлечение и обработка требуют новых, нестандартных подходов [3].

В настоящее время разрабатываются интеллектуальные системы и машины (СиМ), которые смогут взять на себя «творческие» задачи, включая прогнозирование. К таким СиМ относятся интеллектуальные роботы (ИР) различного предназначения [4, 5], интеллектуальные системы управления [6, 7] и другие приложения. Прогнозирование может осуществляться как путем непосредственного анализа временных рядов, так и путем построения и использования моделей процессов. Прогнозирование временных рядов событий с высокой неопределенностью динамики является одной из творческих задач, возлагаемых на интеллектуальные СиМ [8].

Степень разработанности темы. Несмотря на известные успехи методов искусственного интеллекта, задача прогнозирования трудноформализуемых

событий продолжает оставаться полем для исследований. Прогнозирование событий и нейросетевую обработку информации исследовали А.Н. Аверкин, В.Ю. Осипов, Б.В. Соколов, Я.А. Холодов, Р.М. Юсупов, С.Ф. Яцун, Р. Brockwell, E. Egrioglu, I. Goodfellow, L. Haitao, S. Haykin, H. Hu, S. Malik, F. Moretti, A. Sagheer, J. Schmidhuber, Z. Shen, K. Thurow, Y. Wu, B. Yang, T. Zhou и др.

Разработано большое число моделей, методов и средств прогнозирования, однако они не всегда удовлетворяют требованиям по точности, оперативности, глубине прогнозирования. Традиционные методы, такие как сглаживание, регрессионный и авторегрессионный анализ, успешно применимы при краткосрочном прогнозировании, однако не позволяют учитывать изменения в законах поведения временных рядов, описывающих происходящие события. Эвристические подходы зачастую являются вычислительно сложными или не универсальными. Гибридные подходы, в том числе с применением нечеткой логики, генетических алгоритмов, моделей прецедентных рассуждений, комбинации нескольких нейросетевых архитектур работоспособны при определенном наборе условий и не всегда демонстрируют свою эффективность в других ситуациях. Большинство моделей требует существенных временных затрат на обучение, что существенно ограничивает их применение в приложениях реального времени.

Потенциально широкими возможностями по прогнозированию обладают потоковые рекуррентные нейронные сети (РНС). Они позволяют обрабатывать выборки данных произвольной длины, а наличие обратных связей дает возможность реализовывать глубокую обработку информации.

Применение методов прогнозирования на основе нейронных сетей предусматривает обучение и непосредственно прогнозирование. При обновлении данных процесс обучения приходится запускать заново. Применительно к прогнозированию событий в интересах интеллектуальных приложений реального времени, где требуется постоянное обновление прогнозов с учетом вновь поступивших данных, такие системы потребляют значительные аппаратные

ресурсы. При этом в настоящее время отсутствуют нейросетевые системы, способные одновременно выполнять прогнозы и непрерывно обучаться.

Кроме этого требуют развития правила управления ассоциативным вызовом информации из памяти нейронных сетей при прогнозировании событий на основе длинных, коротких и зашумленных выборках временных рядов. Нерешенными остаются и проблемы обеспечения устойчивости многих РНС [9]. Для широко применяемых в настоящее время сетей LSTM характерна проблема накопления ошибок по мере увеличения длины обрабатываемой выборки [10].

Главный недостаток существующих методов заключается в том, что они не обеспечивают должного уровня точности прогнозов событий в условиях высокой неопределенности связанных факторов. Известные методы ограничены в возможностях по обработке информации.

Поскольку многие приложения должны работать в реальном времени, чрезвычайно важно не только получать точные прогнозы, но и иметь возможность непрерывного обучения системы прогнозирования событий для оперативного обновления прогнозных моделей. В настоящее время отсутствуют архитектуры, объединяющие в себе прогнозирование будущих событий и непрерывное обучение в целях формирования прогнозов в реальном времени с учетом вновь поступившей информации. Поэтому необходимо совершенствование научно-методического аппарата.

Целью диссертационной работы является повышение точности получаемых прогнозов трудноформализуемых событий.

Решаемая научная задача: разработка моделей, методов и архитектур программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.

Цель работы достигается совокупным решением следующих **частных задач:**

- 1) анализ процесса прогнозирования трудноформализуемых событий;
- 2) разработка модели системы прогнозирования трудноформализуемых событий с непрерывным обучением;

3) разработка методов прогнозирования трудноформализуемых событий с непрерывным обучением и управлением направленностью вызова сигналов из ассоциативной памяти;

4) разработка архитектур программных систем, реализующих методы прогнозирования трудноформализуемых событий с непрерывным обучением;

5) оценивание полученных результатов, выработка рекомендаций по повышению точности и использованию разработанных моделей, методов и программных систем.

Объектом исследования является процесс прогнозирования трудноформализуемых событий рекуррентными нейронными сетями.

Предметом исследования выступает научно-методический аппарат нейросетевого прогнозирования трудноформализуемых событий.

Научную новизну диссертационной работы составляют:

1. Модель системы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, отличающаяся своей структурой и правилами обработки сигналов, обеспечивающими оперативное прогнозирование с учетом изменений в законах проявления событий.

Модель содержит блок управления прогнозированием и две идентичные по своей структуре рекуррентные нейронные сети (РНС-1 и РНС-2), объединенные в систему. РНС-1 работает в режиме обучения, блок управления прогнозированием выполняет копирование обученной пространственно-временной модели событий из РНС-1 в РНС-2, а РНС-2 реализует прогнозирование. Предложенная модель обеспечивает непрерывность процесса обучения при прогнозировании. Это позволяет обеспечить работу в реальном времени и возможность постоянного формирования прогнозов с учетом изменяющихся законов поведения временных рядов. Отсутствует необходимость переобучения сети при поступлении новых данных. Исключается искажение пространственно-временной модели РНС из-за смены режимов ее функционирования.

2. Методы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением: с временными сдвигами сигналов и без временных

сдвигов сигналов, отличающиеся новыми правилами прогнозирования и управления ассоциативным вызовом информации из нейросетевой памяти и обеспечивающие высокую точность получаемых прогнозов трудноформализуемых событий.

Согласно методу с временными сдвигами, на вход РНС-1 подаются текущий и задержанный временные ряды. В процессе их прохождения вдоль слоев сети осуществляется ассоциативное пространственно-временное связывание прошлых и будущих событий. Блок управления прогнозированием копирует обученную модель из РНС-1 в РНС-2 и подает текущие входные данные в задержанный канал. В результате в текущем канале за счет вызова сигналов из ассоциативной памяти формируется прогноз будущих событий.

В методе прогнозирования без временных сдвигов на вход РНС-1 подается текущий временной ряд. При прохождении его по сети на ее элементах формируется модель событий, которая постоянно обновляется с учетом вновь поступающих данных. Блок управления прогнозированием копирует состояние РНС-1 в РНС-2 и запускает РНС-2 на формирование прогнозов по новым правилам, предусматривающим управление направленностью ассоциативного вызова сигналов из памяти нейронной сети. Согласно этим правилам, если обрабатываемая выборка признается короткой, то перед прогнозированием предлагается удлинять ее за счет ассоциативного вызова из памяти сети предшествующих значений.

3. Параллельная и буферная архитектуры программных систем, отличающиеся новой структурой и правилами функционирования программных систем прогнозирования с непрерывным обучением, обеспечивающие программную реализацию предложенных моделей и методов и расширение их функций.

В параллельной архитектуре эмулируется оба экземпляра нейронных сетей (РНС-1 и РНС-2). Буферная архитектура предполагает наличие только одного экземпляра (модуля) нейронной сети (РНС-1), называемого модулем эмуляции РНС-1 и РНС-2, а также входного буфера и модуля памяти для хранения

состояний нейронов РНС-2. Новизна буферной архитектуры состоит в отказе от выделения памяти для хранения синапсов РНС-2, в выполнении квазипараллельного обучения и прогнозирования, а достигаемый эффект заключается в сокращении объемов требуемой памяти в общем случае в два раза.

4. Практические рекомендации по повышению точности и использованию программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, обеспечивающие повышение точности прогнозов за счет определения наиболее эффективной конфигурации нейросетевых слоев применительно к задаче прогнозирования трудноформализуемых событий и разработки новых правил выбора метода и архитектуры в зависимости от условий, в которых функционирует система прогнозирования.

Теоретическая и практическая значимость работы. Теоретическая значимость полученных научных результатов состоит в развитии научно-методического аппарата прогнозирования трудноформализуемых событий рекуррентными нейронными сетями с непрерывным обучением. Практическая значимость этих результатов состоит в возможности повысить точность прогнозов возможных событий для различных приложений в условиях слабо формализуемых процессов с учетом большого числа неявно связанных факторов. Помимо повышения точности прогнозов, предложенные решения могут найти применение при проектировании перспективных интеллектуальных систем.

Методология и методы исследования. При выполнении диссертационного исследования использованы методы системного анализа и синтеза, интеллектуальной обработки данных, современная теория нейросетевого прогнозирования, а также методы оптимизации программных систем.

На защиту выносятся следующие положения:

1. Модель системы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.

2. Методы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением и управлением направленностью вызова сигналов из ассоциативной памяти.

3. Параллельная и буферная архитектуры программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.

4. Практические рекомендации по повышению точности и использованию программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.

Соответствие диссертации паспорту научной специальности. Представленные результаты соответствуют паспорту специальности 2.3.5 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»

Высокая степень достоверности научных положений обеспечена анализом текущего уровня исследований в данной области, корректным использованием апробированного математического аппарата, согласованностью теоретических выводов с результатами вычислительных экспериментов, сравнением предложенных решений с известными аналогами и одобрением основных положений диссертационной работы на международных и всероссийских научных конференциях.

Апробация и реализация результатов. Основные положения диссертационной работы представлялись на 6 международных конференциях «Digital Transformation And Global Society (DTGS-2019)» (Санкт-Петербург, 19-21 июня 2019 г.) [11], «Experimental Economics and Machine Learning (EEMML-2019)» (Пермь, 25-26 сентября 2019 г.) [12], 5-я Международная научно-практическая конференция «Технологическая перспектива-2019» (Санкт-Петербург, 7-8 ноября 2019 г.) [13], «Digital Transformation And Global Society-2020» (Санкт-Петербург, 17-19 июня 2020 г.), 6-я Международная научно-практическая конференция «Технологическая перспектива-2020» (Санкт-Петербург, 12-13 ноября 2020 г.) [14], 7-я Международная научно-практическая конференция «Технологическая

перспектива-2021» (Санкт-Петербург, 11-12 ноября 2021 г.) и всероссийской конференции «Информационные технологии в управлении» (ИТУ-2020) (Санкт-Петербург, 7-8 октября 2020 г.) [15].

Результаты диссертационной работы использованы в НИР СПб ФИЦ РАН №0073-2019-0001 «Теоретические основы и алгоритмические модели когнитивного управления, взаимодействия и анализа состояния групп гетерогенных робототехнических комплексов», а также в ЦСАМ АО «НТЦ РЭБ» при проведении научных исследований по обнаружению и траекторному сопровождению малоразмерных беспилотных летательных аппаратов для прогнозирования радиолокационной обстановки и выявления аномальных радиосигналов.

Публикации по теме диссертации. По научным результатам диссертационного исследования опубликовано 12 работ, в том числе 2 публикации в журналах из «Перечня рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук», одна из статей в указанном перечне опубликована без соавторов, 5 публикаций в зарубежных изданиях, индексируемых в Scopus/WoS (в том числе 2 публикации в журналах Q1), два свидетельства о государственной регистрации программы для ЭВМ, одно из которых зарегистрировано без соавторов.

Опубликованы статьи в следующих журналах из перечня рецензируемых научных изданий:

- «Информационно-управляющие системы»;
- «Информационные технологии».

Полный перечень публикаций и приравненных к ним работ представлен в Приложении А диссертации.

Личный вклад соискателя. Автором лично разработаны архитектуры и правила функционирования программных систем нейросетевого прогнозирования с непрерывным обучением. Лично автором разработаны практические рекомендации по повышению точности прогнозов и использованию программных

систем нейросетевого прогнозирования с непрерывным обучением. Модель и методы нейросетевого прогнозирования с непрерывным обучением разработаны в соавторстве с научным руководителем, причем вклад соискателя в совместных публикациях был значительным.

Структура и объём диссертации. Текст работы состоит из следующих структурных элементов: введение; основная часть, включающая четыре главы; заключение; список сокращений и условных обозначений; список литературы, содержащий 157 наименований; список иллюстративного материала; три приложения, содержащие список публикаций соискателя по теме диссертации, копии полученных свидетельств об интеллектуальной собственности, а также копии актов внедрения результатов диссертационной работы. Общий объём диссертационной работы – 145 страниц. Работа включает в себя 38 рисунков, 14 таблиц.

Краткое содержание работы.

В первой главе дана общая характеристика процесса прогнозирования трудноформализуемых событий, определены условия и возможности такого прогнозирования, проведен анализ известных моделей и методов. Сформулирована решаемая научная задача.

Во второй главе приведены формальные постановки типовых задач поиска целесообразных методов нейросетевого прогнозирования трудноформализуемых событий. Разработана обобщенная модель нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, а также два метода на основе предложенной модели: метод прогнозирования с временными сдвигами сигналов и метод без временных сдвигов с управлением направленностью вызова сигналов из ассоциативной памяти.

В третьей главе разработаны архитектуры программных систем, реализующих методы нейросетевого прогнозирования с непрерывным обучением. Приведена параллельная архитектура, отличающаяся наличием двух полноценных экземпляров РНС, работающих параллельно. В целях сокращения объемов требуемой памяти разработана буферная архитектура, в которой

используется квазипараллельное обучение и прогнозирование на одном экземпляре РНС. Приведен способ ускоренного расчета синаптических весов в программных системах нейросетевого прогнозирования.

В четвертой главе осуществлено моделирование и оценка эффективности предложенных методов и архитектур на примере задач прогнозирования транспортных потоков и лексического содержания новостных лент. Приведены результаты использования разработанных методов и архитектур в исследованиях по обнаружению БПЛА. Разработана программная реализация системы нейросетевого прогнозирования с непрерывным обучением. Выработаны практические рекомендации по повышению точности и использованию разработанных программных систем, а также рекомендации по взаимодействию разработанных программных систем с другими системами.

В заключении приведены выводы и результаты, полученные в ходе выполнения работы, даны рекомендации по применению разработанных моделей, методов и архитектур, определены перспективы дальнейшей разработки темы.

ГЛАВА 1. АНАЛИЗ ПРОЦЕССА ПРОГНОЗИРОВАНИЯ ТРУДНОФОРМАЛИЗУЕМЫХ СОБЫТИЙ

1.1 Цели и условия прогнозирования трудноформализуемых событий

Прогнозирование событий – актуальная научно-практическая задача, представляющая интерес во многих областях [1, 2]. Прогнозы необходимы при управлении транспортом, экономикой, финансами, социальной сферой, сложными техническими объектами. Они используются при принятии решений на различных уровнях, начиная от отдельных людей и приложений, автоматизированных систем, предприятий, отраслей, и заканчивая органами государственной власти и крупными корпорациями. Горизонты прогнозов могут составлять от нескольких секунд до нескольких лет. В зависимости от этого различаются цели прогнозирования. Примеры целей:

- повышение горизонта прогнозирования при ограничениях на ошибку;
 - увеличение точности прогнозов при ограничениях на ошибку и ресурсы, затрачиваемые на прогнозирование;
 - сокращение времени прогнозирования при обеспечении заданного уровня точности;
 - сокращение ресурсов, затрачиваемых на прогнозирование, при ограничениях на ошибку;
 - предоставление информации для прогнозов более высокого уровня;
- и другие.

Повышение сложности анализируемых процессов и свойственным им событий требуют все более совершенных инструментов прогнозирования [16–18]. Под событиями, подлежащими прогнозированию, понимаются привязанные к пространству и времени совокупности факторов [19]. В ряде случаев законы проявления событий трудно формализуемы и закодированы в объемном информационном потоке. Рассмотрим условия прогнозирования трудноформализуемых событий на примере предсказания в интересах автономных интеллектуальных роботов, а также предсказания новостных

информационных сообщений. Эти условия достаточно полно демонстрируются приведенными примерами, однако не ограничиваются ими, а имеют место и в других приложениях, где требуется прогнозирование в условиях высокой неопределенности.

Так, при прогнозировании событий в интересах интеллектуальных автономных роботов [20, 21] законы проявления событий неявно описываются показаниями с датчиков температуры, скорости, давления [22, 23], расходом заряда батарей [24, 25], расстоянием до препятствий [20, 26, 27]. Могут быть данные и от внешнего источника, передаваемые на робота посредством канала связи. Например, для прогнозирования пробок на городских дорогах в целях планирования своего маршрута беспилотный автомобиль может получать информацию о дорожно-транспортных происшествиях, о состоянии дорог, ремонтных работах, режимах работы светофоров, а также о текущей загруженности городских магистралей [4, 5].

В свою очередь, при анализе потоков новостных сообщений также приходится учитывать, что законы проявления событий записаны в больших массивах текстовых данных, и для предсказания их будущего содержания необходима глубокая обработка лексического состава [3].

Условия, которые имеют место при прогнозировании трудно формализуемых событий, включают следующие.

1. Горизонт прогнозов.

Эксплуатация современных робототехнических изделий отчетливо проявила проблемы, связанные с неспособностью автономных машин передвигаться в суровых и/или непредсказуемых средах [28–40]. Зачастую операторное управление ими затруднено или невозможно в силу большой задержки поступления управляющих сигналов [41]. Такая задержка может быть обусловлена большим расстоянием между роботом и пунктом управления, или другими причинами. Для роверов-планетоходов такая задержка уже сегодня исчисляется в минутах. Задержка между передачей управляющих команд для луноходов может составлять до 20 секунд [41], для марсоходов – от 5 до 20 минут

[42]. В перспективе при исследовании более удаленных планет эта задержка будет составлять несколько часов или даже дней. Заметим, что для наземных робототехнических устройств, функционирующих в условиях преднамеренных помех, задержка также может исчисляться в секундах и минутах. Для ряда перспективных роботов, таких как беспилотные автомобили, операторное управление предполагается использовать только в определенных ситуациях [43]. Таким образом, система прогнозирования должна быть способна давать адекватные прогнозы событий на горизонты, сопоставимые со временем автономности объекта, в интересах которого эти прогнозы выполняются.

Предсказание лексического содержания текстовых потоков в частности может использоваться для различных целей. По полученным данным могут корректироваться планы промышленных и торговых предприятий, государственных учреждений, отдельных людей. На сегодняшний день такая информация в основном используется в качестве вспомогательной при прогнозировании финансово-экономических показателей [44, 45], однако она может быть полезна и в других приложениях. Например, на основе прогнозирования и оценки будущего содержания новостей в интернете возможно предупреждение различных негативных социальных явлений [46]. Для того, чтобы прогнозы содержания новостных сообщений могли быть своевременно использованы, необходимо, чтобы горизонт прогноза был сопоставим со временем реакции на полученные новости и принятия мер.

2. Наличие пропусков и искажений во входных данных.

Использование робототехнических систем особенно актуально в условиях повышенной опасности или вредности окружающей среды, а также в условиях, где человек не может находиться по объективным причинам [47]. В таких условиях может происходить частичный отказ аппаратуры робота, который невозможно устранить с человеческой помощью. Сенсоры, ответственные за получение информации из внешней среды, могут выходить из строя полностью или частично, выдавая данные с помехами. Могут иметь место пропуски данных.

В условиях внешних помех (например, у роботов военного назначения) данные могут быть искажены или отсутствовать.

В свою очередь, потоки новостных сообщений также подвержены различного рода искажениям и пропускам. Для текстовых данных, описывающих новостные события, характерны ошибки, начиная опечатками в словах и заканчивая преднамеренными вбросами недостоверных новостей. Информация о новостях появляется с задержкой, новостные сервисы могут быть недоступны по техническим причинам. Могут изменяться стилистические особенности публикуемых текстов, появляться и исчезать блоки контекстной рекламы.

Существующие методы не всегда в состоянии выполнять адекватную обработку информации в рассмотренных случаях, однако система прогнозирования должна быть способна обрабатывать информацию даже в условиях зашумленности, частичного отсутствия или искажения данных.

3. Учет взаимосвязей между элементами обрабатываемых данных.

Не менее важным условием функционирования системы прогнозирования является необходимость учета взаимосвязей совокупности внешних и внутренних факторов, которые имеют место в среде функционирования объекта, в интересах которого выполняются прогнозы.

На примере автономного робота, к внешним факторам можно отнести следующие. Это проходимость местности, наличие препятствий или опасных зон, температура окружающей среды, присутствие подвижных объектов и характер их движения. Также могут иметь место периодически появляющиеся факторы (например, световые вспышки, взрывы и т.п.). К внутренним факторам относятся: техническое состояние робота (наличие повреждений, отказавшего оборудования), заряд батарей (объем топлива), текущая дополнительная нагрузка (вооружение и т.п.), другие. Перечень факторов, влияющих на робота, зависит от его предназначения и возможностей. Чем больше этих факторов будет учтено, тем точнее прогноз и выше эффективность робота. Современные методы не в состоянии учесть все необходимые факторы, или же их учет требует неприемлемо больших вычислительных затрат.

При прогнозировании новостей важно учитывать взаимосвязи между отдельными элементами обрабатываемого информационного блока. В текстовых данных этими элементами являются отдельные слова [48]. Взаимосвязи слов могут нести в себе информацию о будущих событиях, поэтому необходимы методы, которые способны обработать и ассоциативно связать такие элементы.

Система прогнозирования должна быть способна обрабатывать разнородные данные, несущие информацию о факторах, которые имеют значение в текущий момент, выявлять скрытые внутренние закономерности между ними и выполнять прогнозы, основываясь на этой информации.

4. Учет изменений в законах проявления событий.

Одной из ключевых проблем создания перспективных автономных интеллектуальных машин является наделение их возможностью работы в условиях присутствия динамических препятствий и «загроможденной» местности [49]. Природа объекта и характер изменения положения таких препятствий оказывают влияние на степень его предсказуемости. Наличие дополнительных объектов в поле зрения робота и их взаимодействие с движущейся целью существенно осложняют прогноз. Наиболее же сложной и труднопредсказуемой категорией динамических объектов для интеллектуального робота являются другие роботы, находящиеся в области его видимости [26]. Они не только двигаются по труднопредсказуемой траектории, но и нацелены на взаимодействие с объектами окружающего мира, что делает задачу прогнозирования событий вдвойне нетривиальной. Проблема коллективного взаимодействия может быть относительно просто разрешена при наличии у группы роботов общего управляющего центра. Однако в ряде случаев такого центра нет, и роботы принадлежат к различным или даже враждебным группам. В таких ситуациях разработчики систем автономной навигации сталкиваются с существенными трудностями. Несмотря на сложность проблемы, ее решение является чрезвычайно важным, поскольку будущее лежит именно за кооперативными робототехническими системами [50].

При прогнозировании содержания новостных блоков во многих случаях имеется набор тематик, которые охватывает обрабатываемое информационное сообщение [51]. Переменным во времени является как состав тем, так и занимаемая ими доля, а также тематическая направленность внутри каждой из тем. Темы могут быть неявно связаны друг с другом, что также необходимо учитывать при прогнозах. К примеру, вслед за новостями о ядерных испытаниях в Северной Корее зачастую следуют сообщения о ухудшении инвестиционного климата в странах-соседях.

Из вышеприведенных примеров следует, что система прогнозирования должна не только обеспечивать точные прогнозы, но и одновременно с этим учитывать изменения в факторах, оказывающих влияние на будущие события.

5. Оперативность выполнения прогнозов.

Неотъемлемой чертой перспективных интеллектуальных машин является оперативность принятия решений. Поскольку среда, в которой функционирует робот, является сложной и динамичной, не исключены ситуации, в которых на принятие решения отведено очень ограниченное время. События, представленные в новостных сообщениях, также могут развиваться очень стремительно и требуют быстрой реакции. Это особенно заметно в новостях, связанных с чрезвычайными происшествиями и рынком ценных бумаг. Однако существующие подходы к прогнозированию либо вычислительно сложны, либо требуют переобучения при изменениях в прогнозируемой среде. В обоих случаях между возникновением нового фактора и принятием решения проходит значительное время, что делает существующие методы неприменимыми к прогнозированию трудноформализуемых событий. Система прогнозирования должна быть достаточно оперативной, чтобы давать актуальные прогнозы.

6. Ограниченность аппаратных ресурсов для выполнения прогнозирования.

Аппаратные и программные ресурсы, выделяемые на прогнозирование, не всегда позволяют своевременно получать интересующие прогнозы. Помимо того факта, что бортовой компьютер робота загружен многими задачами, не связанными с планированием поведения, его возможности могут быть

ограничены условиями функционирования. Например, для роверов-планетоходов требуется специальная аппаратура, устойчивая к космической радиации. По этой причине бортовые компьютеры даже современных планетоходов в десятки раз слабее обычных бытовых вычислителей. К примеру, марсоход «Perseverance» имеет на своем борту компьютер с процессором RAD750 с тактовой частотой 200 МГц, 256 мегабайт оперативной памяти и 2 гигабайта памяти для хранения информации [52]. Робот «BigDog» для поддержки пехоты в труднодоступных условиях оснащен компьютером на базе PC104 Pentium для небольших систем [53]. Современные беспилотные летательные аппараты (БПЛА) также имеют крайне скромные аппаратные возможности, что ограничивает выполнение полезных приложений на их бортовых компьютерах [54].

При прогнозировании новостных событий необходим учет широкого контекста. Взаимосвязанность событий из разных сфер жизни приводит к тому, что для точного предсказания требуется обрабатывать одновременно большое количество текстовых блоков из разных тематических сфер. При этом общий объем такой информации экспоненциально увеличивается [55]. Требуется не только учет самих слов, но и связей между ними, что в свою очередь также многократно усложняет обработку. В таких условиях применение известных методов обработки естественного языка на аппаратных мощностях стандартных компьютеров представляет серьезные трудности [56, 57].

Из этого следует, что система прогнозирования событий должна быть не только достаточно интеллектуальной и адаптивной, но и подходить для системы с ограниченными ресурсами [58, 59].

7. Высокий ущерб при недостаточной точности прогнозов.

Непременным условием прогнозирования событий во многих случаях является высокий ущерб в случае недостаточной точности и своевременности прогнозов. В частности, робототехнические системы отличаются высокой стоимостью, а устранить проблемы, возникшие по причине неадекватности прогнозов и, соответственно, выбранного маршрута, на месте не представляется возможным. Повреждения оборудования, не говоря уже об авариях, являются

высокой ценой за некорректное планирование действий. В свою очередь, за некорректным прогнозированием новостных событий могут последовать неверные управленческие решения, ведущие к большим финансовым потерям или различного рода чрезвычайным ситуациям. В связи с этим от системы прогнозирования событий требуется не только оперативность, но и высокая точность.

Вышеизложенные условия прогнозирования событий в той или иной степени относятся и к другим областям прогнозирования трудноформализуемых событий. Требования к величине горизонтов прогнозов, работа в условиях искажений и пропусков в данных, ошибок программного обеспечения и алгоритмов предварительной обработки, внешних и внутренних факторов, влияющих на объекты прогнозирования, делают получение информации о будущих событиях нетривиальной задачей. Во многих случаях необходимо обеспечить быстроту принятия решений и непрерывность обучения системы прогнозирования при весьма ограниченных аппаратных возможностях используемой ЭВМ. В настоящий момент разработан широкий круг подходов к прогнозированию. Однако они во многом не учитывают весь спектр рассмотренных выше условий прогнозирования событий. Рассмотрим эти методы более подробно.

1.2 Анализ известных методов и средств прогнозирования трудноформализуемых событий

Выбор метода прогнозирования параметров обусловлен спецификой задачи, особенностями среды, в которой функционирует система прогнозирования, требованиями по точности, быстродействию, а также программно-аппаратными возможностями оборудования, на котором развернута система [60]. Неослабевающий интерес к изучению различных аспектов прогнозирования событий нашел отражение в многочисленных исследованиях как зарубежных, так и российских авторов.

Всю совокупность методов прогнозирования событий по степени формализации можно разделить на интуитивные и формализованные методы (рис. 1). К интуитивным методам относятся методы индивидуальных и коллективных экспертных оценок. «Они применяются тогда, когда объект прогнозирования либо слишком прост, либо настолько сложен, что аналитически учесть влияния многих факторов практически невозможно» [61].



Рисунок 1 — Классификация методов прогнозирования событий по степени формализации

В свою очередь, формализованные методы делятся на экстраполяционные, системно-структурные, ассоциативные методы, а также методы опережающей информации. Системно-структурные методы включают в себя функционально-иерархическое моделирование, морфологический анализ, матричный анализ, сетевое моделирование и структурный анализ. Ассоциативные методы делятся на имитационное моделирование и историко-логический анализ, а в группу методов опережающей информации входят такие методы, как анализ потоков публикаций, оценка значимости изобретений и анализ патентной информации [61]. Экстраполяционные методы основываются на анализе временных рядов —

«упорядоченных во времени наборов измерений тех или иных характеристик исследуемого объекта, процесса» [61].

Существующие на сегодняшний день методы прогнозирования временных рядов, согласно [62], делятся на традиционные методы, методы на основе нейронных сетей и методы на основе нечеткой логики. Структуру этих методов можно представить в виде рис. 2.

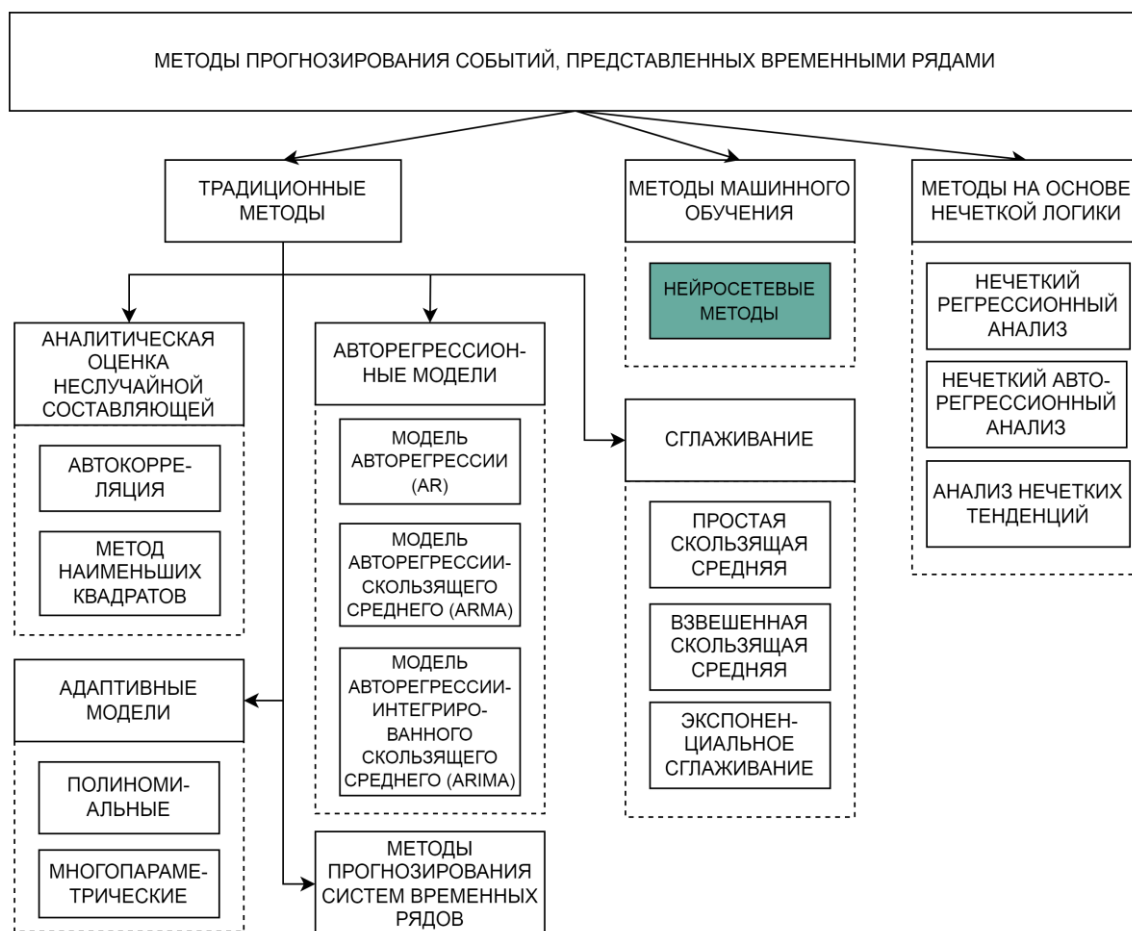


Рисунок 2 — Классификация методов прогнозирования временных рядов

Традиционные методы прогнозирования временных рядов [61, 63–66] заключаются в определении основных аналитических параметров временного ряда, влияющих на его динамику, и экстраполяции на основе известных предыдущих и настоящих значений. Выделяют следующие основные группы методов:

- сглаживание;
- аналитическая оценка неслучайной составляющей;

- авторегрессионные модели;
- адаптивные модели;
- методы прогнозирования систем временных рядов.

Сглаживание временных рядов являются одними из наиболее простых методов. Выделяют сглаживание простой скользящей средней, взвешенной скользящей средней, экспоненциальное сглаживание и их разновидности [61, 67].

Аналитическая оценка неслучайной составляющей временного ряда [63, 64] включает в себя:

- автокорреляцию уровней, когда определяется корреляционная зависимость между последовательными уровнями временного ряда. «Автокорреляция количественно измеряется с помощью коэффициента корреляции между уровнями исходного временного ряда и уровнями этого ряда, сдвинутыми на несколько шагов назад во времени. Коэффициенты автокорреляции характеризуют тесноту линейной связи текущего и предыдущего уровней ряда. По знаку коэффициента автокорреляции нельзя сделать вывод о возрастающей или убывающей тенденции в уровнях ряда. Если наиболее высоким оказался коэффициент автокорреляции первого порядка, исследуемый ряд содержит только тренд. Если наиболее высоким оказался коэффициент автокорреляции порядка τ , ряд содержит колебания с периодом τ . Если все коэффициенты автокорреляции оказались незначимыми, то либо ряд не содержит тренда и периодических колебаний, либо ряд содержит сильную нелинейную тенденцию» [68];

- прогнозирование по линии тренда. Оно осуществляется на основе кривых роста, когда выбирается кривая и оцениваются ее параметры, а затем на их основе выполняется прогноз [69–71];

- прогнозирование периодической компоненты, которая может быть оценена как разность между наблюдаемыми значениями и значениями, сглаженными с использованием соответствующих скользящих средних.

Модель авторегрессии – скользящего среднего и ее разновидности (AR, ARMA, ARIMA и др.) [62, 72]:

– модель авторегрессии (AR) – модель временного ряда, в которой значения в данный момент линейно зависят от предыдущих значений этого же ряда;

– модель авторегрессии – скользящего среднего (ARMA) используется для анализа стационарных временных рядов. Она обобщает в себе более простые модели авторегрессии и скользящего среднего;

– модель авторегрессии – интегрированного скользящего среднего (ARIMA) – модификация алгоритма ARMA для нестационарных временных рядов.

Адаптивное прогнозирование предполагает, что в процессе прогнозирования осуществляется учет вновь поступающей информации и корректировка прогноза. Адаптивные модели могут быть полиномиальными и многопараметрическими [68].

При прогнозировании в некоторых случаях приходится учитывать взаимосвязи различных элементов, состояние каждого из которых описывается отдельным параметром. Совокупности значений этих параметров, регистрируемых в дискретные моменты времени, носит название систем временных рядов. Для прогнозирования таких систем существуют следующие методы [68]:

– трендовая модель системы временных рядов – для каждого ряда, входящего в систему, определяется наличие тренда. Если тренд наблюдается, то для соответствующих временных рядов строится несколько конкурирующих парных моделей, из которых выбирается оптимальная. Если же тренд не обнаружен, то временные ряды сглаживаются скользящей средней;

– совместная гармоническая модель строится для временных рядов, в которых выявлена циклическая составляющая;

– модель векторной авторегрессии – это модель, одновременно описывающая поведение нескольких совместно зависимых временных рядов через изменение собственных предшествующих значений и значений других совместно зависимых рядов.

Как правило, статистические модели применимы для получения краткосрочных прогнозов. В соответствии с этим подходом, когда меняются

законы динамики временных рядов, требуется строить новые модели, адекватные текущей ситуации. Данная особенность делает такие методы малоприменимыми для прогнозирования событий условиях высокой динамичности и неопределенности интересующего параметра. Эти методы не способны реагировать на изменения внешней обстановки с достаточной оперативностью. Поэтому на сегодняшний день наиболее перспективным подходом является использование эвристических моделей [73, 74] на основе накопленного опыта.

Методы нечеткой логики являются одним из активно развивающихся в настоящий момент подходов к прогнозированию временных рядов [75]. «Представление временных рядов в классе нечетких временных рядов основывается на предположении, что возможна лингвистическая интерпретация значений временного ряда, основанная на теории нечетких множеств» [76]. Такая интерпретация может выполняться как на основе абстрактных лингвистических оценок, так и с привлечением экспертов. «Song и Chissom предложили нечеткие модели детерминированных и авторегрессионных временных рядов первого порядка и применили разработанные модели для прогнозирования количества регистрирующихся студентов университета штата Алабама, фазифицировав предварительно четкий временной ряд. Перфильевой И. была разработана методика нечеткого сглаживания временных рядов, которая может быть отнесена к методикам нечеткого приближения на основе нечеткого преобразования» [76].

Среди существующих подходов наиболее гибким и интеллектуальным является подход, основанный на машинном обучении [63, 64, 77–79]. Среди них особым потенциалом обладают искусственные нейронные сети (ИНС) [78, 80–83]. Они нашли широкое применение в роботизированных системах [22–25]. Хайкин С. отмечает следующие полезные свойства нейронных сетей [84]: нелинейность, отображение входной информации в выходную, адаптивность, очевидность ответа, контекстная информация, масштабируемость, отказоустойчивость, единообразие анализа и проектирования, аналогия с нейробиологией. Труды [84, 85] обосновывают возможность успешного применения ИНС для аппроксимации функций, прогнозирования, классификации, фильтрации, управления, а также в

качестве ассоциативной памяти. В качестве примера успешной практической реализации системы, выполняющей локализацию и распознавание объектов на изображении, можно привести архитектуру YOLO [86]. Примеры применения ИНС в задачах фильтрации и управления представлены в [87, 88].

Решения на основе нейронных сетей можно классифицировать по способу воплощения: на аппаратные и программно-аппаратные (рис. 3).

Аппаратные основываются на применении нейронных чипов.

Известен проект «Spinnaker» (сокращение от Spiking Neural Network Architecture) и одноименная архитектура нейроморфного компьютера [89] развиваемая в Advanced Processor Technologies Research Group (APT) университета Манчестера под руководством профессора С. Б. Фарбера. SpiNNaker направлен на создание массово параллельного миллионно-ядерного компьютера, архитектура межсоединений которого вдохновлена характеристиками связности мозга млекопитающих и который подходит для моделирования крупномасштабных нейронных сетей в биологическом реальном времени. В частности, его архитектура позволяет передавать очень большое количество маленьких пакетов данных, каждый из которых явно передает источник и неявное время одного нейронного потенциала действия или «всплеска» [90].

Еще в 2014 году IBM Research представила нейроморфный процессор TrueNorth [91], способный эмулировать 1 млн. цифровых нейронов и 256 млн. синапсов, реализуемых 4096 ядрами. Два года спустя было представлено исследование [92], в котором продемонстрирована рабочая сверточная нейронная сеть реального времени на базе этого нейропроцессора (Рис. 4). Такая сеть может использоваться в качестве средства извлечения ключевых признаков состояния исследуемой системы и последующего прогнозирования [93].

Имеются и отечественные разработки в этой области. В 2020 году ученые из Тюменского университета опубликовали статью [94], в которой описали концепцию биоморфного нейропроцессора. Как утверждают авторы, он способен генерировать новые знания по биологически подобному механизму, что

качественно отличает его от предшествующих разработок, направленных в первую очередь на повышение производительности и энергоэффективности.

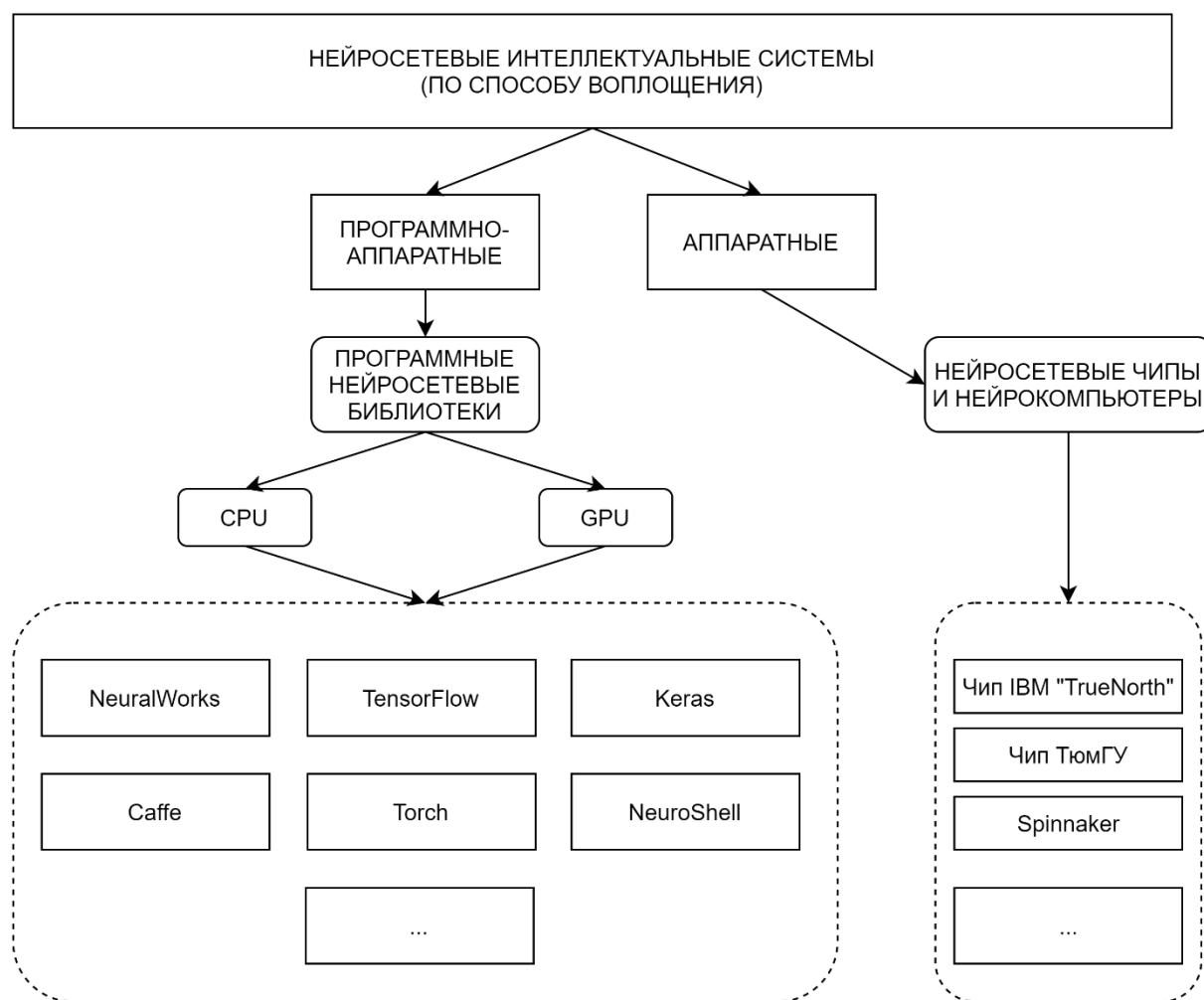


Рисунок 3 — Классификация способов реализации нейросетевых интеллектуальных систем

Аппаратный подход к реализации нейросетевых архитектур в настоящее время не представлен массово и является скорее объектом экспериментов. Подавляющее большинство работ связано с программно-аппаратным подходом, и эти исследования направлены на применение нейросетевых методов к решению различных прикладных задач.

Для типовых нейросетевых архитектур разработан ряд программных пакетов, содержащих наборы функций для эмуляции нейронных сетей. Поскольку нейронные сети по сути представляют собой параллельный распределенный

процессор, для их обучения нашли широкое применение не только центральные (CPU), но и графические процессоры (GPU). Примеры нейросетевых пакетов: NeuralWorks, Caffe, Keras, Tensorflow, Torch, NeuroShell.

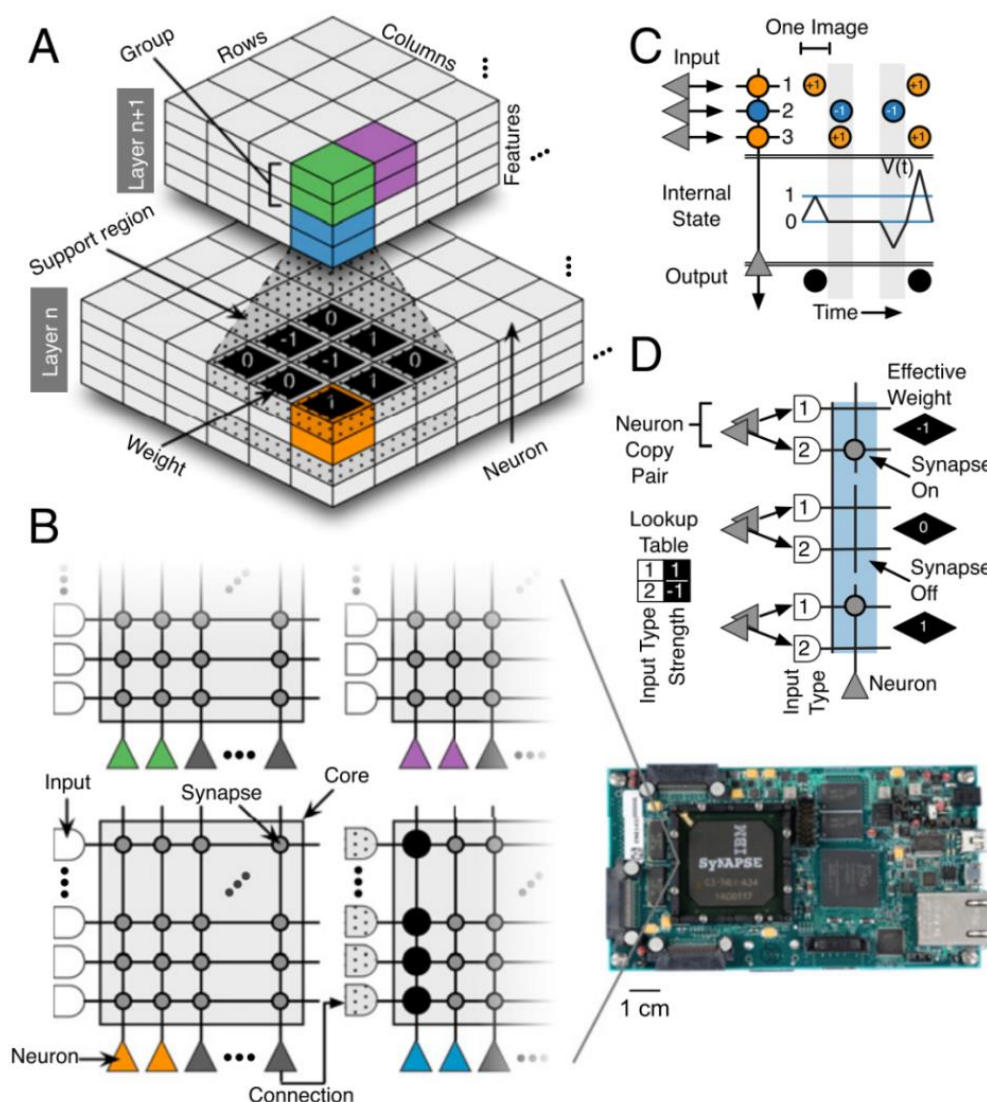


Рисунок 4 — Сверточная нейронная сеть, реализованная на базе нейрочипа TrueNorth: А – слои сверточной нейронной сети; В – структура синапсных ядер; С – динамика нейронов; D – веса сверточного фильтра [92]

Архитектуры нейронных сетей (НС), применяемые для прогнозирования событий, описываемых временными рядами, делятся на две основные группы: НС прямого распространения [95–99] и рекуррентные нейронные сети (РНС) [10, 100–104]. Стоит отметить, что в настоящее время наблюдается тенденция объединять ИНС с другими подходами [105, 106], поэтому такие решения могут быть рассмотрены в качестве отдельной группы гибридных методов (рис. 5).

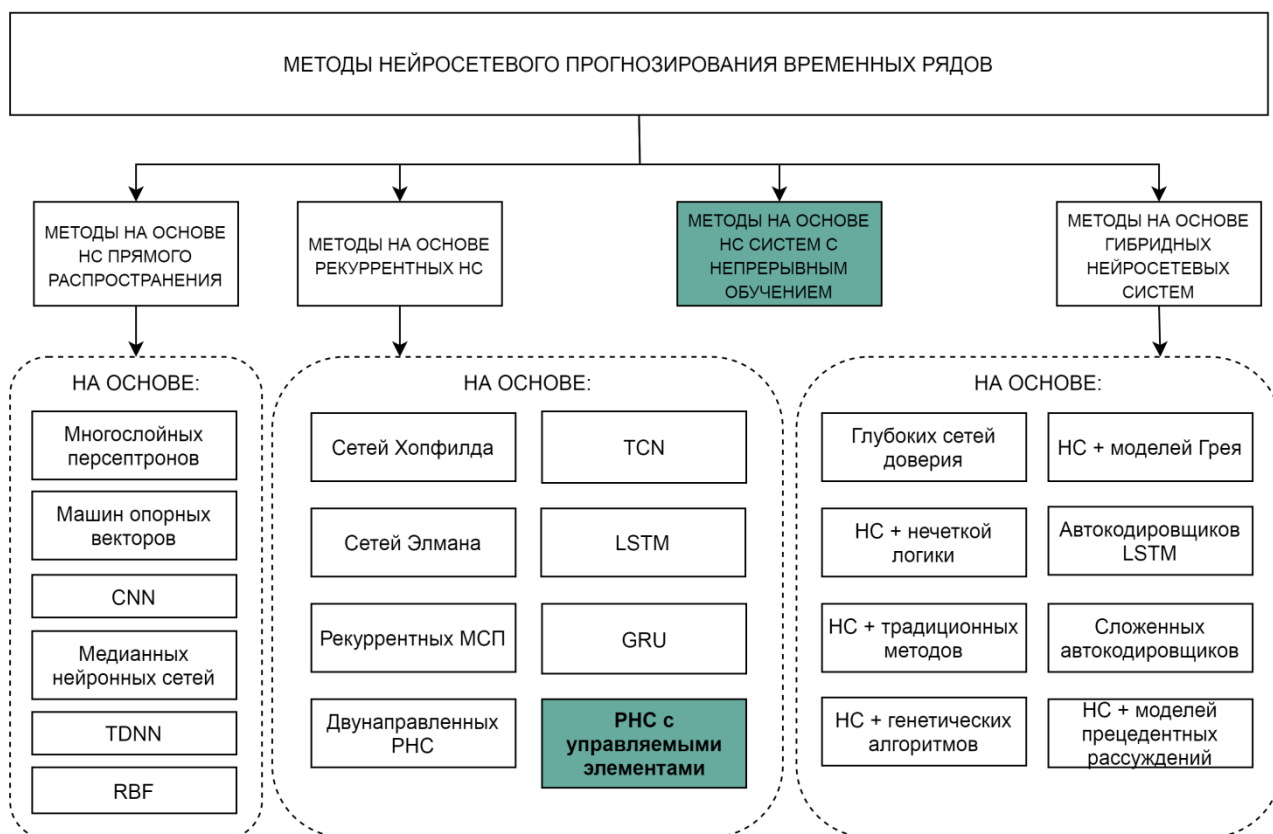


Рисунок 5 — Классификация систем нейросетевого прогнозирования временных рядов

К первой группе относятся ИНС прямого распространения. Такие сети позволяют строить как линейные, так и нелинейные регрессионные модели. С точки зрения функциональности, их можно считать эквивалентными традиционным аналитическим моделям. За многие годы выработаны эффективные рекомендации по созданию нужной архитектуры и организации процесса обучения таких архитектур. Среди последних работ, касающихся применения ИНС прямого распространения для прогнозирования, стоит упомянуть [95–97, 107–110].

Широкое применение в задачах прогнозирования нашли многослойные перцептроны, сверточные сети, сети радиальных базисных функций (РБФ), машины опорных векторов. Суть метода опорных векторов заключается в переводе исходных данных в пространство более высокой размерности и поиске разделяющей гиперплоскости с наибольшим зазором в этом пространстве. Метод опорных векторов в основном используется для краткосрочного прогнозирования

финансовых временных рядов [111–113]. В работе [95] рассматриваются возможности краткосрочного прогнозирования интенсивности движения с использованием сети прямого распространения с двумя скрытыми слоями. Эффективность применения многослойных перцептронов обосновывается в [114]. В [115] используется итеративный подбор оптимальных параметров многослойного перцептрона, позволяющий добиться высокой точности прогнозирования. В работе [116] исследуется применимость трехслойного перцептрона к экономическому анализу на основе валютных индикаторов. Сверточные сети используются в [117]. В работе [97] предлагается модифицировать Пи-Сигма нейронную сеть, используя медианную модель нейрона вместо аддитивной, для повышения точности прогнозирования временных рядов с выбросами.

Основным преимуществом НС прямого распространения является простота их внедрения и обучения. Их главным недостатком является невысокая точность прогноза для больших горизонтов. Поэтому в современных приложениях классические нейронные сети прямого распространения зачастую объединяются с другими методами [69, 80].

Объединение модели Грея и нейронной сети исследовалось в работе [98], ИНС и модели ARIMA – в [118]. В [77] упомянуты комбинации нейросетевого подхода с моделями прецедентных рассуждений. В работе [99] для прогнозирования оптимальные параметры нейронной сети подбирались с использованием группового метода обработки данных. В [107] для прогнозирования транспортных потоков предлагается метод, сочетающий использование линейной модели на основе L1-регуляризации и последовательности слоев гиперболического тангенса. Данный подход позволил улавливать нелинейные пространственно-временные эффекты в анализируемых временных рядах, тем самым повысив точность прогнозирования. Применяются гибриды сетей прямого распространения и нечетких систем [108, 109]. В статье [119] предложена нейросетевая модель прогнозирования, в основу которой положен алгоритм оптимизации параметров нейронной сети на основе

улучшенного синусно-косинусного метода. В работе [110] многослойный перцептрон применяется в сочетании с простым статистическим подходом. Сравнительное описание возможностей прогнозирования транспортных потоков с использованием регрессии вектора релевантности на основе хаотической имитационной оптимизации отжига, а также регрессии опорного вектора и нейронной сети обратного распространения вместе с генетическими алгоритмами приведено в работе [83].

Для прогнозирования временных рядов активно применяются гибридные системы на основе автокодировщиков. Повышение достоверности прогнозирования в условиях высокой неопределенности событий и аномалий в обрабатываемых данных в работе [96] исследуется на примере сложенного автокодировщика. В [120] используется сложная модель кодера-декодера на основе сетей долгой краткосрочной памяти.

Определенное место в системе гибридных методов прогнозирования занимают эволюционные алгоритмы. Так, в работе [121] для оптимизации весов сети РБФ использовался генетический подход. В [122] предложена модель, объединяющая ИНС обратного распространения с алгоритмом адаптивной дифференциальной эволюции для устранения проблемы локального минимума при обучении. Объединение многослойного перцептрона с многоцелевой оптимизацией роя частиц и элитарным генетическим алгоритмом сортировки без доминирования предложено в [123].

К гибридным решениям относятся и те, которые включают в себя нейронные сети более простых архитектур. Глубокие сети доверия, построенные слоях ограниченной машины Больцмана, приведены в [103].

В ряде случаев гибридные методы позволяют нивелировать недостатки отдельных подходов, но при этом являются структурно более сложными и/или менее универсальными. Обзорные работы [79, 124–125] демонстрируют, что ни один из существующих гибридных подходов не является доминирующим, и итоговая эффективность зависит во многом от конкретных ситуаций и условий, в

которых они применяются. Такие методы не дают уверенности в эффективной работе в непредсказуемых условиях.

Расширенными возможностями в прогнозировании временных рядов обладают рекуррентные НС [126]. Наличие обратных связей позволяет осуществлять ассоциативное пространственно-временное связывание информации в таких НС, тем самым повышается глубина обработки информации и, соответственно, точность прогнозов на больших горизонтах.

Ряд публикаций представляет сравнительный анализ возможностей различных типов РНС [127, 128]. Обучение таких сетей может проходить с учителем или без учителя. В первом случае успешность применения обученных сетей с зафиксированными весами синапсов обеспечивается только при стабильности законов изменения значений прогнозируемых параметров, в противном случае их эффективность снижается. Рекуррентные нейронные сети, обучаемые без учителя, лишены этого недостатка. Они модифицируют свою внутреннюю структуру в процессе функционирования, а ассоциативность памяти обеспечивает глубокое выявление взаимосвязей анализируемых временных рядов. При этом эффективность прогнозирования событий рекуррентными нейронными сетями зависит от их структуры, степени обучения, управления ассоциативными взаимодействиями обрабатываемых сигналов и других факторов. Большинство известных РНС характеризуются существенно ограниченными возможностями по формированию, ассоциативному запоминанию и извлечению из памяти моделей наблюдаемых событий. Для них характерны слабая устойчивость, а также трудности при ассоциативном вызове ранее запомненной информации.

Для прогнозирования событий в исследованиях [10, 100–102, 104, 127–131] применялись многослойные рекуррентные перцептроны, сети долгой краткосрочной памяти (Long Short-Term Memory – LSTM), ассоциативная память Хопфилда, двунаправленная ассоциативная память (Bidirectional associative memory – BAM) и другие. Особенности использования рекуррентных нейронных сетей Элмана для прогнозирования временных рядов раскрыты в работе [102]. В [104] применялись мультипликативные рекуррентные нейронные сети. В работах

[10, 100, 101, 129–132] рассмотрены возможности одноуровневого и иерархического прогнозирования транспортных потоков с использованием LSTM. Вариант сверточных нейронных сетей с рекуррентностью, называемый «temporal convolutional networks (TCNs)», применяемый для обработки последовательности данных, упомянут в [133]. Сравнительная характеристика ряда известных методов прогнозирования с использованием РНС отражена в [127, 128, 130].

Несмотря на то, что применение нейронных сетей является наиболее перспективным подходом к прогнозированию, возможности имеющихся архитектур остаются весьма ограниченными. Успешность их применения в значительной степени зависит от особенностей человеческой инженерии, и они не всегда могут быть настроены и применены без помощи экспертов предметной области [125]. Более того, эти методы зачастую требуют довольно большого объема данных для обучения, что в существующих условиях не всегда может быть обеспечено [134].

НС прямого распространения по своим возможностям во многом аналогичны традиционным статистическим методам. Рекуррентные нейронные сети предназначены для обработки временных рядов и являются более приемлемым решением. Они обеспечивают более глубокую обработку информации во времени, но для них характерны проблемы, связанные с устойчивостью. Не отработаны правила ассоциативного вызова информации из памяти таких сетей. Не все архитектуры способны осуществлять параллельную обработку разнородных данных, а также отсутствуют подходы, которые позволили бы одновременно обучать и использовать РНС для выполнения прогнозов. Для архитектуры LSTM, являющаяся на сегодняшний день одной из наиболее эффективных, характерна проблема накопления ошибок [10], при этом ее вычислительная стоимость остается высокой [135].

Из изложенных в предыдущем параграфе условий прогнозирования событий следует, что необходимы системы, не только осуществляющие прогнозирование с высокой точностью, но и непрерывно обучающиеся. Однако в

настоящий момент отсутствуют системы прогнозирования с непрерывным обучением.

Таким образом, необходимы новые модели и методы прогнозирования событий с возможностью непрерывного обучения.

Разрабатываемые в последние годы рекуррентные нейронные сети с управляемыми элементами [19, 136–140], которые являются самообучающимися, имеют значительный потенциал для решения задачи прогнозирования событий. В отличие от известных аналогов, такие РНС могут наделяться линейными, спиральными, петлевыми и другими логическими структурами. Они могут параллельно обрабатывать разнородные данные, при этом за счет надления слоев РНС различными конфигурациями возможно управление пространственно-временным связыванием и ассоциативным вызовом информации из памяти таких сетей. Необходимо разработать модели, методы и архитектуры программных систем прогнозирования с непрерывным обучением, которые можно будет применять в составе интеллектуальных систем прогнозирования временных рядов в условиях высокой неопределенности ситуации и присутствия большого числа переменных факторов.

1.3 Логическая схема исследования

Анализ исследуемого процесса и существующих методов прогнозирования временных рядов позволяет утверждать о необходимости совершенствования этих методов. Требуется поиск новых моделей, методов и архитектур программных систем прогнозирования трудноформализуемых событий с непрерывным обучением. Как показано в предыдущем параграфе, известные методы и подходы не в полной мере применимы для прогнозирования временных рядов, поскольку эти методы не учитывают имеющихся возможностей и условий такого прогнозирования.

Необходимо решить научную задачу разработки новых моделей, методов и архитектур программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, повышающих

точность прогнозирования. Новые модели, методы и архитектуры систем прогнозирования должны также обеспечивать возможность непрерывного обучения, позволяющего гибко учитывать законы поведения временных рядов.

Для решения сформулированной научной задачи в ходе выполнения диссертационных исследований предусматривались:

- 1) анализ процесса прогнозирования трудноформализуемых событий;
- 2) разработка модели системы прогнозирования трудноформализуемых событий с непрерывным обучением;
- 3) разработка методов прогнозирования трудноформализуемых событий с непрерывным обучением и управлением направленностью вызова сигналов из ассоциативной памяти;
- 4) разработка архитектур программных систем, реализующих методы прогнозирования трудноформализуемых событий с непрерывным обучением;
- 5) оценивание полученных результатов, выработка рекомендаций по повышению точности и использованию разработанных моделей, методов и программных систем.

При разработке методов и моделей нейросетевого прогнозирования предлагается применить класс рекуррентных нейронных сетей с управляемыми элементами, что позволит использовать их возможности по наделению слоев различными логическими структурами и задействовать их потенциал по управлению ассоциативным вызовом информации из их памяти.

Разработка новых моделей, методов и архитектур программных систем нейросетевого прогнозирования с непрерывным обучением позволит расширить возможности по прогнозированию временных рядов со сложной динамикой, повысит интеллектуальность использующих их систем и повысит эффективность этих систем.

Выводы по главе 1

В настоящей главе проведён анализ известных методов и систем прогнозирования событий, по результатам которого было установлено, что они не

удовлетворяют условиям, которые имеют место при функционировании временных рядов со сложной динамикой. Показано, что существующие подходы не учитывают ряд особенностей, свойственных таким временным рядам. Нет возможности оперативного учета изменений законов поведения временных рядов. Традиционные подходы к прогнозированию позволяют получать достаточно точные прогнозы лишь на небольших горизонтах, а более перспективные методы машинного обучения являются либо вычислительно сложными, либо недостаточно надежными в силу не проработанности используемых архитектур нейронных сетей. Остаются нерешенными проблемы управления ассоциативным вызовом информации из памяти таких сетей, не развиты правила этого управления. Отсутствуют подходы, позволяющие осуществлять прогнозирование событий с непрерывным обучением, что существенно ограничивает применимость известных решений в условиях высокой неопределенности.

Сформулирована научная задача диссертационного исследования: разработка моделей, методов и архитектур программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.

ГЛАВА 2. МОДЕЛИ И МЕТОДЫ НЕЙРОСЕТЕВОГО ПРОГНОЗИРОВАНИЯ ТРУДНОФОРМАЛИЗУЕМЫХ СОБЫТИЙ С НЕПРЕРЫВНЫМ ОБУЧЕНИЕМ

2.1 Обобщенная модель нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением

С формальной точки зрения прогнозирование событий по временному ряду $X = \{x_1, x_2, \dots, x_T\}$, на горизонт $T_{\text{упр.}}$ сводится к поиску ряда $X_{\text{п.}} = \{x_{T+1}, x_{T+2}, \dots, x_{T+\tau}, \dots, x_{T+T_{\text{упр.}}}\}$, который описывает будущие события, где $\tau = \overline{1, T_{\text{упр.}}}$. Система прогнозирования реализует функцию S от исходного ряда X , позволяющую получить временной ряд $X_{\text{п.}}$, отличающийся от реального $X_{\text{п.}}'$ на величину менее заданной ошибки $E_{\text{зад.}}$:

$$X_{\text{п.}} = S(X), \quad (1)$$

$$\text{Error}(X_{\text{п.}}, X_{\text{п.}}') < E_{\text{зад.}} \quad (2)$$

Могут прогнозироваться различные события. Например, для беспилотных транспортных средств (БТС) это могут быть ряды в виде значений скоростей движения на дорогах в фиксированные моменты времени, объемов дорожного трафика или загруженности дорог [141]. Эти временные ряды с течением времени дополняются новыми значениями, информация обновляется, и выдаваемый прогноз нуждается в корректировке с учетом вновь полученных значений. При этом могут изменяться законы поведения рядов, которые необходимо учитывать для получения точных прогнозов событий. Это требует постоянного обучения систем прогнозирования и отсутствия прерываний этого обучения при получении прогнозов. Известные решения не удовлетворяют этим требованиям, что существенно сказывается на точности формируемых прогнозов. Предлагается новая нейросетевая модель прогнозирования, позволяющая устранить существующее противоречие и повысить точность прогнозирования различных событий [140]. Эта модель нейросетевого прогнозирования событий с непрерывным обучением представлена на рисунке 6.

Согласно рис. 6 в состав системы прогнозирования с непрерывным обучением входят:

- 1) блок предобработки;
- 2) первый экземпляр нейронной сети (РНС-1);
- 3) блок постобработки № 1;
- 4) второй экземпляр нейронной сети (РНС-2);
- 5) блок постобработки № 2;
- 5) блок управления прогнозированием.

Блок предобработки предназначен для преобразования входных данных в последовательность совокупностей единичных образов (СЕО) – формат, воспринимаемый нейронной сетью, а также для определения эффективного способа кодирования данных.

Первый экземпляр нейронной сети, называемый РНС-1, предназначен для обработки данных и непрерывного обучения (то есть построения внутри себя пространственно-временной модели событий, происходящих во внешней среде и отраженных в значениях временного ряда) и распознавания этих данных.

Блок постобработки №1 предназначается для раскодирования результатов распознавания, получаемых с выхода РНС-1. Процедуры раскодирования является обратной используемой в блоке предобработки.

Второй экземпляр нейронной сети, называемый РНС-2, предназначен для получения с блока управления прогнозированием пространственно-временной модели РНС-1, приведения своего состояния в соответствие этой пространственно-временной модели путем присваивания соответствующим элементам полученных значений, обработки скопированных данных внутри себя с формированием прогноза на указанный горизонт и выдачей результатов прогнозирования на выходе.

Блок постобработки №2 по своей структуре и выполняемым операциям аналогичен блоку постобработки №1, за тем исключением, предназначен для раскодирования результатов прогнозов.

Блок управления прогнозированием предназначен для осуществления общего контроля над системой прогнозирования, и в частности, над экземплярами нейронных сетей РНС-1 и РНС-2.

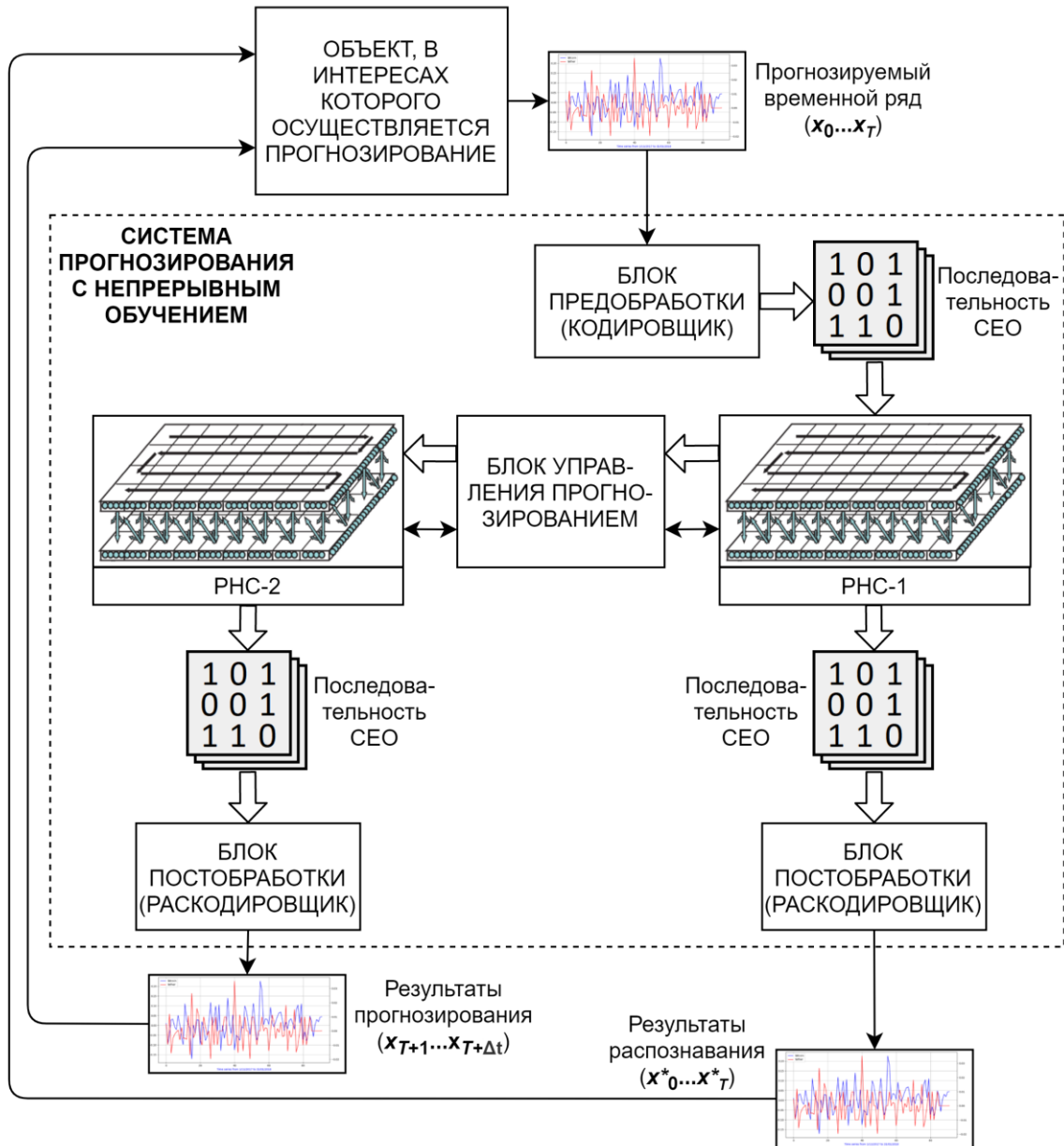


Рисунок 6 — Обобщенная модель системы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением. Объемными стрелками обозначены потоки обрабатываемых данных

Главная идея, положенная в основу предлагаемой модели системы прогнозирования, заключается в использовании двух экземпляров РНС, первый из которых непрерывно работает в режиме обучения, а во втором по необходимости

выполняется прогнозирование на основе обученной пространственно-временной модели, скопированной из первой РНС. Такая схема позволяет, во-первых, производить обучение и прогнозирование в реальном времени, во-вторых, быстро без переобучения выполнять обновление прогнозов, и в-третьих, не искажать обученную пространственно-временную модель из-за смены режимов работы РНС.

Выбор рекуррентных нейронных сетей в качестве основы системы прогнозирования объясняется тем, что такие архитектуры предназначены для обработки временных последовательностей произвольной длины, и это делает их наиболее пригодными для прогнозирования событий в условиях высокой неопределенности. Наличие обратных связей позволяет РНС осуществлять более глубокую пространственно-временную обработку информации и выполнять прогнозы на большие горизонты, чем обычным нейронным сетям прямого распространения. В качестве таких РНС предложено использовать рекуррентные нейронные сети с управляемыми элементами, наделяемые различными логическими структурами слоев [136, 142, 143]. Выбор таких РНС обосновывается их широкими функциональными возможностями по управляемой ассоциативной обработке сигналов.

Ниже приведена последовательность шагов, описывающая работу предлагаемой модели прогнозирования с непрерывным обучением.

Шаг 1. Элементы обрабатываемого временного ряда, на рис. 6 обозначенного $x_0 \dots x_T$, поступают в блок предобработки. Там они преобразуются в формат, воспринимаемый нейронной сетью, и поступают в РНС-1.

Шаг 2. Последовательность закодированных данных обрабатывается в РНС-1. На элементах РНС-1 происходит формирование пространственно-временной модели. При этом осуществляется распознавание и частичное восстановление элементов выборки.

Шаг 3. Блок управления прогнозированием выполняет копирование состояний элементов РНС-1 в РНС-2 и запускает на ней обработку

скопированных данных. В РНС-2 реализуется ассоциативный вызов прогнозных значений временного ряда.

Шаг 4. Осуществляется снятие информации с выхода РНС-1 и раскодирование данных в блоке постобработки №1. Получение результатов распознавания входных данных нейронной сетью ($x^*_0 \dots x^*_T$). Снятие информации с выхода РНС-2 и раскодирование их в блоке постобработки №2. Получение результатов прогнозирования ($x_{T+1} \dots x_{T+\Delta t}$).

В рамках предложенной модели могут реализовываться различные способы прогнозирования трудноформализуемых событий с непрерывным обучением. Рассмотрим математические формулировки задач поиска таких способов.

2.2 Математические формулировки задач поиска целесообразных способов нейросетевого прогнозирования трудноформализуемых событий

С формальной точки зрения задача разработки методов и средств прогнозирования трудноформализуемых событий сводится к поиску такого способа прогнозирования S_0 , при котором достигается экстремум целевой функции $W_0(S_0)$ при ограничениях на вспомогательные показатели эффективности $B_{ij}(S_i)$:

$$W_0(S_0) = \text{extr}_{i \in Q} W_i(S_i), \quad (3)$$

$$B_{ij}(S_i) \in B_{j \text{ зад.}}, \quad j = \overline{1, J}, \quad (4)$$

$$i = \overline{1, I}, \quad (5)$$

где Q – область поиска допустимых способов прогнозирования;

I – число возможных способов прогнозирования;

J – число вспомогательных показателей эффективности прогнозирования;

$B_{j \text{ зад.}}$ – области допустимых значений.

В зависимости от целей и условий их достижения конкретные математические задачи могут существенно отличаться. Области применения и соответственно цели прогнозирования временных рядов очень широки и включают в себя приложения медицины, экономики, социальной сферы, техники.

Кроме этого прогнозирование может осуществляться на различных уровнях иерархии анализируемых процессов. В качестве показателей эффективности систем прогнозирования событий могут выступать различные ошибки и время прогнозирования, затрачиваемые аппаратные и другие ресурсы. Рассмотрим возможные математические формулировки задач поиска целесообразных способов нейросетевого прогнозирования трудноформализуемых событий.

Задача 1. Когда требуется найти целесообразный способ S_0 прогнозирования трудноформализуемых событий, позволяющий минимизировать его ошибку E_0 на заданном горизонте $T_{\text{упр. зад.}}$ при ограничениях на время формирования прогноза $T_{\text{рч. зад.}}$, необходимо решать задачу:

$$E_0(S_0) = \min_{i \in Q} E_i(S_i), \quad (6)$$

$$T_{\text{упр. } i}(S_i) = T_{\text{упр. зад.}} \quad (7)$$

$$T_{\text{рч. } i}(S_i) \leq T_{\text{рч. зад.}} \quad (8)$$

Ошибка прогнозирования $E_i(S_i)$ определяется выражением:

$$E_i(S_i) = \frac{1}{N} \sum_{k=1}^N Err_k(S_i), \quad (9)$$

где N – количество точек в прогнозе;

$Err_k(S_i)$ – оператор ошибки прогнозирования для k -й точки.

Ошибка прогнозирования есть апостериорная величина отклонения прогноза от действительного состояния объекта. Показатели, используемые для определения ошибки прогнозирования $Err_k(S_i)$, можно разделить на абсолютные и относительные [144].

К абсолютным показателям относятся те, которые позволяют количественно определить величину ошибки прогнозирования y_p^* для временного ряда y_p , содержащего $p = 1 \dots N$ точек, в единицах измерения прогнозируемого объекта. К этим показателям относятся:

1) средняя абсолютная ошибка (MAE) [145]:

$$MAE = \frac{1}{N} \sum_{p=1}^N |y_p - y_p^*|, \quad (10)$$

2) среднеквадратическая ошибка [145]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{p=1}^N (y_p - y_p^*)^2}. \quad (11)$$

Недостатком рассматриваемых показателей является то, что значение этих характеристик существенно зависит от масштаба измерения уровней исследуемых явлений.

Относительными показателями точности прогнозов являются такие, в которых ошибка выражается в процентах относительно фактических значений. К ним относится средняя абсолютная процентная ошибка (MAPE), выражаемая формулой [145]:

$$MAPE = \frac{1}{N} \sum_{p=1}^N \frac{|y_p - y_p^*|}{y_p} \cdot 100\%. \quad (12)$$

Данный показатель используется при сравнении точности прогнозов разнородных объектов прогнозирования, поскольку они характеризуют относительную точность прогноза.

Выбор показателей точности прогнозов зависит от решаемых задач. Необходимо учитывать, что на формирование исследуемого явления влияет множество разнообразных факторов. “Поэтому полное совпадение или значительное расхождение прогнозов и реальных данных может быть следствием просто особо благоприятных (или неблагоприятных) стечений обстоятельств. Отсюда следует, что о качестве прогнозов можно судить лишь по совокупности сопоставлений прогнозов и их реализаций” [145].

Задача 1 применима для приложений, в которых требуется принятие оперативных и точных решений. Примером таких приложений являются беспилотные транспортные средства [146]. Особенности их функционирования

таковы, что ошибки в прогнозах могут привести к авариям, повреждениям дорогостоящего оборудования и даже создать угрозу жизни и здоровью людей. Поэтому первоначально важно обеспечить высокую точность прогнозирования будущих событий. Время прогнозирования не должно превосходить заданных значений, в которых учитываются особенности движения БТС в общем транспортном потоке и время реакции на возможные события. В то же время, на борту БТС возможно размещение достаточно мощного бортового компьютера, поэтому роль аппаратных ограничений в данном случае несущественна.

Задача 2. В случае, когда нужно найти целесообразный способ S_0 прогнозирования трудноформализуемых событий при минимуме ошибки прогнозирования на заданный горизонт и ограничениях на объем используемой памяти $H_{\text{рсч. зад.}}$, решаемая задача имеет вид (13) – (15):

$$E_0(S_0) = \min_{i \in Q} E_i(S_i), \quad (13)$$

$$T_{\text{упр. } i}(S_i) = T_{\text{упр. зад.}} \quad (14)$$

$$H_{\text{рсч. } i}(S_i) \leq H_{\text{рсч. зад.}} \quad (15)$$

Задача 2 актуальна для систем, в которых наряду с точностью прогнозов имеет значение экономия аппаратных ресурсов, в частности, памяти. Яркими примерами в данном случае являются роверы-планетоходы [28, 29, 31, 32]. Для данного типа роботов, как и в предыдущем случае, требуются прогнозы высокой точности, поскольку ошибки в прогнозировании могут привести к нештатным ситуациям, устранить которые ввиду большой удаленности планетохода не представляется возможным. Помимо этого, жесткое космическое излучение вынуждает использовать специализированную электронику, защищенную от воздействия радиации, но на порядки менее производительную, чем «земная». Поэтому необходимо, чтобы искомый способ прогнозирования удовлетворял ограничениям на аппаратные ресурсы. В этом случае показатель времени, затрачиваемого на прогнозирование, играет второстепенную роль ввиду того, что скорость движения роверов в целом невысока.

Задача 3. Если необходимо найти целесообразный способ S_0 прогнозирования трудноформализуемых событий для заданного горизонта за минимальное время $T_{р\text{сч. } 0}$ при ограничении на величину ошибки $E_{\text{зад.}}$, то условия задачи с учетом (6) – (8) и (13) – (15) представляются в виде:

$$T_{р\text{сч. } 0}(S_0) = \min_{i \in Q} T_{р\text{сч. } i}(S_i), \quad (16)$$

$$T_{упр. } i(S_i) = T_{упр. } \text{ зад.} \quad (17)$$

$$E_i(S_i) \leq E_{\text{зад.}} \quad (18)$$

Задача 3 применима для тех случаев, когда определяющим показателем является скорость принятия решений. Примером этого класса являются робототехнические системы военного назначения [147–149]. В условиях боя быстрота и маневренность являются ключевыми факторами выживания робота, поэтому именно им следует уделять первостепенное внимание. В то же время, ошибка прогнозирования не должна превышать установленных значений, поскольку неправильное планирование поведения может привести к уничтожению робота.

Задача 4. Когда требуется найти целесообразный способ S_0 прогнозирования трудноформализуемых событий при минимизации объема требуемой памяти $H_{р\text{сч. } 0}$ системы и ограничениях на ошибку прогнозирования с заданным горизонтом, требуется решать задачу:

$$H_{р\text{сч. } 0}(S_0) = \min_{i \in Q} H_{р\text{сч. } i}(S_i), \quad (19)$$

$$T_{упр. } i(S_i) = T_{упр. } \text{ зад.} \quad (20)$$

$$E_i(S_i) \leq E_{\text{зад.}} \quad (21)$$

Обозначения, принятые в (19) – (21), раскрыты при описании (6) – (8) и (13) – (18).

Задача 4 наилучшим образом отражает потребности устройств, примером которых являются компактные бытовые автономные средства (роботы-подметальщики, пылесосы, мойщики окон и т.п.). Такие средства эксплуатируются в достаточно мягких условиях, не требующих большой оперативности прогнозирования. Для таких средств основным показателем

эффективности является стоимость, которая, в свою очередь, напрямую зависит от количества аппаратных ресурсов бортового компьютера. Поэтому фактор их затрат системой прогнозирования является определяющим.

Из рассмотрения оптимизационной задачи (3) – (5) и ее частных случаев (задачи 1-4), следует, что для разных условий требуется применение своих моделей и методов прогнозирования трудноформализуемых событий, а также практических рекомендаций по их применению.

2.3 Метод нейросетевого прогнозирования с непрерывным обучением системы и с временными сдвигами сигналов

Модель нейросетевого прогнозирования событий с непрерывным обучением (см. 2.1) может реализовываться, исходя из различных задач и условий (см. 2.2). Ее основная идея, заключается в наличии двух параллельно работающих экземпляров РНС, первый из которых непрерывно обучается, а второй по необходимости выполняет прогнозы, основываясь на сформированной в первой РНС пространственно-временной модели наблюдаемого процесса. Исходя из требований к системе прогнозирования с непрерывным обучением оно может реализовываться различными методами.

Рассмотрим метод нейросетевого прогнозирования с непрерывным обучением с временными сдвигами сигналов [140]. В основу этого метода положены рекуррентные нейронные сети с управляемыми элементами [136, 142, 143] (рис. 7, 8). Такие нейронные сети представляют собой структуры, состоящие из двух слоев нейронов, двух блоков единичных задержек (ЕЗ), двух блоков динамических синапсов (ДС) и блока управления (рис. 7). Нейроны могут находиться в трех состояниях: ожидания, возбуждения и рефрактерности (невосприимчивости). Изначально все нейроны находятся в состоянии ожидания. Когда суммарный заряд на входе нейрона превышает некоторый порог возбуждения, нейрон возбуждается и генерирует одиночный импульс (сигнал), после чего переходит в состояние невосприимчивости входных сигналов. В этом состоянии нейрон не может накапливать заряд и возбуждаться. Время

рефрактерности нейронов больше времени суммарной задержки одиночных импульсов в двухслойном контуре РНС [150].

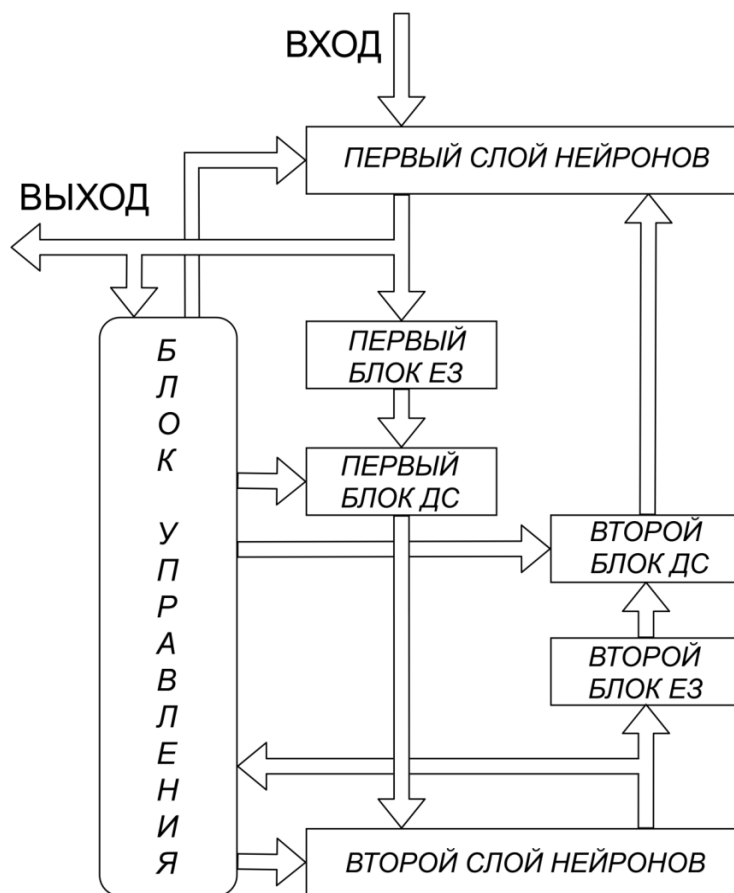


Рисунок 7 — Структурная схема рекуррентной нейронной сети с управляемыми элементами

Нейронная сеть оперирует с данными, представленными в виде совокупностей единичных импульсов возбужденных нейронов, называемых совокупностями единичных образов (СЕО). Сгенерированные нейронами импульсы задерживаются в блоке единичных задержек и через блок динамических синапсов (ДС) передаются на нейроны другого слоя. Блок ДС связывает синапсами каждый нейрон одного слоя со всеми нейронами другого слоя. Амплитуда единичного образа на выходе каждого синапса, связывающего i -й нейрон с j -м нейроном на момент времени t , равна амплитуде входного единичного образа, умноженной на вес $w_{ij}(t)$. Веса определяются по формуле [151, 152]:

$$w_{ij}(t) = k_{ij}(t) \times \beta(r_{ij}(t)) \times \eta(r_{ij}(t)), \quad (22)$$

где $k_{ij}(t)$ – весовой коэффициент;

$\beta(r_{ij}(t))$ – функция ослабления расходящихся единичных образов;

$\eta(r_{ij}(t))$ – функция ослабления сходящихся единичных образов.

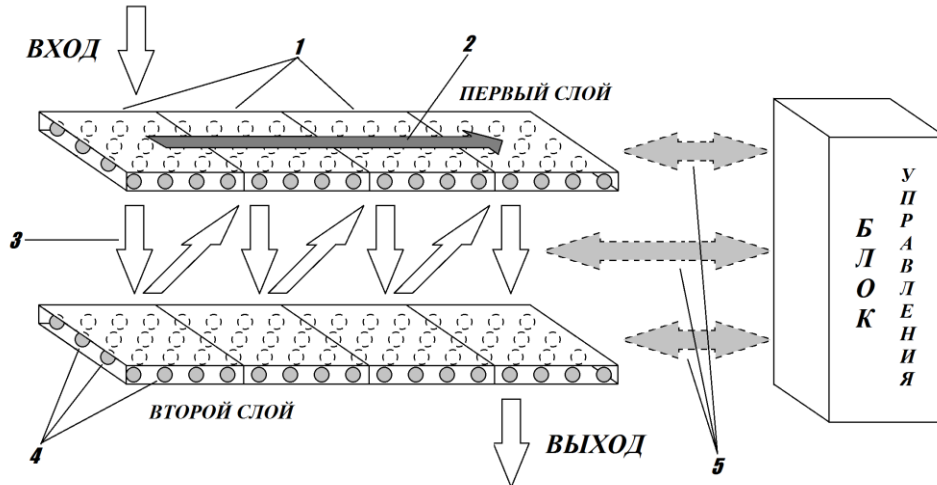


Рисунок 8 — Рекуррентная нейронная сеть с линейной структурой слоев: 1 – логические поля, на которые делятся слои сети; 2, 3 – направления движения СЕО вдоль слоев и между слоями соответственно; 4 – нейроны, 5 – управляющие воздействия блока управления [13]

Весовые коэффициенты $k_{ij}(t)$ изменяются в зависимости от воздействия на соответствующие синапсы проходящих через них импульсов и являются элементами долговременной ассоциативной памяти. При прохождении по синапсу импульс снимает с него информацию о предыдущих воздействиях и оставляет информацию о своем воздействии посредством изменения весового коэффициента [143, 150]. В зависимости от состояний i -го и j -го нейронов весовой коэффициент изменяется по следующим правилам [151]:

$$k_{ij}(t) = th(\gamma * g_{ij}(t)), \quad (23)$$

$$g_{ij}(t) = g_{ij}(t - \Delta t) + \Delta g_{ij}(t), \quad (24)$$

$$\Delta g_{ij}(t) = \begin{cases} g^+, & \text{если } i - \text{й нейрон сгенерировал сигнал, а } j - \text{й возбудился,} \\ g^-, & \text{если } i - \text{й нейрон не генерировал сигнал, а } j - \text{й возбудился,} \\ 0, & \text{во всех остальных случаях.} \end{cases} \quad (25)$$

где $g_{ij}(t)$, $g_{ij}(t - \Delta t)$ – предыстория обучения на моменты времени t и Δt соответственно, $g_{ij}(0) = 0$;

γ – положительный коэффициент;

g^- , g^+ – некоторые отрицательное и положительное значения соответственно.

Функции $\beta(r_{ij}(t))$ ослабления расходящихся единичных образов зависят от условной удаленности $r_{ij}(t)$ связываемых синапсом нейронов [151]:

$$\beta(r_{ij}(t)) = \frac{1}{1 + \alpha^h \sqrt{r_{ij}(t)}}, \quad (26)$$

где α , h – положительные коэффициенты.

При передаче одиночных импульсов между слоями блок управления определяет состояние принимающего слоя. Заметим, что используемые рекуррентные сети за счет реализации пространственных сдвигов сигналов при передаче от одного слоя к другому наделяются логическими структурами. Сдвиги осуществляются путем изменения соответствующих условных расстояний $r_{ij}(t)$. Благодаря сдвигам слои РНС разбиваются на логические поля одинакового размера. Если логические поля имеют размер $d \times q$, то соответствующие размеры слоев L и M определяются из выражений [140]:

$$L = d \times D, M = q \times B, \quad (27)$$

где D, B – положительные целые числа.

За счет сдвигов сигналов при их передаче от слоя к слою слои РНС могут конфигурироваться в различные логические структуры: линейные, спиральные, петлевые. На рисунке 8 представлена РНС со линейной структурой слоев.

С учетом вышеизложенного, параметр $r_{ij}(t)$ из выражений (22), (26) определяется следующим образом [136]:

$$r_{ij}(t) = \sqrt{(\Delta x_{ij}(t) + n_{ij}(t)d)^2 + (\Delta y_{ij}(t) + m_{ij}(t)q)^2}, \quad (28)$$

$$n_{ij}(t) = \pm 0, 1, \dots, D - 1; m_{ij}(t) = \pm 0, 1, \dots, B - 1, \quad (29)$$

где $\Delta x_{ij}(t)$, $\Delta y_{ij}(t)$ - проекции связи j -го нейрона с i -м нейроном на оси X и Y без учета пространственных сдвигов.

Выбор РНС с управляемыми элементами в качестве модельной основы предлагаемого метода объясняется следующими причинами. В отличие от известных методов, сети данной архитектуры наделены прозрачной логической структурой, что расширяет их возможности по управлению ассоциативным вызовом информации. Их слои могут быть сконфигурированы в различные структуры, при этом данные будут продвигаться от входа к выходу по различным маршрутам. Структура РНС с управляемыми элементами допускает параллельную обработку нескольких нейросетевых каналов. Именно это свойство будет использовано в предлагаемом методе нейросетевого прогнозирования с временными сдвигами.

Рассмотрим более подробно предлагаемый метод [140]. Его выполнение описывается следующим алгоритмом (рис. 9).

Шаг 1. Ввод данных в систему прогнозирования, разбиение их на два канала: первый (текущий) и второй (задержанный на заданное время). Кодирование данных в последовательность СЕО.

Шаг 2. Ассоциативная обработка данных в РНС-1. Формирование в сети пространственно-временной модели наблюдаемых событий.

Шаг 3. Считывание информации о состоянии РНС-1 в блок управления прогнозированием. Копирование состояний синапсов РНС-1 в РНС-2, а также первого (текущего) канала РНС-1 во второй (задержанный) канал РНС-2

Шаг 4. Запуск и реализация процесса обработки данных в РНС-2 с усилением ассоциативного вызова в первом канале РНС-2.

Шаг 5. Обратное преобразование результатов прогнозов в РНС-2 из формата СЕО в соответствующие им данные исходного формата.

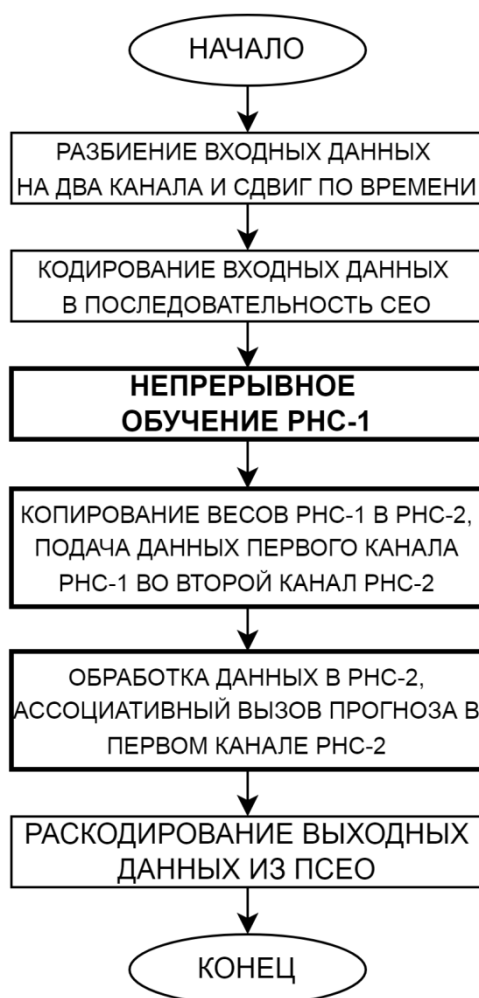


Рисунок 9 — Блок-схема, описывающая выполнение метода прогнозирования с непрерывным обучением системы с временными сдвигами сигналов

Структурная схема системы прогнозирования, реализующей указанный метод, представлена на рисунке 10, а сущность процесса формирования пространственно-временной модели и ассоциативного вызова данных на рисунке 11. На нем представлена РНС с линейной конфигурацией слоев, которая обрабатывает временной ряд, обозначенный как $F(t)$. Слои разделены на 6 логических полей, 3 из которых несут элементы обрабатываемой информации. Логические поля условно разделены на верхний нейросетевой канал (элементы данных $F(T-2)$, $F(T-1)$, $F(T)$) и нижний нейросетевой канал, задержанный относительно верхнего на время ΔT (элементы $F(T-\Delta T-2)$, $F(T-\Delta T-1)$, $F(T-\Delta T)$).

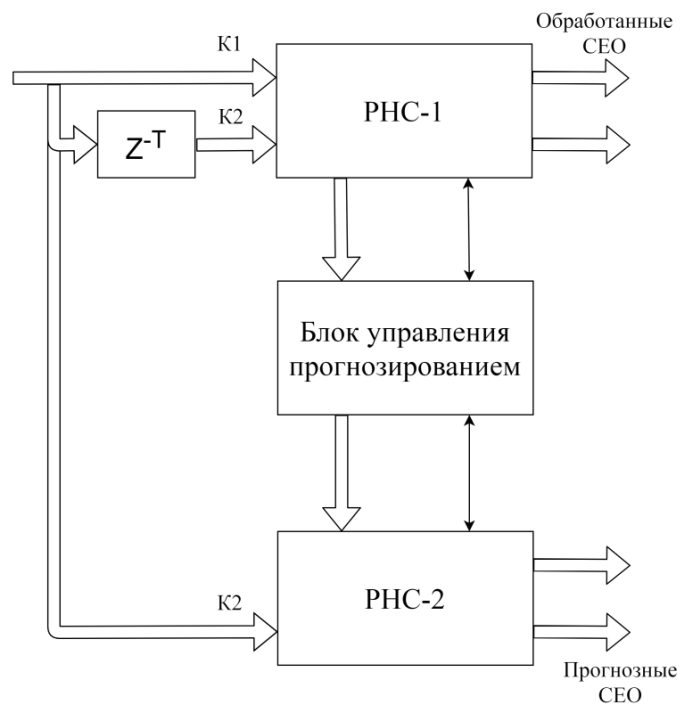


Рисунок 10 — Структурная схема системы нейросетевого прогнозирования с временными сдвигами

В верхней части рисунка 11 продемонстрировано формирование ассоциативных связей между элементами обрабатываемых данных. Связи обозначены пунктирными стрелками. В нижней части рисунка продемонстрирован процесс прогнозирования. За счет возбуждающего потенциала нейронов, несущих информацию об элементах $F(T-2)$, $F(T-1)$, $F(T)$, а также выстроенных ранее ассоциативных связей, обозначенных тут сплошными стрелками, происходит вызов из памяти сети прогнозных значений $F(T+\Delta T-2)$, $F(T+\Delta T-1)$, $F(T+\Delta T)$. На рисунке они оттенены.

На уровне нейросетевых каналов метод прогнозирования с временными сдвигами сигналов демонстрируется рисунком 12. В данном случае используется РНС с конфигурацией слоев в виде спирали постоянного радиуса. Серые стрелки представляют нейросетевые каналы. Белые стрелки обозначают поступление входных и выходных данных РНС, а черные – ассоциативные связи, формируемые в процессе обучения и используемые при прогнозировании.

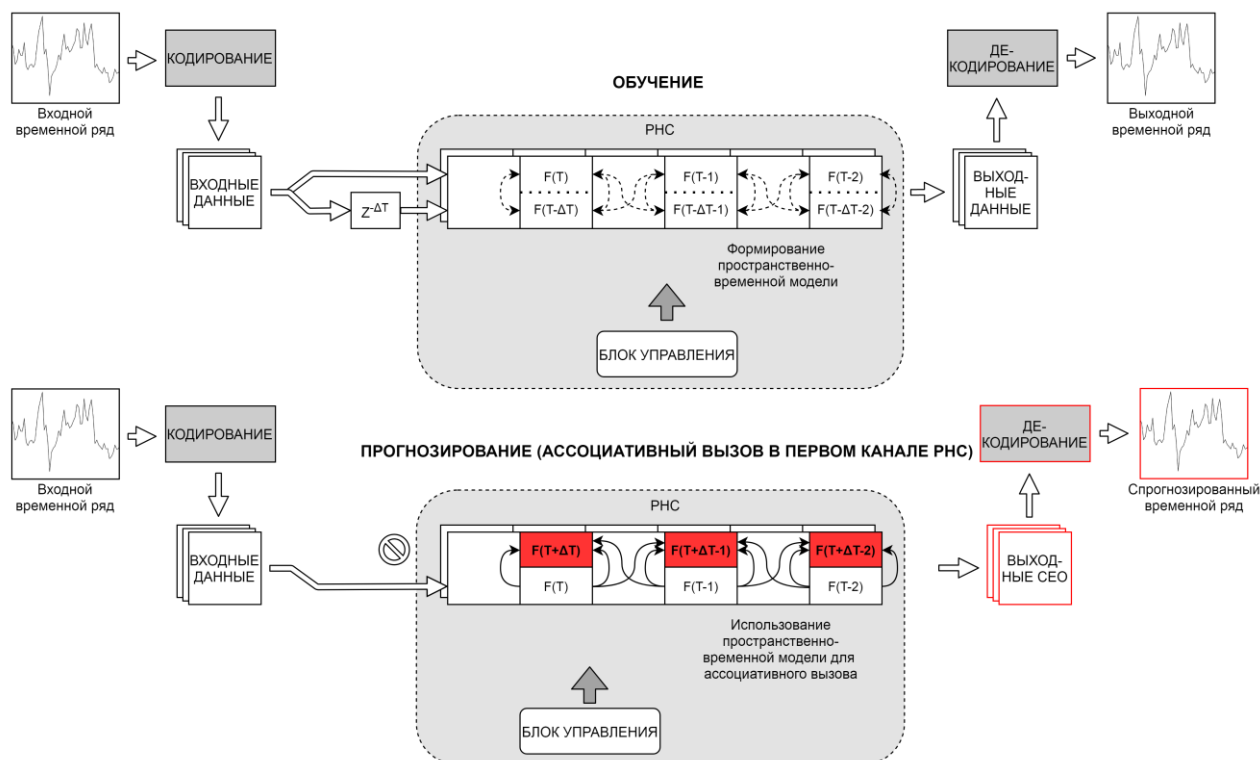


Рисунок 11 — Сущность процесса формирования пространственно-временной модели в РНС с управляемыми элементами (вверху) и последующего ассоциативного вызова информации из памяти сети (внизу)

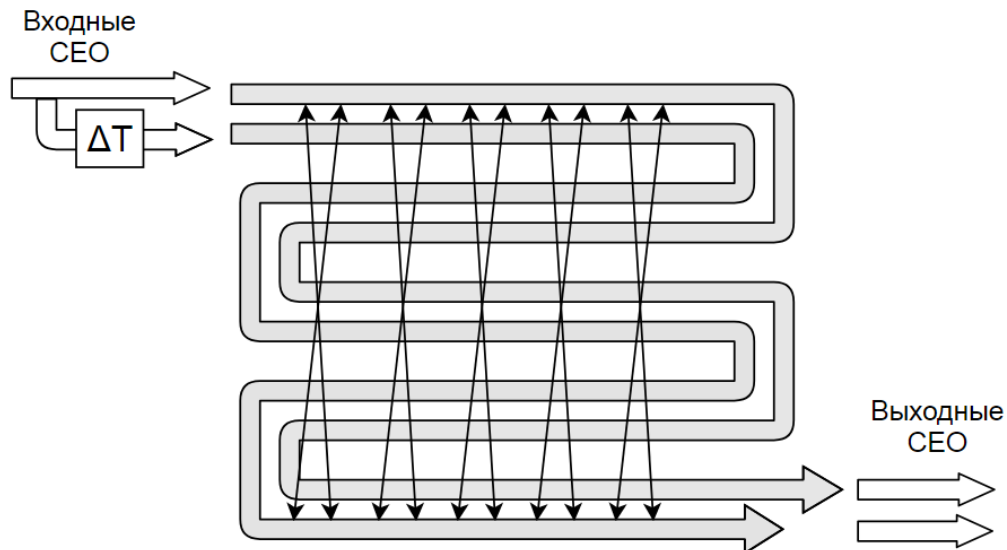


Рисунок 12 — Канальная структура РНС-1 при обучении с временными сдвигами сигналов

С формальной точки зрения, согласно методу прогнозирования с временными сдвигами временной ряд $X_t = \{x_1, x_2, \dots, x_t\}$ сдвигается на интервал τ и формируется ряд $X_{t-\tau} = \{x_{1-\tau}, x_{2-\tau}, \dots, x_{t-\tau}\}$. В РНС на вход подается пара X_t ,

$X_{t-\tau}$. В процессе прохождения пары временных рядов РНС формирует пространственно-временную ассоциативную модель путем изменения весов синапсов.

Процесс обучения и прогнозирования по методу с временными сдвигами представляет собой выражения:

$$W_{t+1} = \varphi(X_t, X_{t-\tau}, W_t), \quad (30)$$

$$X_t = \psi(X_{t-\tau}, W_t), \quad (31)$$

↓

Копирование весов синапсов РНС-1 в РНС-2

↓

$$X_{t+\tau} = \psi^*(X_t^*, W_t^*), \quad (32)$$

где φ – функция изменения весов синапсов; ψ – функция переходов процесса из одних состояний в другие, W_t – пространственно-временная модель событий.

2.4 Метод нейросетевого прогнозирования без временных сдвигов с управлением направленностью вызова сигналов из ассоциативной памяти

Метод с временными сдвигами сигналов, изложенный в предыдущем разделе, предполагает наличие двух нейросетевых каналов, сдвинутых по времени один относительно другого. Такая схема прогнозирования позволяет осуществлять плотное пространственно-временное связывание данных, но требует их дублирования и, как следствие, увеличения размеров логических полей РНС. Кроме того, горизонт прогнозирования при таком методе постоянен и равен величине сдвига. Такие ограничения не всегда позволяют эффективно использовать данный метод.

В качестве альтернативы предлагается метод прогнозирования без временных сдвигов последовательностей СЕО. Как и в предыдущем случае, используется пара РНС с управляемыми элементами. Однако, в отличие от метода с временными сдвигами, в данном случае используется только один нейросетевой канал. Пространственно-временное связывание осуществляется внутри этого канала, а прогнозирование основывается на ассоциативном вызове будущих

значений за счет предыдущих СЕО. Рассмотрим алгоритм выполнения этого метода [140, 151] (рис. 13).

Шаг 1. Ввод данных в систему прогнозирования и кодирование в последовательность СЕО.

Шаг 2. Ассоциативная обработка данных в РНС-1. Формирование в сети пространственно-временной модели наблюдаемых событий.

Шаг 3. Считывание информации о состоянии РНС-1 в блок управления прогнозированием. Копирование состояний элементов из РНС-1 в РНС-2. Определение параметров обрабатываемой выборки. Если данные в основном сосредоточены до середины слоев РНС, то есть обрабатываемая выборка является короткой, то осуществляется переход к шагу 5.

Шаг 4. Запуск и реализация процесса обработки данных в РНС-2 с ускоренным вызовом из ассоциативной памяти результатов прогноза в направлении входа этой сети. Переход к шагу 7.

Шаг 5. Запуск и реализация процесса обработки данных в РНС-2 с ускоренным вызовом из ассоциативной памяти результатов в направлении выхода этой сети в течение времени, зависящего от длины и положения выборки на слоях РНС.

Шаг 6. Изменение направленности ассоциативного вызова информации из памяти РНС-2 в сторону входа сети и продолжение реализации процесса обработки ПСЕО с формированием прогноза.

Шаг 7. Обратное преобразование результатов прогнозов в РНС-2 из формата СЕО в соответствующие им данные исходного формата.

Как видно из представленного алгоритма, метод прогнозирования без временных сдвигов позволяет не только получать прогнозы произвольной длины, но и наделен новыми правилами управления ассоциативным вызовом информации из памяти. Дополнение выборки предшествующими значениями при обработке коротких выборок позволяет полнее использовать обученную пространственно-временную модель, то есть, учитывать большой ассоциативный контекст при формировании прогнозов.

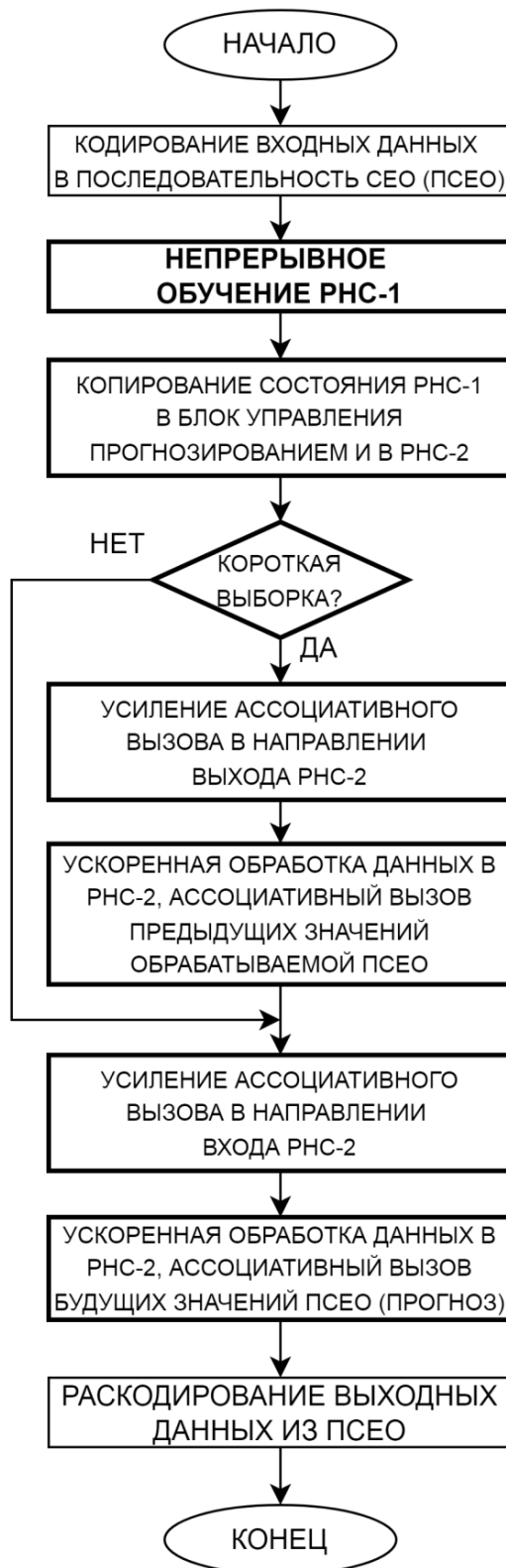


Рисунок 13 — Блок-схема, описывающая выполнение метода прогнозирования временных рядов без временных сдвигов с управлением направленностью вызова сигналов из ассоциативной памяти

Структурная схема системы прогнозирования, реализующей указанный метод, представлена на рис. 14. Сущность процесса формирования пространственно-временной модели и ассоциативного вызова данных представлена на рис. 15. Конфигурация РНС и используемые обозначения идентичны рисунку 11.

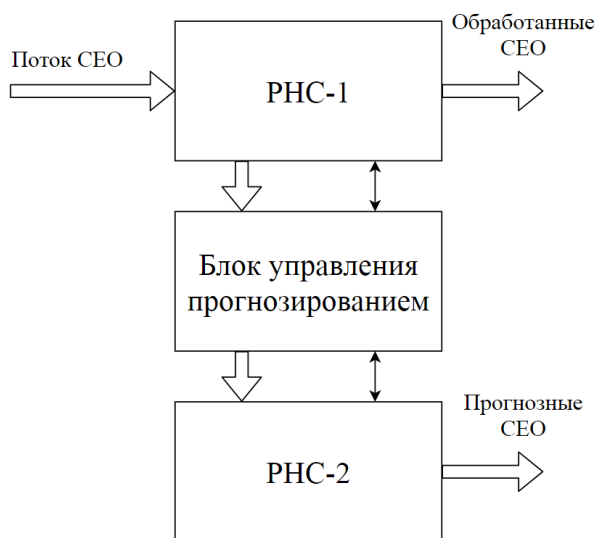


Рисунок 14 — Структурная схема системы нейросетевого прогнозирования без временных сдвигов

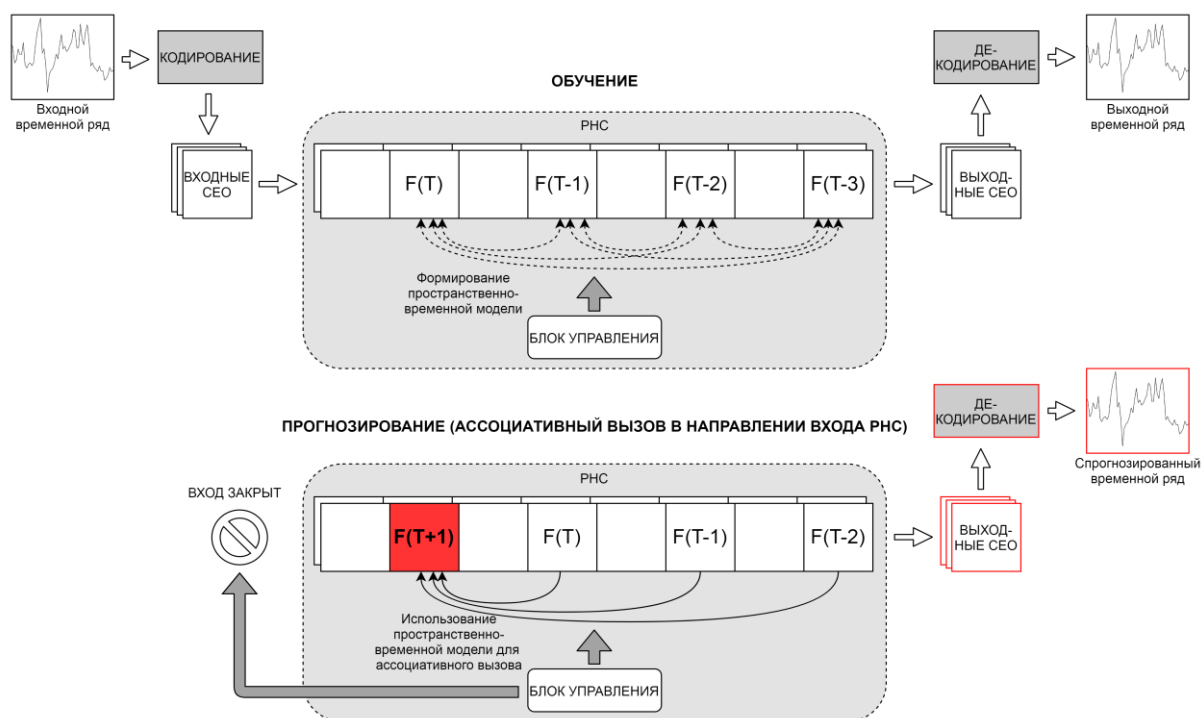


Рисунок 15 — Сущность процесса формирования пространственно-временной модели в РНС с управляемыми элементами (вверху) и последующего ассоциативного вызова информации из памяти сети (внизу)

В верхней части рисунка 15 продемонстрировано формирование ассоциативных связей между элементами обрабатываемых данных. Связи обозначены пунктирными стрелками. В нижней части рисунка продемонстрирован процесс прогнозирования. Вход РНС в режиме прогнозирования закрыт, а используемые ассоциативные связи усиливаются в направлении входа РНС для вызова прогнозных значений. Такое состояние системы прогнозирования соответствует шагам 4 и 6 алгоритма.

На уровне нейросетевых каналов метод прогнозирования с временными сдвигами демонстрируется рисунком 16. Обозначения и конфигурация слоев РНС идентичны рисунку 12. В отличие от схемы на рисунке 12, в этой схеме используется только один канал.

С формальной точки зрения описание процесса обучения и прогнозирования нейросетевой системой без временных сдвигов выглядит следующим образом:

$$\mathbf{W}_{t+1} = \varphi(\mathbf{X}_t, \mathbf{W}_t), \quad (33)$$

$$\mathbf{X}_t = \psi(\mathbf{X}_{t-1}, \mathbf{W}_t), \quad (34)$$

↓

Копирование весов синапсов РНС-1 в РНС-2

↓

$$\mathbf{X}_{t+\tau} = \psi^*(\mathbf{X}_t^*, \mathbf{W}_t^*). \quad (35)$$

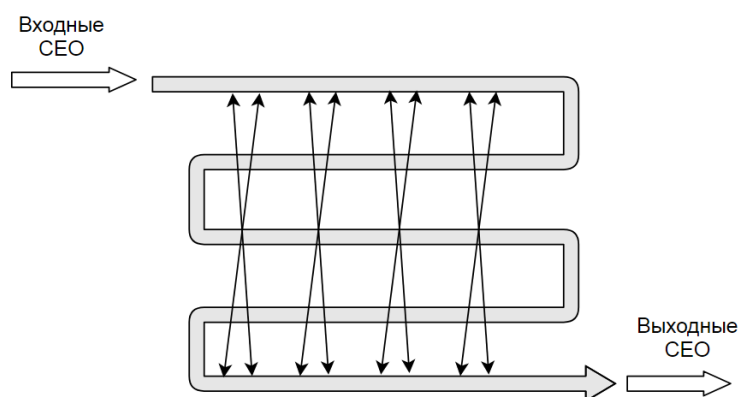


Рисунок 16 — Канальная структура РНС-1 при обучении в интересах прогнозирования без временных сдвигов

При обучении на вход подается ряд \mathbf{X}_t . В процессе прохождения его по слоям РНС-1 формируется пространственно-временная модель событий \mathbf{W}_t в

виде матрицы весов синапсов и текущих состояний нейронов. Когда необходимо произвести прогноз на горизонт $\tau = \overline{1, T_{\text{упр.}}}$, выполняется копирование пространственно-временной модели $W_t \rightarrow W_t^*$ и копирование состояний слоев $X_t \rightarrow X_t^*$ из РНС-1 в РНС-2, после чего на РНС-2 запускается ускоренная обработка данных с заданным коэффициентом ускорения. В результате на выходе РНС-2 получаем спрогнозированный временной ряд $X_{t+\tau}$.

Суть метода прогнозирования в условиях малых выборок, наличия шума и пропусков, являющегося составной частью метода прогнозирования без временных сдвигов и определяющего правила управления направленностью вызова сигналов из ассоциативной памяти нейронной сети, продемонстрирована на рисунке 17. Она заключается в оценке параметров обрабатываемой выборки $T_1 \dots T_N$ и, если та признается короткой, осуществляется удлинение ее за счет усиления ассоциативного вызова сигналов в направлении выхода РНС и получение таким путем предыдущих значений $T_{-M} \dots T_0$. Получение же прогноза $T_{N+1} \dots T_{N+K}$ предполагается за счет использования уже дополненного ряда $T_{-M} \dots T_N$ при усилении ассоциативного вызова сигналов в направлении входа РНС.

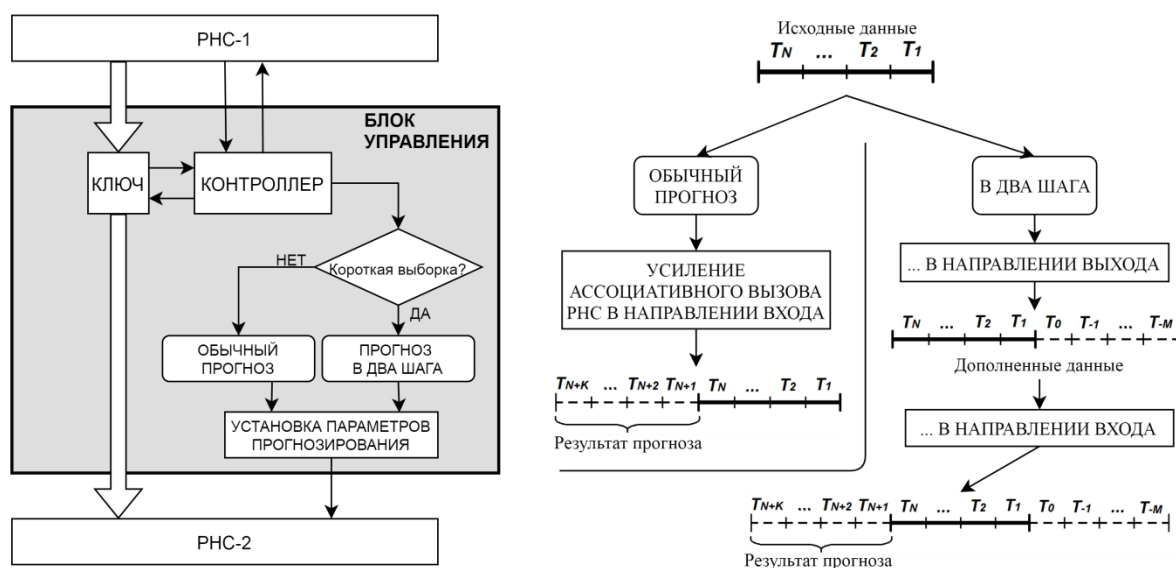


Рисунок 17 — Схема управления прогнозированием в условиях коротких выборок, шумов и пропусков

Таким образом, метод без временных сдвигов позволяет осуществлять прогнозирование событий с меньшими затратами памяти, поскольку отсутствует необходимость в дублировании данных и ресурсов, выделяемых на их обработку. Кроме того, для него предусмотрены правила управления ассоциативным вызовом информации из памяти РНС, что повышает качество использования обученной пространственно-временной модели.

Выводы по главе 2

Разработана обобщенная модель системы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, отличающаяся своей структурой и правилами обработки сигналов, обеспечивающими оперативное прогнозирование с учетом изменений в законах проявления событий. Она содержит блок управления прогнозированием и две идентичные по своей структуре рекуррентные нейронные сети (РНС-1 и РНС-2), объединенные в систему. РНС-1 работает в режиме обучения, а РНС-2 реализует прогнозирование.

Предложенная модель обеспечивает непрерывность процесса обучения при прогнозировании. Это позволяет, во-первых, обеспечить работу в реальном времени и возможность постоянного обновления прогнозов. Во-вторых, отсутствует необходимость переобучения сети при поступлении новых данных. В-третьих, исключается искажение пространственно-временной модели РНС из-за смены режимов функционирования сети.

Математически сформулированы задачи по поиску целесообразных способов прогнозирования.

В рамках разработанной модели предложено два новых метода нейросетевого прогнозирования с непрерывным обучением: с временными сдвигами сигналов и без временных сдвигов, отличающиеся новыми правилами прогнозирования и управления ассоциативным вызовом информации из нейросетевой памяти и обеспечивающие высокую точность получаемых прогнозов трудноформализуемых событий.

Согласно методу с временными сдвигами, в РНС-1 подается система из текущего и задержанного временных рядов, и в процессе их прохождения вдоль слоев сети осуществляется ассоциативное пространственно-временное связывание прошлых и будущих значений. По команде блок управления прогнозированием копирует обученную модель из РНС-1 в РНС-2 и подает текущие данные в задержанный канал, в результате чего в текущем канале происходит ассоциативный вызов будущих значений.

В методе прогнозирования без временных сдвигов на этапе обучения предлагается подавать только текущий временной ряд в РНС-1, в результате прохождения которого от ее входа к выходу формируется модель событий. По команде с блока управления прогнозированием предлагается осуществлять копирование состояния РНС-1 в РНС-2 и запускать на РНС-2 процесс прогнозирования по новым правилам управления направленностью ассоциативного вызова сигналов из памяти нейронной сети. Согласно этим правилам, если обрабатываемая выборка признается короткой, то перед прогнозированием предлагается удлинить ее за счет ассоциативного вызова из памяти сети предшествующих значений.

Метод без временных сдвигов позволяет использовать РНС с размером слоев в два раза меньшим, чем в методе с временными сдвигами, то есть является более экономичным в части затрат памяти и вычислений. В этом методе разработаны правила управления ассоциативным вызовом данных из нейросетевой памяти, повышающие качество использования пространственно-временной модели РНС.

ГЛАВА 3. АРХИТЕКТУРЫ ПРОГРАММНЫХ СИСТЕМ НЕЙРОСЕТЕВОГО ПРОГНОЗИРОВАНИЯ ТРУДНОФОРМАЛИЗУЕМЫХ СОБЫТИЙ С НЕПРЕРЫВНЫМ ОБУЧЕНИЕМ

3.1 Архитектура программной системы нейросетевого прогнозирования с непрерывным обучением

3.1.1 Требования, предъявляемые к программной архитектуре

Особенности программной архитектуры, реализующей методы прогнозирования временных рядов, во многом определяют их эффективность. Неудачная архитектура снижает эффект от использования методов, и наоборот, хорошая архитектура позволяет расширить возможности методов, положенных в ее основу. Первым шагом разработки архитектуры выступает формулирование требований к ней.

Для программной системы важно не только реализовывать в полной мере заявленный функционал, но и обеспечивать бесперебойное функционирование в течение длительного времени. Чем проще программная архитектура, тем выше ее надежность и отказоустойчивость, тем ниже вероятность возникновения ошибок. Надежность является одним из основных требований, предъявляемых к программной архитектуре. Особенно это касается многопоточных приложений, где риск возникновения трудно отлаживаемых ошибок особенно велик. В таких условиях важно обеспечить, с одной стороны, отсутствие лишних элементов в архитектуре. С другой стороны, должны быть предусмотрены средства мониторинга состояния программной системы, которые можно было бы задействовать при необходимости.

Программная реализация должна обеспечивать как автоматическое, так и ручное управление процессом ее функционирования. Несмотря на то, что внедренная система не требует участия оператора в своей работе, должны быть предусмотрены основные функции для ручного управления, такие как остановка, запуск, перезапуск, обновление прогноза. Кроме того, на этапе исследования

возможностей разрабатываемой системы и ее отладки необходимы средства настройки параметров и режимов работы программного обеспечения.

Как было отмечено в параграфе 2.1, для каждого типа приложений существуют свои задачи и ограничения, из-за чего необходимы различные методы прогнозирования в интересах таких приложений. Области применения прогнозирования временных рядов отличаются решаемыми ими задачами, средой функционирования системы прогнозирования, требованиями к точности и горизонту прогнозов, аппаратными возможностями компьютера, на который возложено выполнение прогнозов. Во второй главе предложены обобщенная модель и два метода прогнозирования с непрерывным обучением. Первый из них предусматривает использование двух нейросетевых каналов и временных сдвигов сигналов, а второй – одного канала и ускорения внутреннего времени РНС. При этом второй метод включает в себя правила управления направленностью вызова сигналов из ассоциативной нейросетевой памяти. Программная архитектура должна поддерживать возможность использования любого из методов. Независимо от конкретизации, необходимо поддержание двух экземпляров нейронных сетей идентичной конфигурации. Они должны функционировать в параллельных потоках с поддержанием асинхронного взаимодействия между блоком управления прогнозированием, РНС-1 и РНС-2.

Следующим ключевым требованием является возможность системы приводить входные данные, поступающие на обработку в программную систему прогнозирования, в формат нейронных сетей, и выполнять обратные операции на выходе. То есть архитектура должна поддерживать механизмы кодирования и декодирования данных нейросетевой обработки.

Согласно главе 2, функционирование РНС-1 и РНС-2 определяется их собственными внутренними блоками управления, а работу системы в целом контролирует блок управления прогнозированием. Все блоки управления имеют набор параметров, определяемых пользователем при инициализации системы или самим блоком в процессе работы системы. Особенности работы РНС заключается в том, что не всегда удастся формализовать подбор значений этих параметров,

поэтому для эффективного использования программной системы необходимы механизмы, позволяющие контролировать и задавать такие параметры.

Таким образом, требования к структуре разрабатываемой программной архитектуры включают следующие пункты:

- наличие средств мониторинга состояния программной системы;
- наличие автоматического и ручного управления программной системой;
- наличие двух экземпляров РНС и блока управления прогнозированием, работающих в отдельных потоках;
- возможность реализации методов нейросетевого прогнозирования с временными сдвигами сигналов и без временных сдвигов с управлением направленностью вызова сигналов из ассоциативной памяти;
- наличие механизмов приема входных данных в исходном формате и кодирования/раскодирования их в формат, воспринимаемый нейронной сетью;
- наличие средств для контроля и редактирования параметров, связанных с функционированием РНС-1 и РНС-2 и программной системы в целом.

3.1.2 Структура, состав и предназначение элементов архитектуры

Предлагаемая архитектура программной системы нейросетевого прогнозирования с непрерывным обучением [151, 153, 154] представлена на рисунке 18, а на рисунке 19 для нее приведена диаграмма классов UML.

Архитектура включает в себя следующие программные модули:

- 1) экземпляр нейронной сети РНС-1;
- 2) экземпляр нейронной сети РНС-2;
- 3) модуль предобработки входных данных;
- 4) модуль постобработки выходных данных №1;
- 5) модуль постобработки выходных данных №2;
- 6) модуль чтения/записи параметров №1;
- 7) модуль чтения/записи параметров №2;
- 8) графический интерфейс, состоящий из кнопочных панелей и графических модулей для каждого экземпляра РНС.

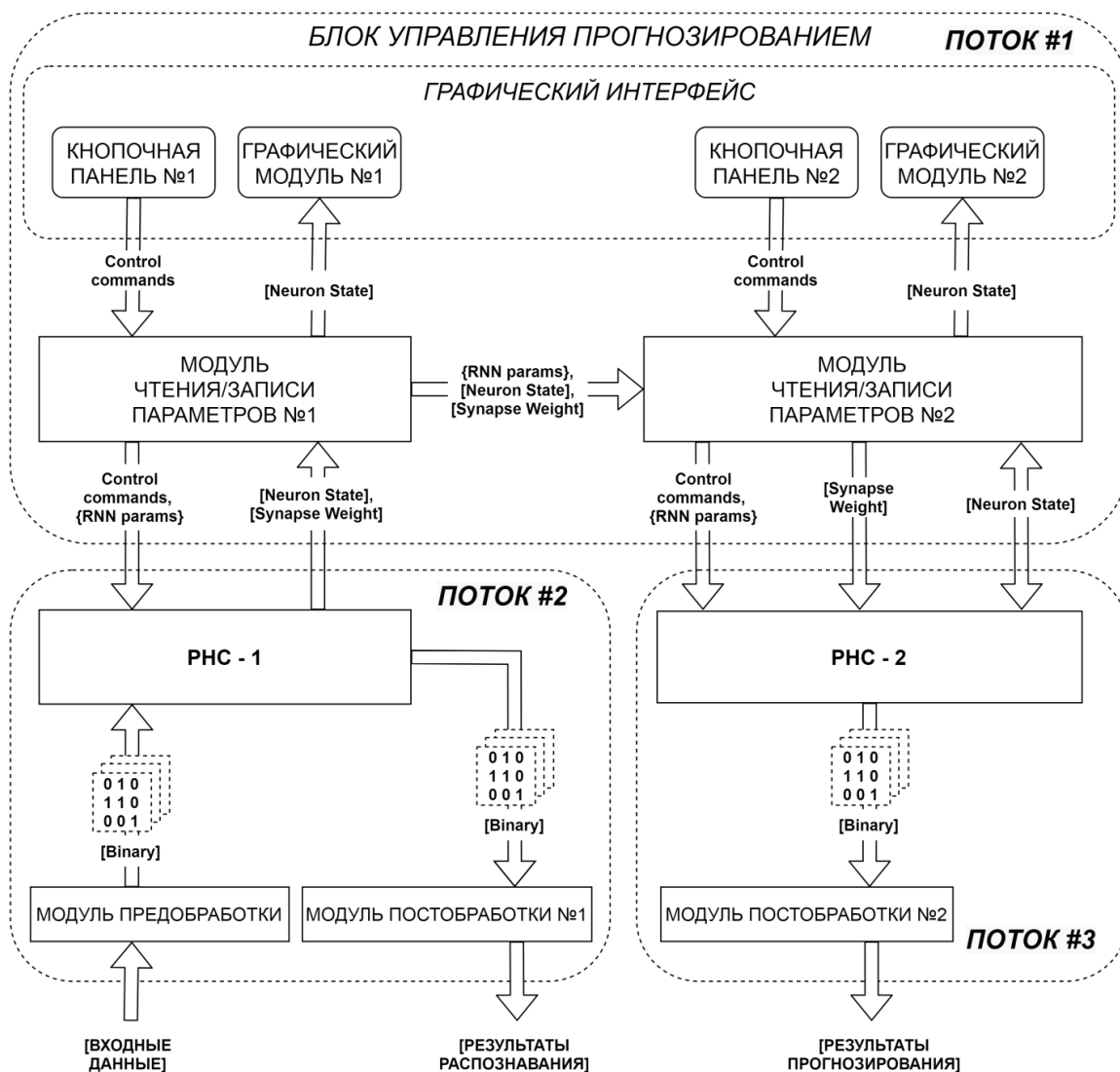


Рисунок 18 — Архитектура программной системы нейросетевого прогнозирования с непрерывным обучением. Здесь Control commands – управляющие команды, {RnnParams} – программная структура с информацией о параметрах функционирования РНС, [Binary], [Neuron State], [Synapse Weight] – массивы с бинарными данными CEO, состояниями нейронов и весами синапсов соответственно

Экземпляр нейросети РНС-1 является реализацией абстрактной нейросетевой программной структуры, описывающей функционал рекуррентных нейронных сетей с управляемыми элементами. Поскольку РНС-1 предназначена для обучения, в данном экземпляре предусматривается возможность изменения синаптических весов сети. Запуск экземпляра РНС-1 осуществляется в отдельном потоке (на рис. 18 «поток #2»). Обмен данными с другими программными

модулями системы прогнозирования осуществляется с применением одного из механизмов межпотокowego взаимодействия, таких как очередь. В процессе функционирования системы РНС-1 связан с другими модулями следующим образом. Из модуля предобработки он получает закодированные в формат СЕО данные о значениях временного ряда в виде массивов бинарных данных ([Binary]). Из модуля чтения-записи параметров №1 на него поступают управляющие команды (Control commands), а также передаются параметры функционирования РНС-1 в виде структуры {RnnParams}. В обратную сторону поступает информация о состоянии слоев РНС в виде массива состояний нейронов [Neuron State], если включен режим их отображения на пользовательском интерфейсе. При получении команды на прогноз туда же передается пространственно-временная модель РНС-1 в виде совокупности состояний ее нейронов ([Neuron State]) и весов синапсов ([Synapse Weight]). Выходные данные, являющиеся результатами распознавания на слоях РНС-1, передаются в модуль постобработки №1.

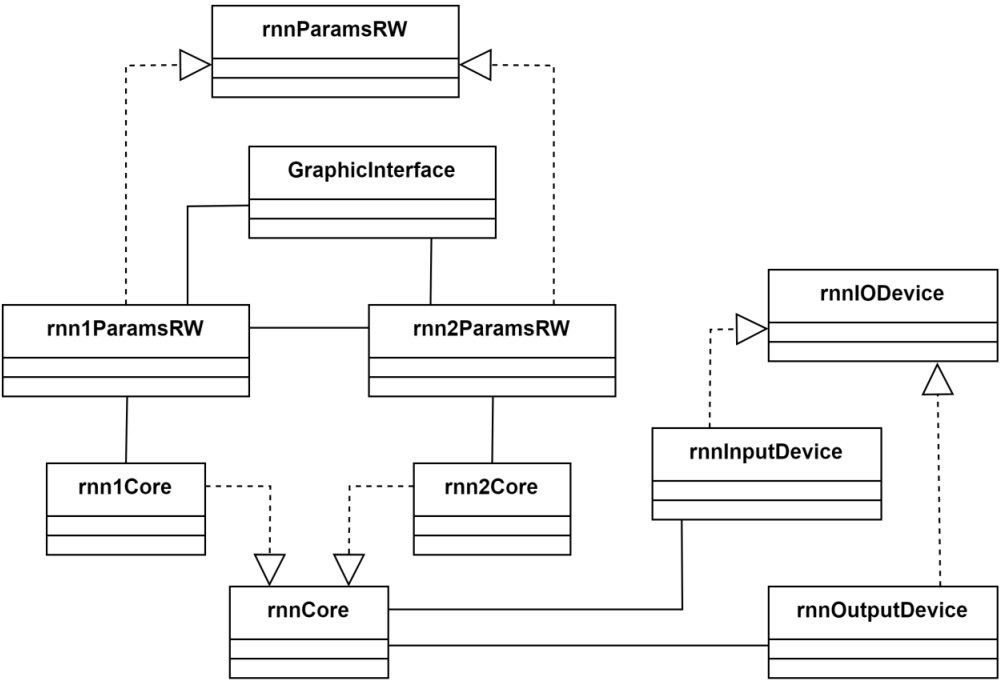


Рисунок 19 — Диаграмма классов UML для параллельной архитектуры программной системы нейросетевого прогнозирования с непрерывным обучением

Экземпляр нейросети РНС-2 – это реализация абстрактной нейросетевой программной структуры по аналогии с РНС-1. В отличие от РНС-1, этот экземпляр не наделяется функционалом, связанным с обучением. Как и РНС-1, выполнение РНС-2 вынесено в отдельный поток («поток #3»), а информационный обмен осуществляется с применением механизмов межпоточкового взаимодействия. По аналогии с РНС-1, экземпляр РНС-2 осуществляет информационный обмен с модулем чтения-записи параметров №2. На вход модуля РНС-2 передаются управляющие команды и скопированная пространственно-временная модель РНС-1 в виде массивов состояний нейронов и весов синапсов. Выходные данные РНС-2 в виде бинарных данных, кодирующих прогнозные значения, передаются в модуль постобработки №2. При включенном режиме визуализации состояний слоев РНС с этого экземпляра также передаются состояния нейронов.

На диаграмме классов UML (рис. 19) абстрактный класс нейросетевого ядра носит название `rnnCore`, а его реализации РНС-1 и РНС-2 – `rnn1Core` и `rnn2Core` соответственно.

Модуль предобработки входных данных реализует абстрактный класс кодировщика и ответственен за выполнение кодирования. Он получает на вход произвольные данные из внешней среды, а на выходе в соответствии с заданными правилами формирует закодированное представление в бинарном виде, которое передается на обработку в РНС-1.

Модули постобработки выходных данных №1 и №2 по аналогии с модулем предобработки являются реализациями абстрактного класса кодировщика, с тем отличием, что вместо кодирования данных в формат РНС выполняют процедуру декодирования. Модули постобработки получают на вход закодированные бинарные данные с выходов РНС-1 и РНС-2, а на выходе формируют элементы данных в исходном формате.

На рис. 19 абстрактный класс кодировщика обозначен `rnnIODevice`, а классы модулей предобработки и постобработки – `rnnInputDevice` и `rnnOutputDevice` соответственно.

Модуль чтения/записи параметров №1 реализует абстрактный класс, отвечающий за хранение, редактирование и передачу в РНС-1 параметров, связанных с управлением этим экземпляром нейронной сети и программной системы в целом. Все основные потоки данных в программной системе проходят через этот блок. С кнопочной панели графического интерфейса на модуль чтения/записи параметров №1 поступают управляющие команды, отправляемые с интерфейса. Информация о состояниях слоев с РНС-1 передается на графический модуль. Через модуль чтения/записи параметров №1 с РНС-1 на модуль чтения/записи параметров №2 передается скопированная пространственно-временная модель. Кроме того, модуль чтения-записи параметров №1 располагает собственным диалоговым окном, предназначенным для ручного редактирования его параметров.

Модуль чтения/записи параметров №2 является реализацией того же класса, что и модуль чтения/записи №1. Отличиями от первого модуля является отсутствие параметров, связанных с управлением системой в целом. Через модуль чтения-записи параметров №2 с модуля №1 на РНС-2 передается пространственно-временная модель РНС-1. По аналогии с модулем №1, этот модуль имеет вызываемое диалоговое окно для редактирования своих параметров.

На диаграмме классов абстрактный класс модуля чтения-записи параметров приведен под именем `rnnParamsRW`, а модули чтения-записи №1 и №2 – под именами `rnn1ParamsRW` и `rnn2ParamsRW` соответственно.

Графический интерфейс включает в себя программно реализованные кнопочные панели №1 и №2, предназначенные для отдачи пользовательских команд по управлению системой прогнозирования, и графические модули №1 и №2, на которые поступает с соответствующих модулей чтения/записи параметров информация о состояниях слоев нейронных сетей. Графические модули выполняют визуализацию состояний слоев РНС-1 и РНС-2 на главном окне программы. Графический интерфейс на диаграмме классов носит название `GraphicInterface`.

Модули чтения-записи параметров №1, №2 и графический интерфейс составляют блок управления прогнозированием, который работает в отдельном от РНС-1 и РНС-2 потоке (на рис. 18 – «поток #1»).

3.1.3 Правила функционирования программной архитектуры

Функционирование программной архитектуры системы прогнозирования представлено на блок-схеме (рис. 20). Оно может быть условно разделено на три этапа, каждый из которых имеет свои особенности. Рассмотрим подробно каждый из этапов.

Этап 1. Инициализация системы.

Запуск программы начинается с получения последних актуальных настроек модулями чтения/записи параметров №1 и №2. Если реализация программной системы предусматривает хранение настроек во внешнем файле или в другом источнике, то модули инициализируются с использованием этой информации. В противном случае используются настройки по умолчанию.

Исходя из полученной информации, определяются параметры нейронных сетей РНС-1 и РНС-2 (размеры слоев и логических полей, параметры синапсов, нейронов, конфигурация нейросетевых слоев и т.д.), модулей предобработки и постобработки (определяется источник данных для обработки: файл, tcp-сокеты и т.д., формат данных, способ кодирования), а также различные системные флаги и настройки. Создаются экземпляры нейросетей, организуются каналы межпоточного взаимодействия. Экземпляры РНС помещаются в отдельные потоки и запускаются.

Этап 2. Прогнозирование с непрерывным обучением.

Прогнозирование с непрерывным обучением является основной фазой работы программной системы. Данная фаза выполняется в многопоточном режиме. На рисунке 20 имеется 3 потока. Первый – основной, в котором работает блок управления прогнозированием. Во втором и третьем потоках функционируют РНС-1 и РНС-2 соответственно. Для передачи информации между потоками используется в общем случае 9 каналов, предназначение которых раскрыто в таблице 1.

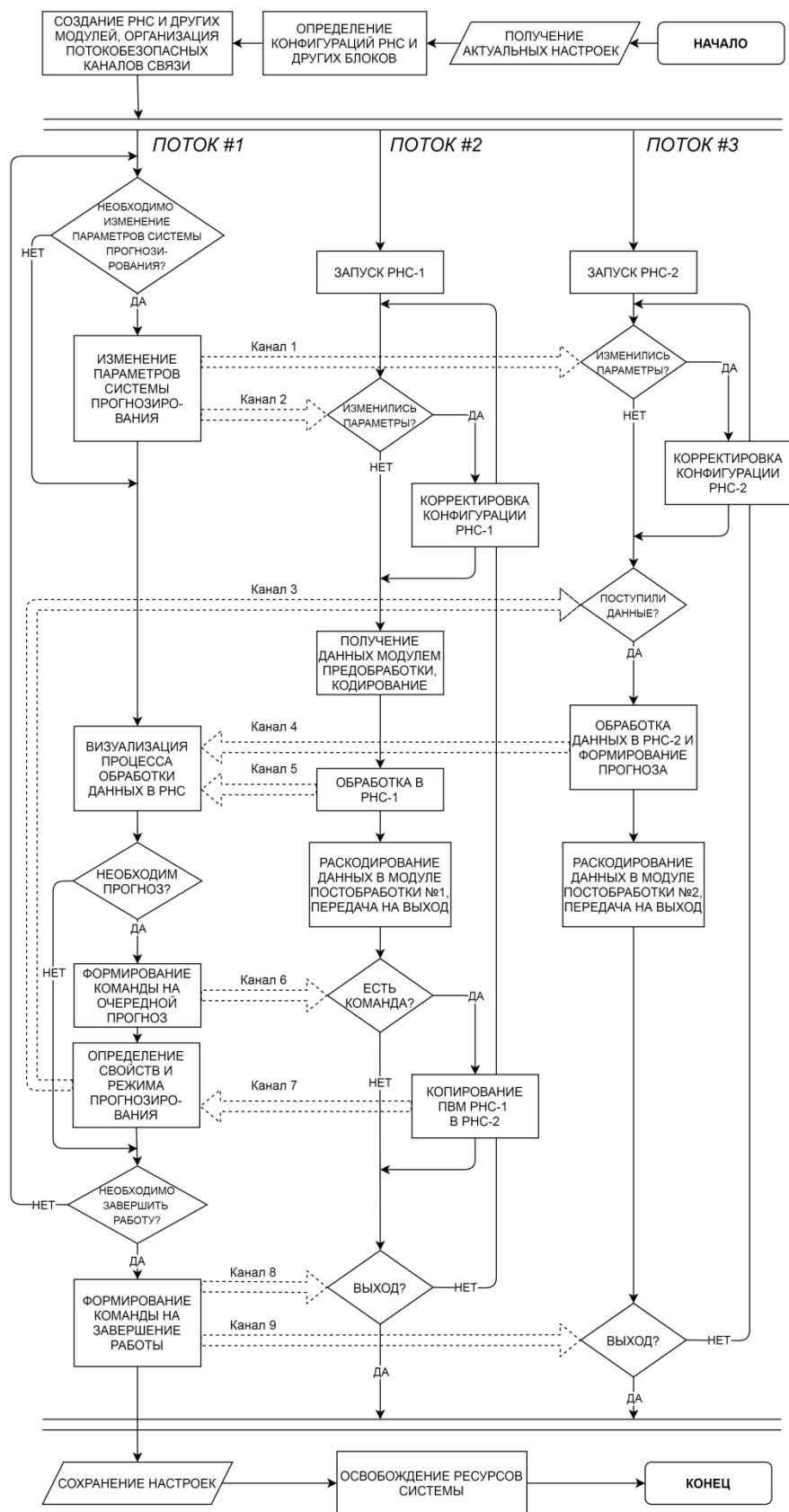


Рисунок 20 — Блок-схема, демонстрирующая процесс функционирования программной архитектуры. Каналы 1-9 (пунктирные стрелки) обозначают передачу данных между потоками

Выполнение данной фазы работы программной архитектуры включает в себя четыре основных подпроцесса:

2.1. *Чтение, изменение и задание параметров системы.* После того, как программная система инициализирована, параметры ее функционирования могут быть изменены без остановки работы. Для этого с кнопочной панели графического интерфейса вызывается команда на редактирование, после чего модуль чтения-записи параметров вызывает диалоговое окно, где отображаются текущие параметры. После редактирования новые параметры передаются в соответствующий экземпляр РНС. Блок на рис. 20 с названием «Корректировка конфигурации РНС», производит оценку степени изменений, которые требуется внести. В зависимости от параметров, которые были изменены, выполняются различные процедуры. Так, если изменение порогов возбуждения нейронов не требует каких-либо действий кроме установки нового значения соответствующей переменной экземпляра РНС, то изменение геометрии слоев повлечет за собой полную переинициализацию программной системы прогнозирования.

Таблица 1 — Каналы передачи информации в программной архитектуре нейросетевого прогнозирования с непрерывным обучением

Идентификатор канала (согласно рис. 20)	Выполняемая функция
Канал 1	Передача параметров функционирования РНС-2 из блока управления прогнозированием.
Канал 2	Передача параметров функционирования РНС-1 из блока управления прогнозированием.
Канал 3	Передача пространственно-временной модели и параметров прогнозирования из блока управления прогнозированием в РНС-2
Канал 4	Передача состояний слоев РНС-2 в блок управления прогнозированием для визуализации работы РНС-2
Канал 5	Передача состояний слоев РНС-1 в блок управления прогнозированием для визуализации работы РНС-1
Канал 6	Передача команды на выполнение прогноза с блока управления прогнозированием в РНС-1
Канал 7	Передача пространственно-временной модели из РНС-1 в блок управления прогнозированием
Канал 8	Передача команды на завершение работы программной системы в РНС-1
Канал 9	Передача команды на завершение работы программной системы в РНС-2

2.2. Обучение. Процедура обучения является цикличной и осуществляется в РНС-1 потактово на протяжении всего времени выполнения Этапа 2. Она предполагает получение данных из внешней среды и преобразование их в блоке предобработки в формат РНС. Далее происходит передача данных в РНС-1 и нейросетевая обработка. В процессе прохождения данных через РНС-1 формируется пространственно-временная модель. На выходе РНС-1 имеются результаты распознавания данных нейронной сетью. Они передаются на блок постобработки №1, а затем на выход системы. На каждом такте обучения РНС-1 проверяет наличие команды на прогноз. Если такая команда поступила, РНС-1 передает пространственно-временную модель в блок управления прогнозированием. Также проверяется наличие команды на завершение работы системы. Если такая команда поступила, то выполнение потока РНС-1 завершается.

2.3. Прогнозирование. Выполнение прогноза начинается с получения РНС-2 пространственно-временной модели, передаваемой с блока управления прогнозированием. РНС-2 путем присвоения своим нейронам и синапсам соответствующих значений пространственно-временной модели приводит свое состояние в соответствие с РНС-1. Кроме этого, РНС-2 получает с блока управления прогнозированием порядок и способ выполнения прогнозов, определенный ею по состоянию слоев РНС-1. Затем выполняется обработка скопированных данных в РНС-2 с формированием прогнозных данных. Эти данные передаются на блок постобработки №2, на выходе которого имеются прогнозные данные в исходном формате. По аналогии с РНС-1, при получении команды на завершение работы поток РНС-2 завершается.

2.4. Визуализация процесса обработки. Процесс визуализации включает в себя потактовую передачу состояний слоев РНС-1 и РНС-2 в блок управления прогнозированием и отрисовку этих данных на графическом интерфейсе. Данный процесс является опциональным и выполняется по усмотрению пользователя.

Этап 3. Завершение работы. Завершение работы начинается с подачи блоком управления прогнозированием соответствующих команд на РНС-1 и РНС-

2. При этом происходит завершение работы потоков №2 и №3. Уничтожаются экземпляры блоков пред- и постобработки. Последние актуальные настройки системы, хранимые в блоках чтения/записи параметров №1 и №2, сохраняются на внешнем носителе, если такая функция предусмотрена реализацией. После этого происходит освобождение ресурсов системы и выполнение программы завершается.

3.2 Алгоритм рационального использования памяти для систем нейросетевого прогнозирования с непрерывным обучением

В предыдущем параграфе была рассмотрена программная архитектура, реализующая методы прогнозирования с непрерывным обучением. Она предполагает наличие двух экземпляров РНС, работающих параллельно. РНС-1 работает в режиме обучения, а по команде на прогноз состояния ее элементов копируются в РНС-2 и запускается прогноз. Однако такая схема требует хранения в памяти достаточно больших объемов информации. В то же время, объем аппаратных ресурсов, выделяемый на прогнозирование с непрерывным обучением, в отдельных случаях является ограниченным (см. (19) - (21)).

Известная структура РНС с управляемыми элементами [136] предполагает наличие двух слоев нейронов, причем каждый нейрон одного слоя соединен со всеми нейронами другого слоя. Для эмуляции нейронной сети, содержащей N нейронов в слое, требуется дополнительно эмулировать N^2 синапсов. Каждый синапс в простейшем случае описывается двумя вещественными переменными: весовым коэффициентом (23) и функцией ослабления расходящихся единичных образов (26).

Также отметим, что от количества нейронов в слое РНС зависят ее возможности по интеллектуальной обработке информации [84]. Недопустимо уменьшение количества нейронов в целях экономии памяти, так как это негативно отразится на точности прогнозов. Уменьшение размеров логических полей РНС приведет к излишнему сжатию данных при кодировании. С другой стороны, сокращение количества таких полей ограничит возможности по пространственно-

временному связыванию информации, что также приведет к ошибкам в прогнозах.

В этих условиях отсутствует возможность полноценного использования предлагаемых методов на устройствах с ограниченными аппаратными возможностями. Необходим поиск решений, которые позволят без ущерба для точности прогнозов снизить объем памяти, потребляемой системой.

Для рационального использования памяти предлагается архитектура, сущность которой представлена на рисунке 21. В отличие от параллельной архитектуры (параграф 3.1), в которой используется пара полноценных РНС, в предлагаемой архитектуре (далее называемой «буферной») вводятся модуль эмуляции РНС-1 и РНС-2, модуль памяти и входной буфер, а также связи между ними [155]. Диаграмма классов UML для предлагаемой архитектуры приведена на рисунке 22.

Модуль эмуляции РНС-1 и РНС-2 представляет собой экземпляр РНС, который наделяется возможностями по смене режимов функционирования (обучение/прогнозирование), ускорении внутреннего времени и заморозке синаптических весов. Модуль памяти имеет размер, необходимый и достаточный для хранения состояний нейронов модуля эмуляции РНС-1 и РНС-2. Входной буфер устанавливается между входами системы прогнозирования и модуля эмуляции и наделяется возможностью по задержке входных данных при получении соответствующей команды от блока управления прогнозированием, а также по ускоренной подаче задержанных данных в модуль эмуляции.

На диаграмме классов UML (рис. 22) модуль эмуляции РНС-1 и РНС-2 обозначен `rnn12Core`, модуль чтения-записи параметров – `rnn12ParamsRW`, модуль памяти – `memoryUnit`, а входной буфер носит обозначение `rnnInputBuffer`. Имена классов графического интерфейса и модулей пред- и постобработки идентичны именам в параллельной архитектуре.

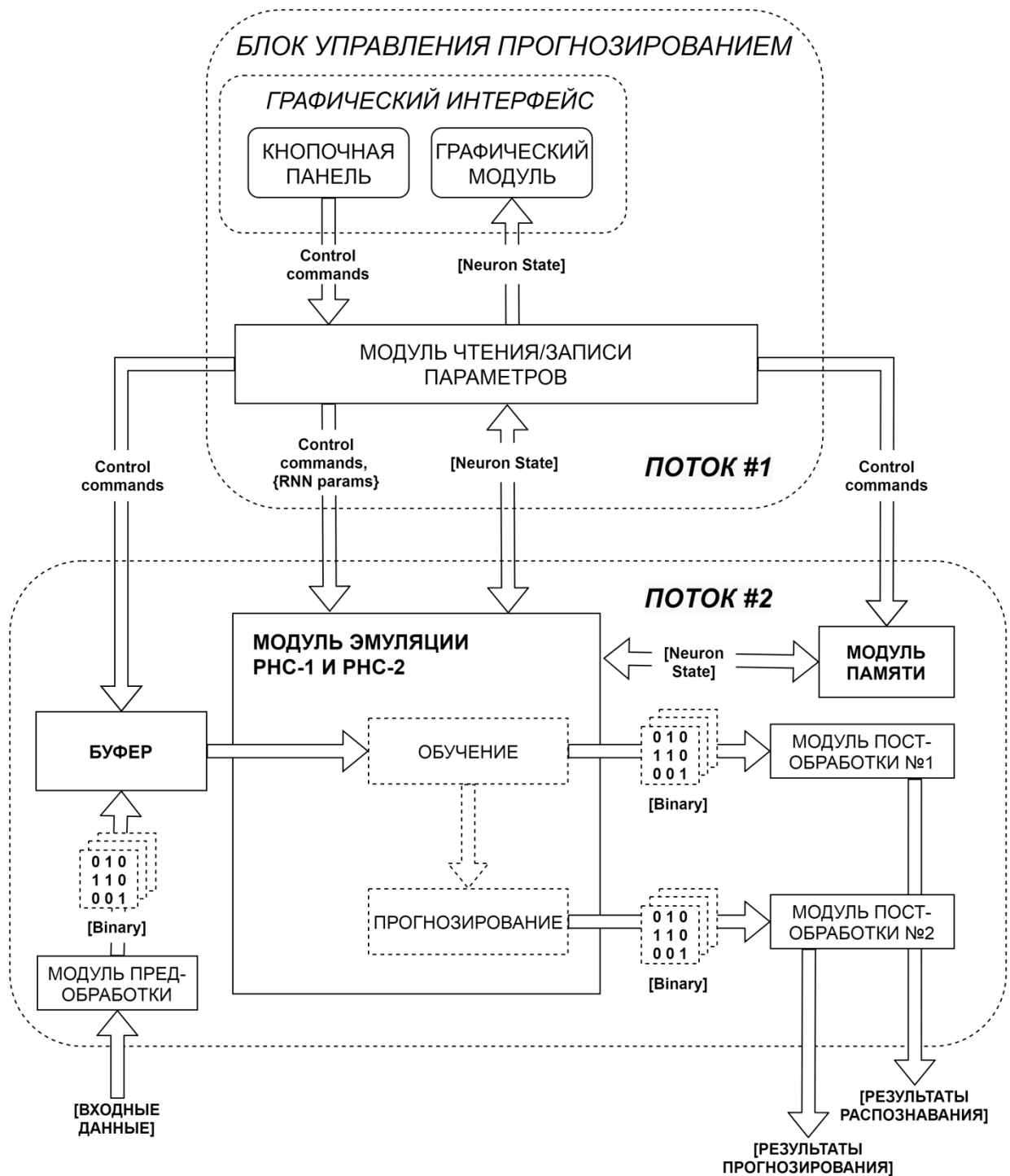


Рисунок 21 — Архитектура программной системы нейросетевого прогнозирования с рациональным использованием памяти. Обозначения потоков данных аналогичны рис. 18

Функционирование системы на этапе обучения происходит следующим образом. Входные данные кодируются в формат РНС и проходят без задержки через буфер и через модуль эмуляции, который в это время обучается. В

результате на синапсах РНС модуля эмуляции выстраивается пространственно-временная модель наблюдаемых событий. В отличие от параллельной архитектуры, в предлагаемой системе при получении команды на прогноз происходит заморозка синаптических весов модуля эмуляции, сохранение состояний его нейронов в модуле памяти, после чего модуль эмуляции переходит в режим прогнозирования и выполняет его. Если во время прогнозирования на вход системе поступают данные, то они задерживаются в буфере. Когда прогнозирование окончено, веса РНС модуля эмуляции размораживаются, а состояния нейронов восстанавливаются из модуля памяти. Таким образом, модуль эмуляции приводит свое состояние к такому, в каком он находился на момент получения команды на прогнозирование. После этого продолжается обучение. Если в буфере накопились данные, то внутреннее время модуля эмуляции ускоряется, и он обрабатывает эти данные в ускоренном режиме, чтобы синхронизировать свое состояние с внешней средой.

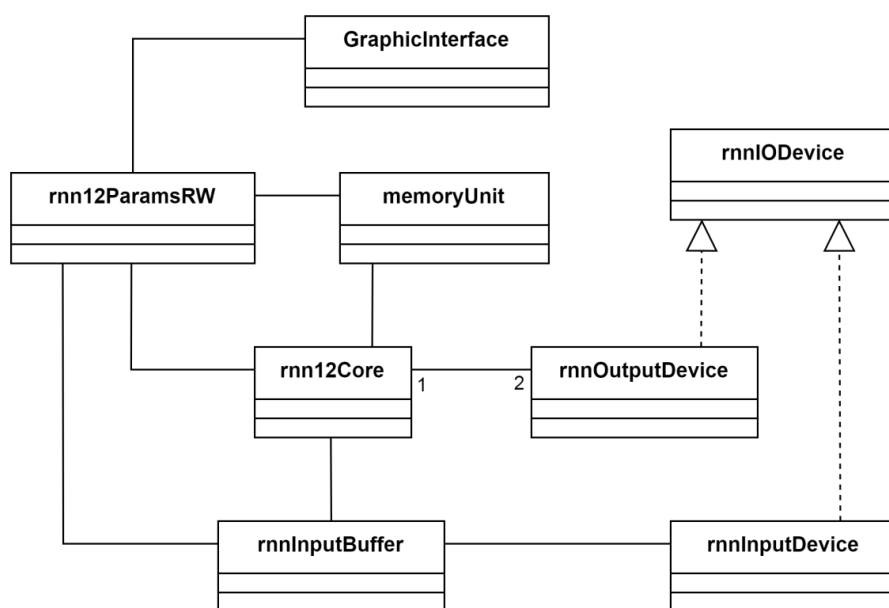


Рисунок 22 — Диаграмма классов UML для буферной архитектуры программной системы нейросетевого прогнозирования с непрерывным обучением

Главным достоинством предлагаемой архитектуры является сокращение объемов памяти, требуемых для функционирования программной системы

прогнозирования. Эффект достигается тем, что в предлагаемом подходе не требуется хранить синапсы второго экземпляра РНС. При этом стоит отметить, что блоки синапсов РНС являются наиболее ресурсоемкими элементами всей системы, в то время как введение буфера перед входом блока эмуляции не требует каких-либо существенных затрат. Поскольку количество синапсов растет в квадратичной зависимости от количества нейронов РНС, то предлагаемая архитектура позволяет снизить потребление памяти в общем случае в два раза.

Блок-схема, отражающая процесс функционирования программной архитектуры в соответствии с предлагаемым решением, приведен на рисунке 23. Вместо 3 потоков в базовом методе на ней имеется только два, связанные друг с другом с помощью механизмов межпоточкового взаимодействия. Список каналов передачи информации и их предназначение приведены в таблице 2.

Таблица 2 — Каналы передачи информации в программной архитектуре нейросетевого прогнозирования с сокращенными затратами памяти

Идентификатор канала (согласно рис. 23)	Выполняемая функция
Канал 1	Передача параметров функционирования блока эмуляции из блока управления прогнозированием.
Канал 2 Канал 3	Передача состояний слоев блока эмуляции в блок управления прогнозированием для визуализации работы РНС
Канал 4	Передача команды на выполнение прогноза с блока управления прогнозированием в блок эмуляции
Канал 5	Передача состояний слоев блока эмуляции в блок управления прогнозированием
Канал 6	Передача параметров прогнозирования из блока управления прогнозированием в блок эмуляции
Канал 7	Передача команды на завершение работы программной системы в блок эмуляции

Как видно из блок-схемы, предлагаемая архитектура является квазипараллельной. Это означает, что при обработке временных рядов с малой задержкой между поступлением входных данных или при частом обновлении прогнозов может иметь место запаздывание обучения, а также переполнение буфера. Эти особенности предлагаемого подхода необходимо учитывать при выборе архитектуры.

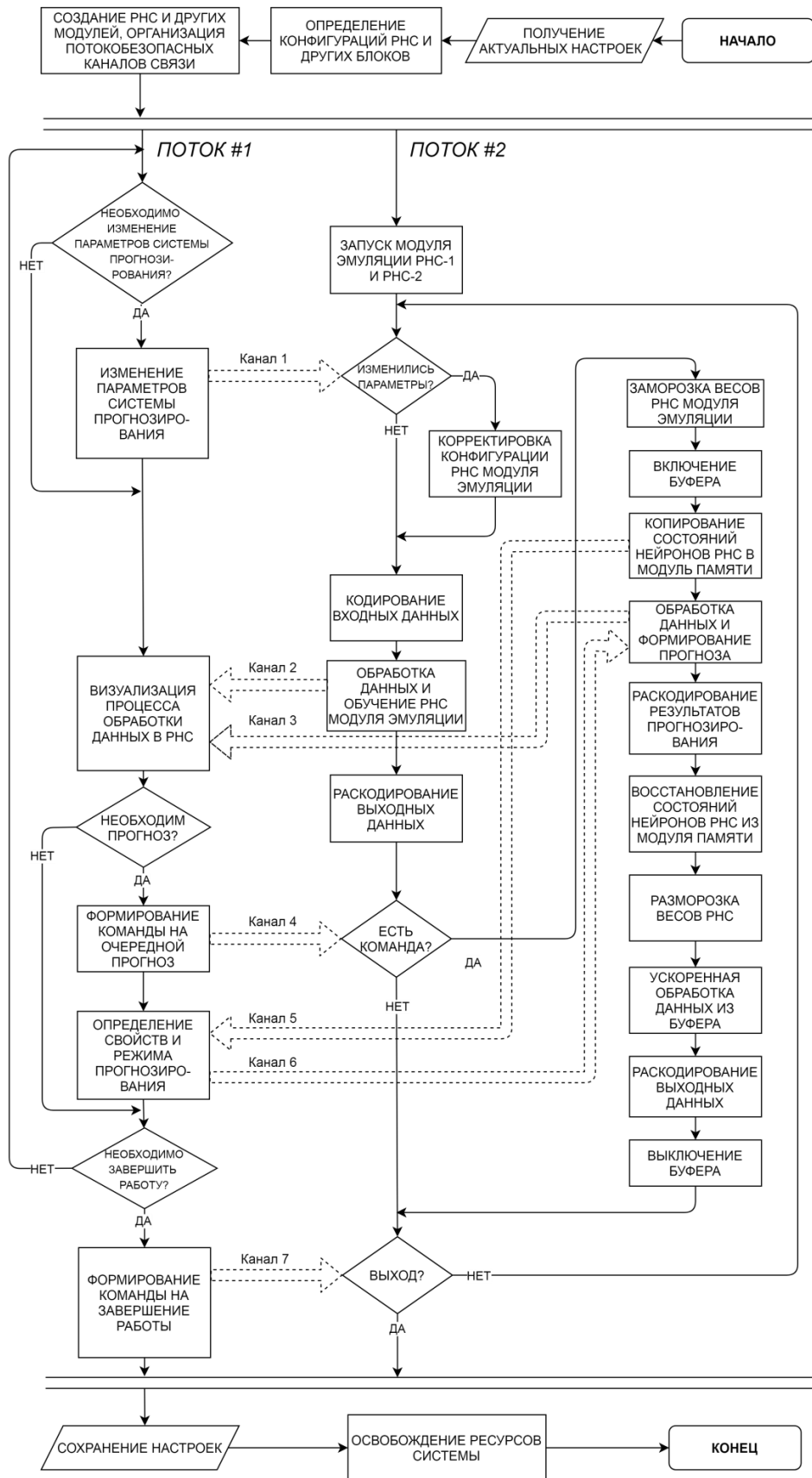


Рисунок 23 — Блок-схема, демонстрирующая функционирование программной архитектуры при сокращенных затратах памяти. Обозначения идентичны рис. 20

Параллельная и буферная программные архитектуры могут быть воплощены в программные средства на языке C++ (Qt) для CPU и для GPU и собраны компиляторами MSVC и MinGW под ОС семейства Windows и Linux.

3.3 Способ ускоренного расчета синаптических весов в программных системах нейросетевого прогнозирования событий

Особенности выполняемых задач по прогнозированию в ряде случаев накладывают ограничения не только на объем памяти, но и на скорость выполнения операций, или объем вычислений (см. (16) - (18)). Системы реального времени требуют своевременного получения результатов прогнозирования. В таких условиях разрабатываемая архитектура должна быть наделена возможностью сокращения объемов производимых ею вычислений для получения прогнозов при заданных ограничениях на точность прогнозирования.

В качестве решения этой проблемы предлагается особый способ вычисления весовых коэффициентов синапсов РНС [155].

Согласно (23) - (25), весовые коэффициенты РНС с управляемыми элементами формализуются функцией гиперболического тангенса. В силу этого, при выходе параметра $g_{ij}(t)$ за пределы некоторой области условной чувствительности синапсов $[G^-, G^+]$ его изменения становятся пренебрежимо малы (рис. 24).

Одновременно, поскольку $\Delta g_{ij}(t)$ может принимать определенное положительное (Δg^+) либо отрицательное (Δg^-) значение, существует дискретное множество возможных значений параметра $g_{ij}(t)$.

Принимая во внимание эти два условия, можно определить конечное множество $G = \{g_{ij}(t) \mid g_{ij}(t) = \pm n \times R(\Delta g^+, \Delta g^-), n = 1, 2, \dots, N; g_{ij}(t) \in [G^-, G^+]\}$ и соответствующее ему множество $K = \{k_{ij}(t) \mid k_{ij}(t) = \text{th}\left(\gamma \times \frac{g_{ij}(t)}{2}\right), g_{ij}(t) \in G\}$ значений возможных весовых коэффициентов синапсов. Здесь

$R(\Delta g^+, \Delta g^-)$ есть наибольшее значение, на которое нацело делятся Δg^+ и Δg^- , γ – некоторый положительный коэффициент.

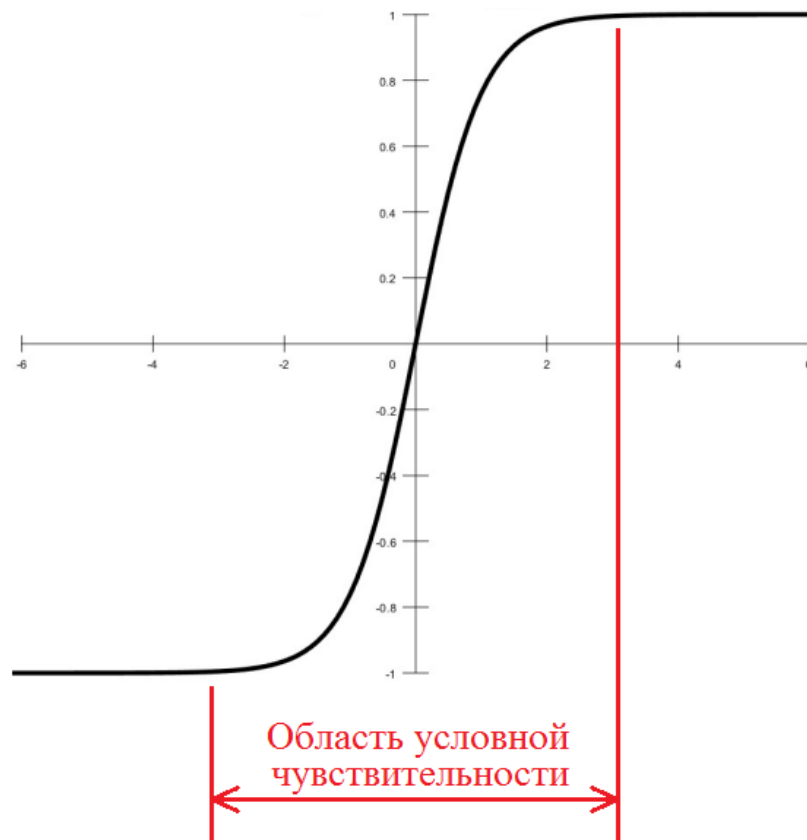


Рисунок 24 — Функция гиперболического тангенса и ее область условной чувствительности

В случаях, когда множество G невелико, предлагается при инициализации сети рассчитывать значения множеств G и K и хранить их в отдельных массивах. При этом отсутствует необходимость в выделении памяти под полноценный массив элементов $k_{ij}(t)$ для каждого синапса, а весовые коэффициенты рассчитываются в процессе выполнения тактов РНС следующим образом. Для ij -го синапса, у которого необходимо вычислить весовой коэффициент $k_{ij}(t)$, оценивается величина $g_{ij}(t)$. Если $g_{ij}(t) > G^+$, то $k_{ij}(t) = 1$; когда $g_{ij}(t) < G^-$, $k_{ij}(t) = -1$. В остальных случаях определяется элемент множества G , равный $g_{ij}(t)$, после чего определяется соответствующий ему элемент множества K . Значение этого элемента и является весовым коэффициентом ij -го синапса.

Предлагаемый подход позволяет в несколько раз сократить время, затрачиваемое на процесс прогнозирования в программной архитектуре, а также добиться небольшого сокращения времени обучения. Количественные оценки времени выполнения различных процессов при использовании предлагаемого подхода и без него будут изложены в главе 4.

Выводы по главе 3

На основе моделей и методов нейросетевого прогнозирования с непрерывным обучением, описанных в главе 2, предложены оригинальные архитектуры программных систем, отличающиеся новой структурой и правилами функционирования программных систем прогнозирования с непрерывным обучением, обеспечивающие программную реализацию предложенных моделей и методов и расширение их функций.

Параллельная архитектура предполагает наличие двух экземпляров РНС, работающих в параллельных потоках, модулей предобработки входных данных, постобработки выходных данных №1 и №2, модулей чтения/записи параметров №1 и №2, а также графического интерфейса, включающего в себя программно реализованные кнопочные панели для управления программной системой и графические модули для визуализации процессов нейросетевой обработки. Разработана UML-диаграмма и блок-схема, отражающая выполнение основных процессов в программной архитектуре: инициализация, обучение, прогнозирование, визуализация, редактирование параметров, завершение работы. Рассмотрены особенности и сущность каждого процесса.

Для рационального использования памяти в условиях существенных ограничений на аппаратные ресурсы предложена буферная архитектура программных систем нейросетевого прогнозирования. Суть ее заключается в отказе от физической реализации второго полноценного экземпляра РНС с отдельным потоком для него. При этом процессы обучения и прогнозирования выполняются на вновь введенном модуле эмуляции РНС-1 и РНС-2, в сущности представляющем собой идентичный экземпляр РНС с дополнительными

возможностями. Когда поступает команда на прогноз, из модуля эмуляции в модуль памяти копируются состояния нейронов, веса РНС замораживаются, и модуль эмуляции переходит в режим прогнозирования так, как это происходило бы на РНС-2 в параллельной архитектуре. Когда прогноз завершен, модуль эмуляции размораживает свои веса и восстанавливает состояния нейронов из модуля памяти. В отличие от параллельной архитектуры, в буферной имеется входной буфер, который задерживает поступающие на обработку данные во время прогнозирования. По его завершении модуль эмуляции обрабатывает эти данные в ускоренном режиме. Приведена UML-диаграмма и блок-схема, на которой раскрыты особенности выполнения процессов в буферной архитектуре. Новизна предлагаемого алгоритма заключается в отказе от выделения памяти под хранение синапсов РНС-2, в выполнении квазипараллельного обучения и прогнозирования, а достигаемый эффект заключается в сокращении объемов требуемой памяти в общем случае в два раза.

В целях повышения скорости обработки данных предложен способ ускоренного расчета синаптических весов синапсов РНС. Исходя из особенностей функции гиперболического тангенса и правил обучения синапсов, предложено рассчитывать конечное дискретное множество состояний истории обучения синапсов и соответствующих им весовых коэффициентов.

ГЛАВА 4. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ И ПРЕДЛОЖЕНИЯ ПО СОВЕРШЕНСТВОВАНИЮ НЕЙРОСЕТЕВОГО ПРОГНОЗИРОВАНИЯ ТРУДНОФОРМАЛИЗУЕМЫХ СОБЫТИЙ

4.1 Результаты и рекомендации по нейросетевому прогнозированию на примере предсказания событий, связанных с дорожным движением

Для проверки работоспособности предлагаемых методов и программных архитектур был проведен ряд исследований. В качестве исходных данных для прогнозирования использовалось три набора данных. Первый набор собирался автором самостоятельно, второй и третий являются общедоступными.

Первый набор собран с интернет-сервиса, информирующего в реальном времени о транспортной ситуации в г. Санкт-Петербург [156]. Сервис предоставлял информацию о средних скоростях движения на текущий момент в конкретной точке, а также давал прогноз для нее на следующий день с помощью встроенных в него алгоритмов. Было выбрано 8 точек, в которых в период с 14 по 27 февраля 2019 года в моменты времени {00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00, 21:00} регистрировалась информация о реальной скорости движения и прогнозе, даваемом сервисом на 1-дневный горизонт. Таким образом, для обучения и тестирования была сформирована база данных из $8 \times 9 \times 14 \times 8 \times 2 = 1792$ значений.

Второй и третий наборы – данные, собираемые Лабораторией исследования транспортных данных (TDRL) [157]. Этот набор содержит информацию о дорожной ситуации в городах-побратимах Миннесота и Ротчестер, полученную с микроволновых петлевых датчиков. Данные снимаются с 4500 датчиков, интервал записи составляет 30 секунд. Данные представлены в бинарном формате. Файлы второго набора имеют расширение «.v30» и содержат информацию об объеме трафика (англ. volume – количество транспортных средств, прошедших точку на дороге в течение 30-секундного интервала), записанную в 1-байтном формате. Файлы третьего набора с расширением «.c30» содержат данные о занятости (англ.

оссурансу – процент времени, в течение которого точка на дороге занята транспортными средствами [152]) в 2-байтном виде.

Второй набор данных содержит информацию о занятости дорог в период с 12 по 24 января 2018 года с детекторов №№ 17, 22, 24, 52, 62, 83.

Третий набор содержит данные об объеме трафика в период с 12 по 24 января 2018 года с детекторов №№ 51, 56, 61, 77, 80, 90, 453.

Эксперименты по определению точности прогнозирования скоростей движения. В первой серии экспериментов ставилась задача определения точности прогнозирования скоростей движения в Санкт-Петербурге с применением моделей и методов прогнозирования с непрерывным обучением. Для этого данные из первого набора были разбиты на две части. Первые 12 дней использовались для обучения РНС, а под прогнозирование оставались данные за последние 2 дня. Горизонт прогнозов также равнялся двум суткам.

Имеющиеся данные сначала кодировались в формат, воспринимаемый нейронной сетью. Диапазон скоростей движения разбивался на 8 поддиапазонов, каждому из которых на логических полях РНС ставились в соответствие определенные нейроны. Пример кодирования относительно размера 8×4 нейронов логических полей РНС приведен в Таблице 3.

Таблица 3 — Кодирование поддиапазонов скоростей движения на логических полях размером 8×4 нейронов

Скорость, км/ч	Матрица логического поля
[0-5]	1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
(5-10]	0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0
(10-15]	0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0
(15-20]	0 0 0 1 0 0 0 1

	0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0
(20-25]	0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0
(25-30]	0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0
(30-35]	0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0
(35-40]	0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1

Для изучения возможностей прогнозирования методом с временными сдвигами были сформированы две идентичные РНС с числом нейронов в каждом слое, равным 1920 единицам. Слои сетей были разделены на логические поля размером 8×8 нейронов. Количество таких полей в каждой строке составляло 6, а в каждом столбце – 5. Каждое логическое поле делилось на два подполя. При обучении РНС сдвинутый временной ряд кодированных скоростей подавался в первое подполе входного поля, а текущий временной ряд – во второе подполе. Кодирование скоростей движения отражено в Таблице 3. Наличие в матрице элементов «1» соответствует возбуждению соответствующих нейронов в логическом подполе.

Пример состояния первого слоя при обучении РНС-1 показан на рис. 25. Согласно этому рисунку, в каждом поле первые четыре линии нейронов относятся к первому подполю, а вторые четыре линии – ко второму подполю. При прогнозировании текущая последовательность СЕО, несущая информацию о скорости движения в определенной точке улицы, передавалась на первый канал РНС-2. На второй канал ничего не подавалось. В процессе функционирования РНС-2 результаты прогнозирования проявлялись во вторых подполях логических полей. В частном случае, показанном на рис. 26, на определенном этапе работы

сети они отражаются в 4, 7, 10, 13, 16, 22, 25 полях по направлению продвижения СЕО.

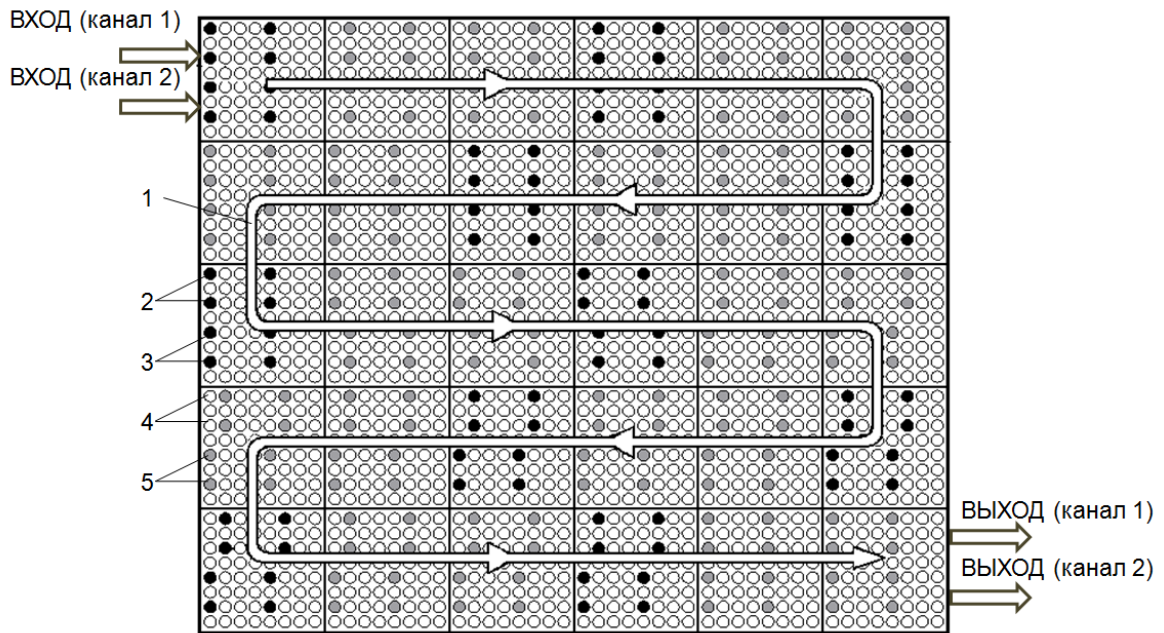


Рисунок 25 — Пример состояния первого слоя РНС-1 при обучении. Здесь 1 – направление движения СЕО вдоль слоя; 2, 3 – возбужденные нейроны (черные узлы) в первом и втором подполях соответственно; 4, 5 – нейроны в состояниях ожидания (белые узлы) и невосприимчивости (серые узлы) соответственно

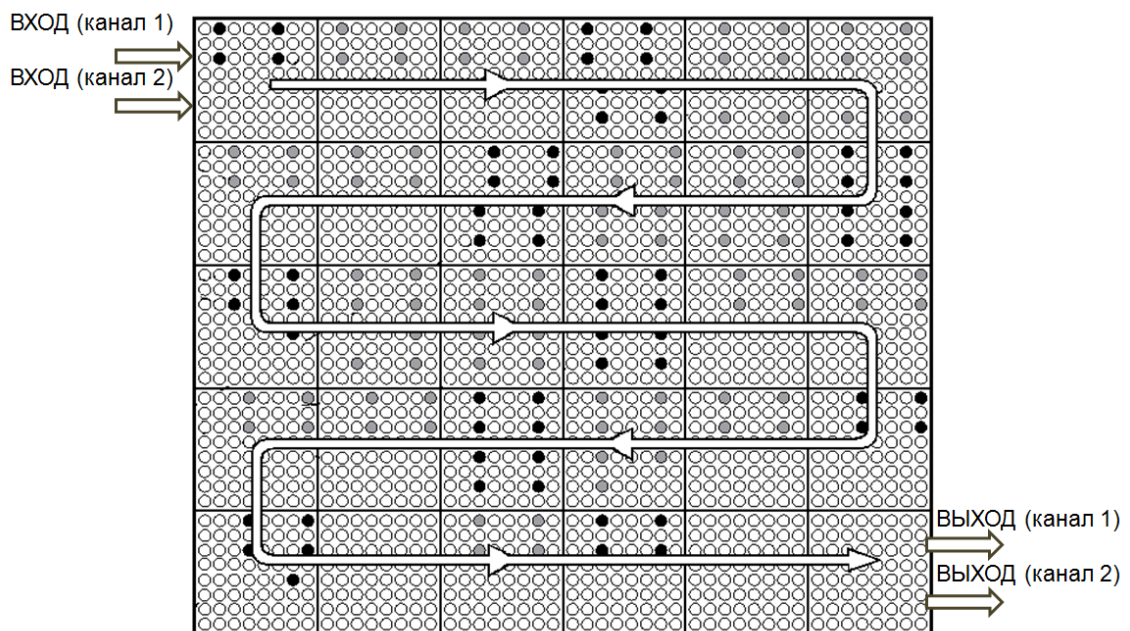


Рисунок 26 — Пример состояния первого слоя РНС-2 во время прогнозирования

Также была исследована возможность прогнозирования скоростей движения по методу без временных сдвигов. Количество нейронов в каждом слое

РНС в этих экспериментах составляло 960 единиц. Каждый слой этих сетей был разделен на 30 логических полей размером 8×4 нейрона. Отличие этих РНС от предыдущей версии состояло лишь в уменьшенных размерах логических полей. Схема продвижения СЕО по слоям сети аналогична предыдущему примеру. На РНС-1 также была возложена задача непрерывного обучения, связывания обрабатываемых сигналов в пространстве и времени, хранения их на элементах сети. Для предсказания состояние РНС-1 также копировалось в РНС-2. После этого при прямом прогнозировании значительно ускорилась обработка сигналов в РНС-2. Пример состояния первого слоя РНС-2 в режиме прогнозирования показан на рис. 27.



Рисунок 27 — Состояние первого слоя РНС-2 при прогнозировании методом без временных сдвигов. Здесь 1 – условный символ, отделяющий результаты прогнозирования (слева) от обрабатываемых данных

Ошибка прогнозирования определялась с использованием метрик средней абсолютной ошибки (MAE, формула (10)), среднеквадратической ошибки (RMSE, формула (11)) и средней абсолютной процентной ошибки (MAPE, формула (12)). В таблице 4 приведены результаты прогнозирования. Графики на рис. 28 и 29 демонстрируют полученные результаты прогнозирования в сравнении с реальными скоростями движения. Результаты прогнозирования на показаны сплошными линиями, а реальные данные – пунктирными.

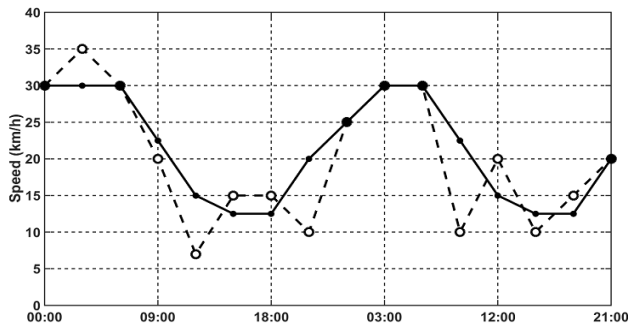
Таблица 4 — Показатели эффективности предлагаемых методов

Точки наблюдения дорожной ситуации	Метод прогнозирования с временными сдвигами			Метод прогнозирования без временных сдвигов		
	MAE (км/ч)	MAPE (%)	RMSE (км/ч)	MAE (км/ч)	MAPE (%)	RMSE (км/ч)
Улица 1	3.31	29.10	5.01	3.47	27.20	4.47
Улица 2	4.28	29.10	5.44	4.75	29.80	5.69
Улица 3	2.16	27.30	2.85	2.39	26.70	2.95
Улица 4	2.22	18.00	3.89	2.69	20.70	4.06
Улица 5	1.72	11.50	2.61	1.88	12.00	2.69
Улица 6	5.97	44.20	9.15	5.87	39.00	8.02
Улица 7	1.88	9.60	3.06	3.44	15.80	4.15
Улица 8	1.63	16.60	2.77	1.94	17.50	2.88
Среднее	2.89	23.20	4.35	3.30	23.60	4.36

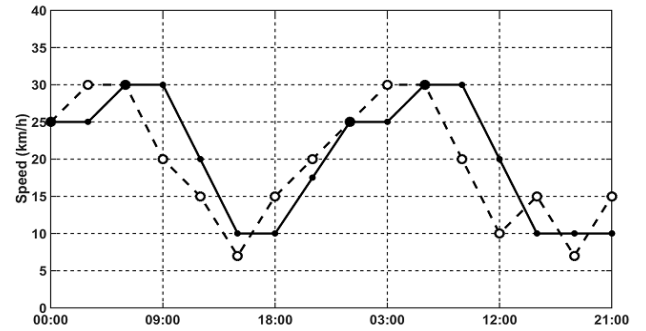
Как видно из таблицы 4, оба метода характеризуются высокими показателями для рассматриваемых условий. Несколько лучше с точки зрения MAE и MAPE отмечается метод прогнозирования с временными сдвигами (первый метод). Однако первый метод сложнее реализовать, чем второй. Кроме того, второй метод (без временных сдвигов) за счет ускорения обработки информации в РНС позволяет получать прогнозы быстрее, чем первый метод.

Используя те же исходные данные, была оценена эффективность прогнозирования транспортных потоков с использованием интернет-сервиса [156], а также модели ARIMA и нейронной сети LSTM. Результаты сравнения предложенных методов с этими подходами приведены в таблице 5.

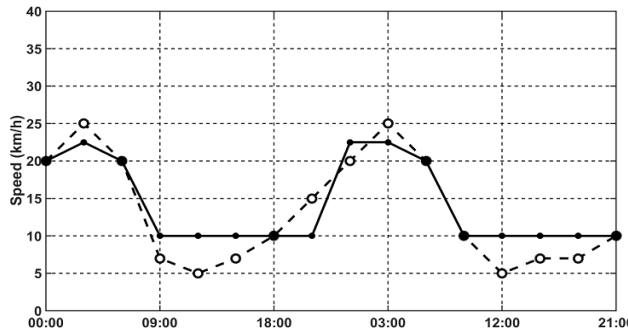
Анализ табл. 5 показывает, что предложенные методы в метриках MAE, MAPE, RMSE превосходят известные решения. В частности, выигрыши предложенного метода с временными сдвигами по сравнению с методом с LSTM составляют: 11.1% MAPE, 31.5% MAE, 15.5% RMSE.



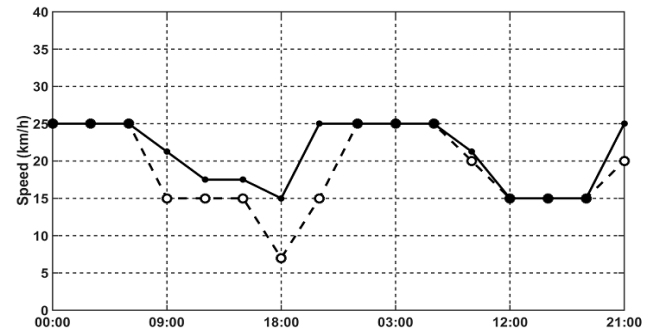
Улица 1



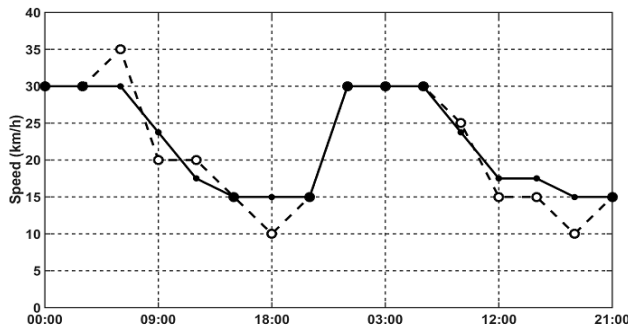
Улица 2



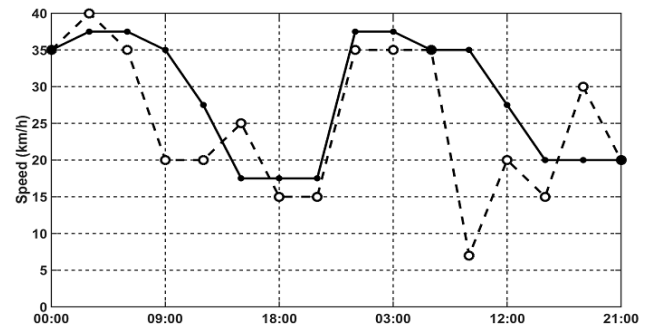
Улица 3



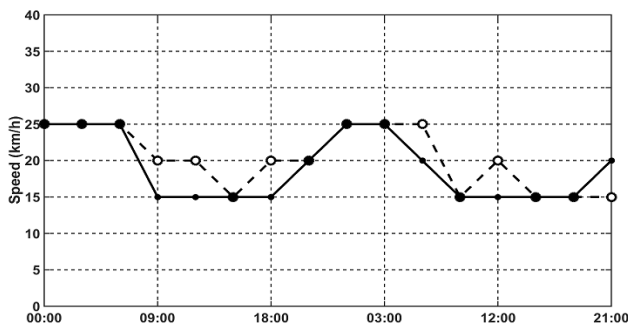
Улица 4



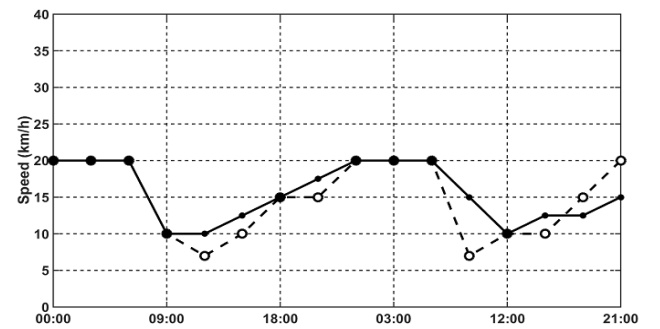
Улица 5



Улица 6

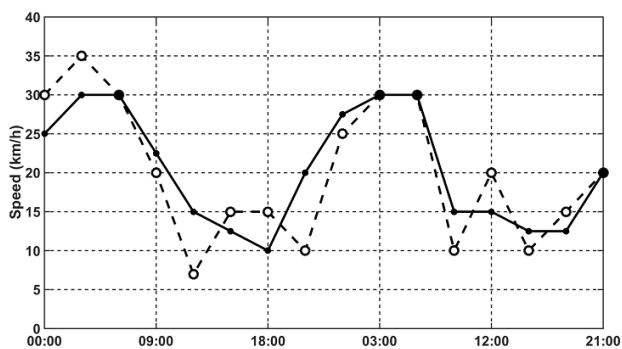


Улица 7

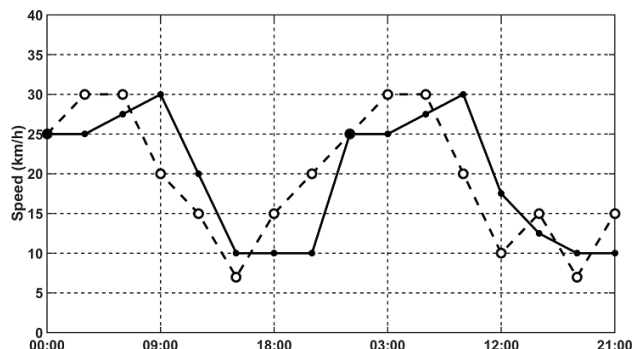


Улица 8

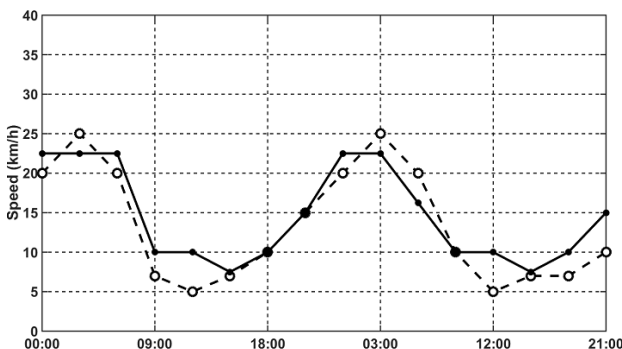
Рисунок 28 — Результаты прогнозирования методом с временными сдвигами. Пунктирными линиями с незакрашенными узлами обозначена реальные скорости, сплошной линией с закрашенными узлами – прогнозируемые скорости



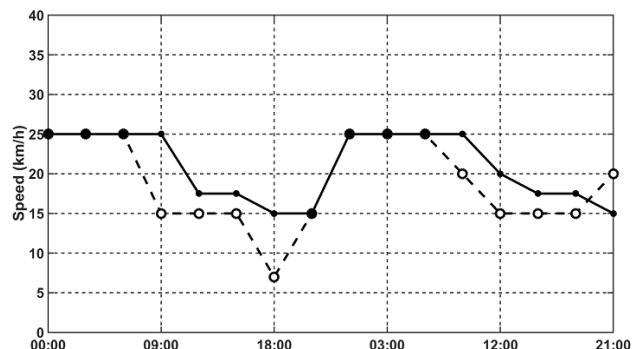
Улица 1



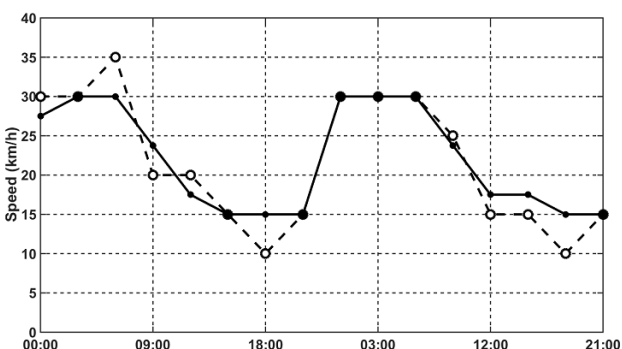
Улица 2



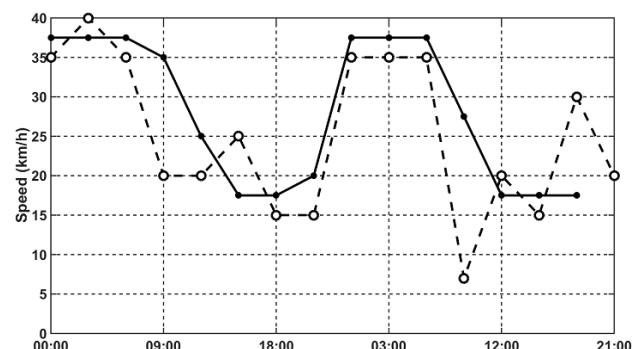
Улица 3



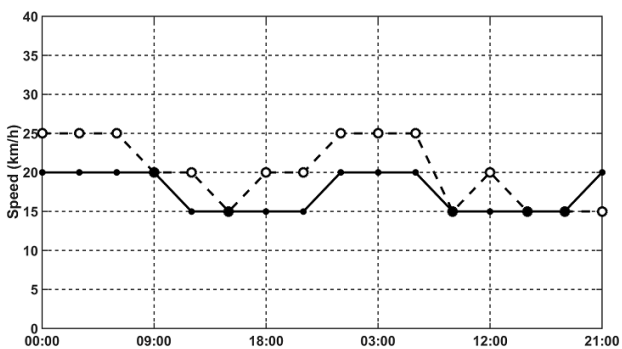
Улица 4



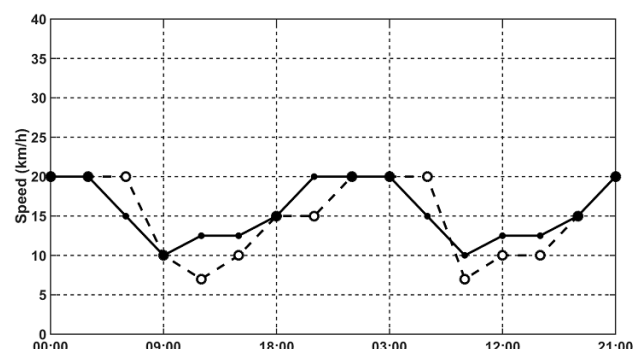
Улица 5



Улица 6



Улица 7



Улица 8

Рисунок 29 — Результаты прогнозирования методом без временных сдвигов. Обозначения аналогичны рис. 28

Таблица 5 — Сравнение показателей эффективности прогнозирования транспортных потоков по предлагаемым методам и известным аналогам

Метод	MAE (км/ч)	MAPE (%)	RMSE (км/ч)
Интернет-сервис «Яндекс.Пробки»	3.64	28.60	4.96
Модель ARIMA	4.18	30.00	4.90
Нейронная сеть LSTM	4.22	26.10	5.15
Предлагаемый метод с временными сдвигами	2.89	23.20	4.35
Предлагаемый метод без временных сдвигов	3.30	23.60	4.36

Для известных методов, основанных на векторной авторегрессии, многослойных LSTM со сверточными нейронными сетями [130], среднеквадратические ошибки прогноза для 1-часового горизонта могут варьироваться от 9.35 до 17.43 км/ч. В нашем случае для прогноза на 24 часа RMSE не превышает 4.36 км/ч.

Для полноты сравнения была проведена оценка производительности программного комплекса, реализующего предложенные методы. Скорость оценивалась на компьютере под управлением ОС Windows 7 x64, с процессором Intel Core i7-4790 3,60 ГГц и 32 Гб оперативной памяти. Результаты приведены в таблице 6.

Таблица 6 — Результаты оценки времени прогнозирования по методу с временными сдвигами и без сдвигов

Время прогнозирования по методу без временных сдвигов, мс	Время прогнозирования по методу с временными сдвигами, мс
31	74

Время прямого прогнозирования по предложенному методу с временными сдвигами транспортного потока составило 74 мс, а по методу без временных

сдвигов – 31 мс. Предложенные методы несколько уступают известным решениям по времени непосредственно прогнозирования, но при этом выигрывают за счет непрерывности обучения и отсутствия необходимости периодически переобучать систему. В известных методах полный цикл переобучения занимал от 19229 до 28674 мс [83].

Результаты сравнения показывают, что предложенные методы превосходят известные решения по точности прогноза более чем на 10% по показателю MAPE.

Эксперименты по определению точности прогнозирования объема дорожного трафика. Эффективность предложенных методов прогнозирования с непрерывным обучением нейронных сетей также оценивалась на втором наборе данных из Исследовательской лаборатории Transportation Data Research Laboratory [159]. Данные из этого набора за 12-23 января 2018 года использовались для обучения РНС, прогнозирование выполнялось на горизонт, равный 1 суткам (24 января 2018 года). Интервал составлял 3 часа, а количество прогнозных значений – 8.

Для предложенных методов получены показатели, приведенные в таблице 7. Из нее видно, что в отличие от предыдущих оценок по первому набору данных, метод без временных сдвигов несколько превосходил по эффективности метод с временными сдвигами транспортных потоков. Это можно объяснить следующими аргументами. Метод с временными сдвигами транспортных потоков хорошо работает при обработке больших, однородных временных выборок. На малых выборках его эффективность невысока из-за ограниченных возможностей ассоциативного вызова сигналов из памяти нейронной сети. Метод без временных сдвигов транспортных потоков с непрерывным обучением нейронной сети позволяет успешно прогнозировать как на малых, так и на больших выборках. При прогнозировании транспортных потоков по методу без временных сдвигов также возможно управление ассоциативным взаимодействием обрабатываемых сигналов в нейронной сети, что позволяет осуществлять более эффективное ассоциативное извлечение информации из нейросетевой памяти.

Таблица 7 — Показатели точности прогнозирования по предлагаемым методам на втором наборе данных

Метод	MAE	MAPE (%)	RMSE
С временными сдвигами	6.50	7.48	7.58
Без временных сдвигов	6.23	7.15	7.49

Таким образом, к основным преимуществам предложенных методов можно отнести:

- 1) более высокую точность прогнозирования по сравнению с известными решениями;
- 2) отсутствие необходимости предварительной подготовки и переподготовки модели, что дает возможность работать в режиме реального времени.

Эксперименты по определению точности прогнозирования занятости дорог. Третья серия экспериментов проводилась на третьем наборе данных и ставила целью определение оптимальной конфигурации нейросетевых слоев для прогнозирования событий. Рекуррентные нейронные сети с управляемыми элементами [153] имеют возможность наделения слоев различными конфигурациями: линейными (рис. 30А), петлевыми (рис. 30Б), спиральными с постоянным радиусом (рис. 30В), сходящимися (рис. 30Г) и расходящимися (рис. 30Д) спиралями.

В данной серии экспериментов при прочих равных параметрах РНС оценивались возможности прогнозирования рекуррентными нейронными сетями с управляемыми элементами. При оценке точности использовались те же метрики, что и в предыдущих экспериментах.

Данные из третьего набора за 12-23 января 2018 года использовались для обучения РНС, в то время как прогнозирование выполнялось на горизонт, равный 1 суткам (24 января 2018 года). Интервал составлял 3 часа, а количество прогнозируемых значений – 8.

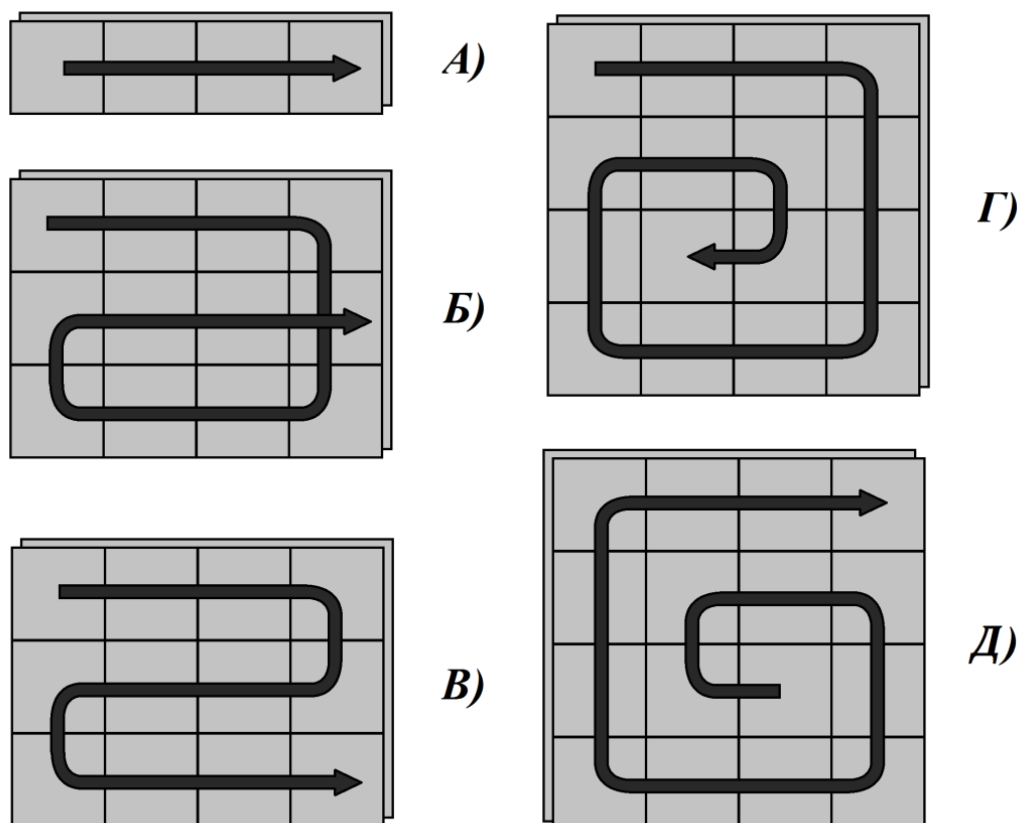


Рисунок 30 — Примеры конфигураций слоев РНС с управляемыми элементами [13]

Результаты оценивания точности прогнозирования по выбранным метрикам для рассмотренных пяти конфигураций программной системы представлены в таблице 8. В этой же таблице для сравнения приведены оценки, полученные на таких же исходных данных с помощью модели Бокса-Дженкинса (ARIMA).

Таблица 8 — Результаты оценивания точности прогнозирования

Вид структуры РНС	Показатель	Датчик 1	Датчик 2	Датчик 3	Датчик 4	Датчик 5	Датчик 6	Датчик 7	Среднее
Линейная	MAE	1,63	1,39	0,34	0,53	0,30	0,68	0,69	0,79
	MAPE	27,90	19,20	22,70	10,00	11,60	30,50	20,40	20,30
	RMSE	2,22	2,18	0,44	1,06	0,43	0,98	0,98	1,18
Спираль с	MAE	1,15	1,07	0,47	0,53	0,18	0,68	0,63	0,67

постоянным диаметром	MAPE	18,80	13,80	25,90	10,00	8,90	30,50	11,30	17,00
	RMSE	1,80	1,78	0,65	1,06	0,21	0,98	0,96	1,06
Петлевая	MAE	1,29	1,44	0,33	0,60	0,18	0,69	0,63	0,74
	MAPE	20,70	24,00	22,30	14,40	8,90	30,50	11,30	18,90
	RMSE	2,01	2,37	0,42	1,08	0,21	0,99	0,96	1,15
Спираль сходящаяся	MAE	1,29	1,27	0,33	0,59	0,18	0,65	0,63	0,71
	MAPE	21,10	16,60	22,30	10,80	8,90	28,50	11,30	17,10
	RMSE	1,87	2,04	0,42	1,09	0,21	0,97	0,96	1,08
Спираль расходящаяся	MAE	0,97	1,13	0,44	0,55	0,18	0,71	0,63	0,66
	MAPE	15,50	14,50	25,80	10,30	8,90	32,40	11,30	17,00
	RMSE	1,72	1,95	0,55	1,07	0,21	0,99	0,96	1,06
ARIMA	MAE	1,72	1,34	0,61	0,89	0,36	0,80	0,65	0,91
	MAPE	30,10	19,00	30,80	17,40	13,90	41,20	18,70	24,40
	RMSE	2,50	2,03	1,12	1,16	0,50	1,19	0,94	1,35

Результаты показывают, что все пять конфигураций продемонстрировали достаточно высокую точность прогнозирования объема дорожного трафика. Она выше точности ARIMA. Наилучшие результаты характерны спиральным структурам, которые оказались предпочтительнее, чем структуры линейные.

Наблюдаемый результат объясним преимуществом спиральных нейросетевых конфигураций с точки зрения плотности ассоциативного связывания вызываемых сигналов с теми СЕО, которые уже находятся на слоях РНС. В то же время линейная структура может обеспечить относительно плотное связывание только ближайших СЕО, а влияние других единичных образов по мере удаления становится пренебрежимо малым.

4.2 Результаты и рекомендации по нейросетевому прогнозированию лексического содержания новостных лент

В целях прогнозирования лексического содержания новостных информационных лент в период с 16 по 22 марта 2019 г. через промежутки времени 15 мин. собиралась информация с сайта Lenta.ru. Полученные данные предварительно обрабатывались, включая фильтрацию (удаление служебных символов, исключение слов, которые не являются частью предварительно скомпилированной онтологии, и т.д.) [3]. В результате для каждого момента времени имелись наборы слов, отражающие лексическое содержание новостей в этот момент.

Списки слов кодировались на полях РНС размером 306 и 1000 нейронов (по размерам словарей) в виде булевых матриц. Значение «1» матрицы означало присутствие слова, поставленного в соответствие данному элементу матрицы, а значение «0» – отсутствие. Пример последовательности совокупностей единичных образов (СЕО), несущей информацию о лексическом содержании новостей, для словаря размером 306 слов, приведен на рисунке 31.

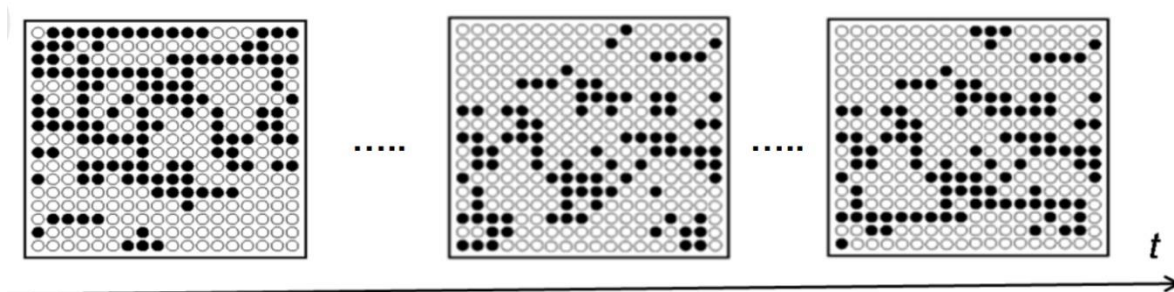


Рисунок 31 — Пример последовательности СЕО, несущей информацию о лексическом содержании новостных текстов, для словаря размером 306 слов [3]

Затем полученные последовательности СЕО подавались в систему прогнозирования. На выходе имели место спрогнозированные СЕО, информация из которых после раскодирования сравнивались с фактическим состоянием новостных лент. С учетом специфики данных и условной важности слов, ошибки прогнозирования оценивались следующим образом:

$$PE0 = \frac{1}{N} \sum_{t=1}^N \frac{\text{Errors of type "0"}_t}{M} \cdot 100\%, \quad (36)$$

$$PE1 = \frac{1}{N} \sum_{t=1}^N \frac{\text{Errors of type "1"}_t}{M} \cdot 100\%, \quad (37)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N \sum_{p=1}^M |(\delta_{tp} - \delta_{tp}^*) \cdot TF_IDF_{tp}|, \quad (38)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{\sum_{p=1}^M |(\delta_{tp} - \delta_{tp}^*) \cdot TF_IDF_{tp}|}{\sum_{p=1}^M \delta_{tp} \cdot TF_IDF_{tp}} \cdot 100\%, \quad (39)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N \left(\sum_{p=1}^M |(\delta_{tp} - \delta_{tp}^*) \cdot TF_IDF_{tp}| \right)^2}, \quad (40)$$

где *Errors of type "0"* – дефекты типа «0» (слово есть, но не спрогнозировано), *Errors of type "1"* – дефекты типа «1» (слова нет, но оно спрогнозировано), δ_{tp} – булевы значения присутствия ($\delta_{tp} = 1$) либо отсутствия ($\delta_{tp} = 0$) p -го слова из словаря размером M на момент времени t ; δ_{tp}^* - прогноз присутствия этого слова; $TF_IDF_{tp} = f_{tp} \cdot \log\left(\frac{W}{w_p}\right)$ есть коэффициент важности p -го слова из словаря, вычисляемое как частота встречаемости f_{tp} этого слова в тексте на момент времени t , умноженное на логарифм отношения количества исследуемых наборов слов W к количеству наборов w_p , в которых есть p -е слово.

Эксперименты проводились для словарей размером 306 и 1000 слов, на горизонты 1 час и 4 часа, с интервалами 15 и 60 минут. Результаты приведены в таблицах 9-11.

Результаты нейросетевого прогнозирования демонстрируют, что для горизонта 1 час при 15, 30, 45, 60 минутах средние ошибки типа «0» в процентах равны соответственно 7.7, 12.8, 15.6, 20.0%. Средние ошибки типа «1» составляют 12.2, 9.4, 11.0, 20.5% соответственно. Показатель MAPE изменяется с 30.9% до 39.7%. Обратим внимание, что в этих экспериментах размеры полей, на которых

разделены слои РНС, более чем в три раза больше, чем в экспериментах со словарем 306 слов. Однако количество таких полей равно 6 против 10 в экспериментах с размером словаря 306. Благодаря большему словарю и размерам РНС общее количество ошибок в рассматриваемом варианте прогнозирования составляет на 10.5% меньше, чем в первом варианте (таблица 9) [3].

Таблица 9 — Результаты прогнозирования лексического содержания новостных лент для горизонта 1 час и словаря 306 слов

Момент времени	15 мин	30 мин	45 мин	60 мин
Число слов	126	118	117	115
Errors of type “0”	5.9	7.1	8.2	9.7
Errors of type “1”	22.4	31.6	30.3	29.4
PE0	4.7	6.0	7.0	8.4
PE1	17.8	26.8	25.9	25.5
MAE	0.18	0.26	0.28	0.29
MAPE	18.6	21.8	22.7	23.1
RMSE	0.19	0.29	0.31	0.31

Таблица 10 — Результаты прогнозирования лексического содержания новостных лент для горизонта 1 час и словаря 1000 слов

Момент времени	15 мин	30 мин	45 мин	60 мин
Число слов	174	179	181	171
Errors of type “0”	13.4	23.0	28.3	34.1
Errors of type “1”	21.1	16.9	20.0	35.0
PE0	7.7	12.8	15.6	20.0
PE1	12.2	9.4	11.0	20.5
MAE	0.14	0.17	0.23	0.36
MAPE	30.9	34.0	36.8	39.7
RMSE	0.19	0.22	0.29	0.44

Таблица 11 — Результаты прогнозирования лексического содержания новостных лент для горизонта 4 часа и словаря 1000 слов

Момент времени	15 мин	30 мин	45 мин	60 мин
Число слов	122	125	124	126
Errors of type “0”	15.5	28.5	32.2	37.9
Errors of type “1”	33.6	30.1	40.2	44.2
PE0	12.7	22.7	25.9	30.1
PE1	27.6	24.0	32.3	35.1
MAE	0.32	0.40	0.51	0.62
MAPE	42.3	40.8	41.2	41.9
RMSE	0.35	0.44	0.53	0.64

Согласно таблице 11 для 4-часового горизонта в 1, 2, 3, 4 часа ошибки типа «0» составляют 12.7, 22.7, 25.9, 30.1% соответственно, а ошибки типа «1» составляют 27.6, 24.0, 32.3, 35.1%. Значения MAPE составляют 42.3, 40.8, 41.2, 41.9% соответственно. В этом случае значения суммарной ошибки значительно выше, чем в экспериментах с горизонтом 1 час. Это обусловлено увеличением горизонта и интервалов прогнозирования. С их увеличением возможности учета пространственно-временных отношений между словами в ленте новостей значительно ограничиваются [3].

Для оценки возможностей рекуррентных нейронных сетей с управляемыми элементами в сравнении с известными подходами аналогичные эксперименты были проведены с использованием нейронной сети LSTM. Сравнение проводилось для словаря объемом 306 слов на 1-часовом горизонте и для словаря объемом 1000 слов на 1-часовом и 4-часовом горизонтах. Результаты представлены в таблице 12.

Анализ данных этой таблицы показывает, что как на 1-часовом, так и на 4-часовом горизонтах предлагаемый метод превосходит архитектуру LSTM. Так, для словаря из 306 слов общая процентная ошибка типа «0» и типа «1» для LSTM выше на 12.2%, а MAPE выше на 7.6%. Для словаря из 1000 слов и 1-часового

горизонта прогноза эти показатели составляют 20.0% и 4.0% соответственно, а для 4-часового горизонта они составляют 10.4% и 2.9%, соответственно [3].

Таблица 12 — Сравнительные показатели эффективности прогнозирования нейронными сетями LSTM и РНС с управляемыми элементами

Размер словаря	306 слов		1000 слов		1000 слов	
Период прогнозирования	1 час		1 час		4 часа	
Интервал прогнозирования	15 мин		15 мин		60 мин	
Метод прогнозирования	Предл. метод	LSTM	Предл. метод	LSTM	Предл. метод	LSTM
Errors of type “0”	7.7	24.4	24.7	55.1	28.5	39.4
Errors of type “1”	28.4	26.3	23.3	28.0	37.0	38.9
PE0	6.5	20.5	14.0	31.3	22.9	31.7
PE1	23.9	22.1	13.2	15.9	29.7	31.3
MAPE	0.25	0.39	0.23	0.36	0.46	0.57
MAE	21.5	29.1	35.6	39.6	41.6	44.5
RMSE	0.26	0.40	0.25	0.38	0.48	0.59

После анализа результатов прогнозирования наборов слов в лентах новостей, полученных с помощью предлагаемых моделей и методов, был сделан вывод о том, что они вполне приемлемы для горизонта прогноза до двух часов [3].

4.3 Результаты оценивания эффективности архитектурных решений по нейросетевому прогнозированию

В целях оценки эффективности архитектурных решений, изложенных в главе 3, а также выработки рекомендаций по применению программных систем проводился дополнительный эксперимент. Исследованы три варианта программного обеспечения.

Первый вариант программного обеспечения (ПО) отражал стремление максимального сокращения объема вычислений, необходимых для

функционирования программной системы прогнозирования (см. (16) – (18)). В нем значения синапсов $\beta(r_{ij}(t))$, $k_{ij}(t)$ и $w_{ij}(t)$ хранятся в массивах.

Второй вариант ПО аналогичен первому варианту, за тем исключением, что для хранения значений $k_{ij}(t)$ и $w_{ij}(t)$ память не выделяется, и они рассчитываются на ходу. Иными словами, данный вариант отражает идею максимальной экономии памяти (см. (19) - (21)).

Третий вариант ПО использует ускоренный расчет весов синапсов. Весовые коэффициенты $k_{ij}(t)$ хранятся в массиве, но хранение организовано специальным образом, что позволяет без потери точности сократить размер массива. Данный вариант отражает комбинированный подход между экономией памяти и вычислительных операций.

Для проведения экспериментов использовался компьютер под управлением ОС Windows 7 x64, с процессором Intel Core i7-4790 CPU 3.6 GHz, 32 ГБ ОЗУ. Программное обеспечение разрабатывалось на языке программирования C++, с применением фреймворка Qt 5.14.1. Сборка производилась компиляторами MinGW 7.3.0 64bit и MSVC 14.0 amd64.

Тесты проводились на примере задачи по прогнозированию наборов слов. Эти наборы кодировались на матрицах размером 20×50 (размер словаря равен 1000 слов), в которых за каждым нейроном закрепляется определенное слово из словаря [3]. По аналогии с экспериментами, описанными в параграфе 4.2, размер РНС для прогнозирования составлял 4×6 логических полей, количество тактов рефракторности нейронов равнялось 7. Таким образом, количество нейронов в слое было 24000. Сеть обучалась на 50 наборов слов, прогнозировались следующие 10 наборов.

Результаты приведены в таблицах 13 и 14. Они показывают, что расчет параметров синапсов «на ходу» (второй вариант ПО) является более дорогостоящим с точки зрения быстродействия. В то же время, комбинированный метод (третий вариант ПО) демонстрирует практически столь же низкие затраты памяти, как и второй вариант ПО, при этом на прогнозирование затрачивается в 5-6 раз меньше времени. Это свидетельствует об эффективности предложенных

решений по ускоренному расчету синаптических весов РНС. Однако в ситуациях, когда критерий времени является определяющим, предпочтительно использовать первый вариант ПО, несмотря на то, что он требует в два раза больше памяти.

Таблица 13 — Значения показателей эффективности ПО (компилятор MinGW)

	Первый вариант ПО	Второй вариант ПО	Третий вариант ПО
Объем памяти, МБ	2153.8	1078.1	1078.4
Время обучения, мс	88960	119104	52133
Время прогнозирования, мс	698	40260	7823

Таблица 14 — Значения показателей эффективности ПО (компилятор MSVC)

	Первый вариант ПО	Второй вариант ПО	Третий вариант ПО
Объем памяти, МБ	2153.8	1078.1	1078.4
Время обучения, мс	45745	151093	148211
Время прогнозирования, мс	1240	39604	7074

4.4 Результаты использования разработанных методов и архитектур в исследованиях по обнаружению БПЛА

Разработанные методы и программные архитектуры были использованы в исследованиях по обнаружению и траекторному сопровождению малоразмерных беспилотных летательных аппаратов в АО «НТЦ РЭБ». Одной из задач, стоящих перед системой обнаружения БПЛА, является регистрация его радиоизлучения на фоне сложной радиоэлектронной обстановки. Проблема заключается не только в присутствии сторонних радиоизлучений, мешающих выявлению сигнала БПЛА, но и в том, что зачастую структура такого сигнала априорно неизвестна. В настоящее время рынок малоразмерных беспилотных воздушных средств весьма разнообразен и продолжает расти, а структура их сигналов крайне разнообразна. В таких условиях использование предобученных систем малоэффективно.

Требуются средства, способные выявлять новизну в радиосигналах для дальнейшего анализа на предмет принадлежности их беспилотным аппаратам.

Разработанная система прогнозирования событий с непрерывным обучением была использована для обнаружения новизны в радиоэлектронной обстановке. Идея заключалась в том, чтобы обучить РНС-1 путем подачи в нее сигналов в виде спектров, которые представляют собой нормальную радиоэлектронную обстановку. При обучении в системе прогнозирования будет сформирована пространственно-временная модель радиоэлектронного фона. Затем предлагается осуществлять прогнозы по методу без временных сдвигов на один временной шаг вперед и сравнивать спрогнозированный спектр со спектром реального радиосигнала, записанного в прогнозируемый момент времени. Путем оценки величины расхождения между реальным и спрогнозированным спектрами и сравнения этой величины с установленным порогом возможно выявлять аномальные сигналы, не присутствовавшие в радиоэлектронной обстановке во время обучения.

На рисунке 32 приведены примеры спектров сигналов в полосе 970...990 МГц в обычной обстановке и при появлении сигнала БПЛА «Орлан-10».

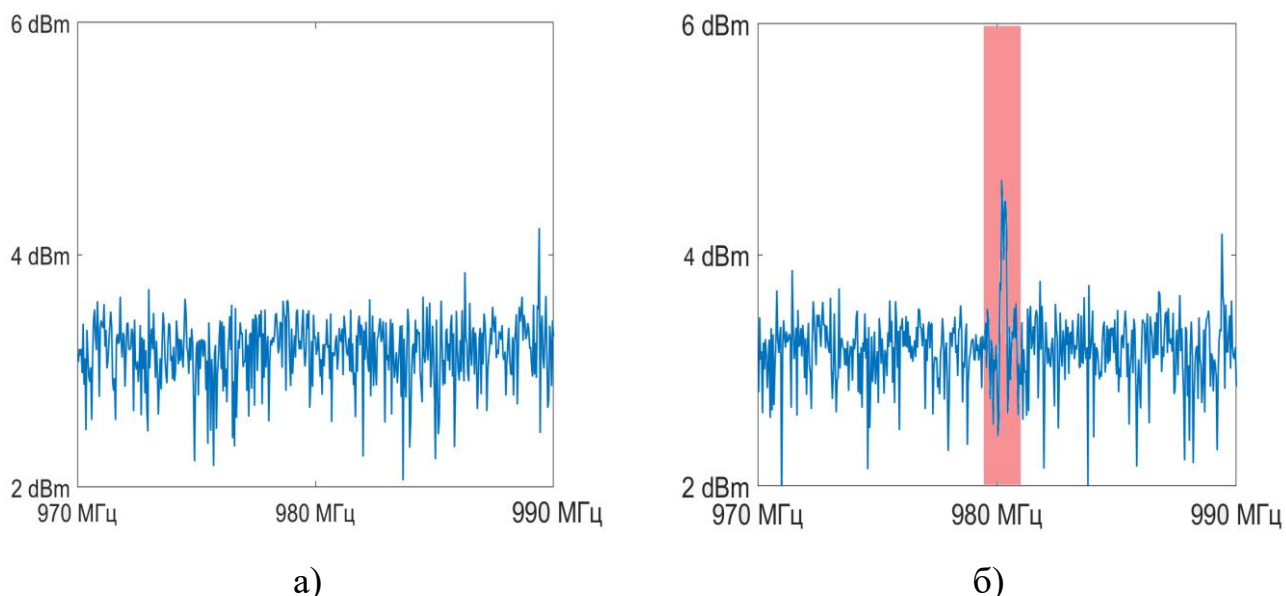


Рисунок 32 — Спектральное представление радиосигнала в полосе 970...990 МГц в обычной обстановке (а) и при появлении аномального сигнала (канал управления БПЛА «Орлан-10») на частоте 980 МГц (б). Новый сигнал выделен красной полосой

Предложенный метод сравнивался с альтернативным подходом, основанным на применении автокодировщика. Суть его заключалась в обучении нейросети распознаванию спектров сигналов обычной радиоэлектронной обстановки. Во время работы в автокодировщик также подаются сигналы, принимаемые из эфира, и оценивается ошибка восстановления как разность между исходным спектром и откликом нейросети на него. При подаче в автокодировщик сигнала, не присутствовавшего в обучающем наборе, такой сигнал не будет корректно распознан, и ошибка восстановления превысит порог обнаружения.

Согласно проведенным экспериментам на радиосигналах в диапазоне 970...990 МГц, вероятность ложного срабатывания у предлагаемого подхода составила 0.09 против 0.14 для автокодировщика, а вероятность пропуска – 0.05 против 0.06 у автокодировщика. Стоит отметить, что подход, основанный на использовании пары РНС, требует больших аппаратных ресурсов, однако за счет обработки информации не только в пространстве, но и во времени, способен обеспечить большую точность.

4.5 Программная система нейросетевого прогнозирования событий с непрерывным обучением

Программная система нейросетевого прогнозирования, реализующая предложенные модели и методы и использовавшаяся в экспериментах, реализована на языке программирования C++ с использованием фреймворка Qt версии 5.14.1.

Разработанная программная система является конкретизацией архитектур, изложенных в главе 3. Ее файловый состав представлен на рисунке 31. Рассмотрим более подробно реализацию каждого из элементов программной реализации.

Первый и второй экземпляры РНС являются экземплярами классов `rnn1Core` и `rnn2Core` соответственно, которые в свою очередь реализуют класс `rnnCore`. Часть функционала нейросетевых ядер и вспомогательных функций вынесены в

отдельные файлы `rnnCore_WCoeff`, `rnnCore_Structures`, `rnnCore_Support`, `rnnCore_Synapses`. Класс `rnnCore` описывает основную логику функционирования нейросетевого ядра. В файлах `rnnCore_Structures` вынесены объявления и реализация основных элементов нейронной сети, таких как нейрон, синапс, логическое поле и т.д. Класс `rnnCore_Synapses` содержит описание блоков динамических синапсов РНС, в класс `rnnCore_WCoeff` вынесены переменные и функции, связанные с обучением сети. Класс `rnnCore_Support` предоставляет вспомогательные функции, такие как определение индекса нейрона в массиве по его координатам на слоях РНС и тому подобные.

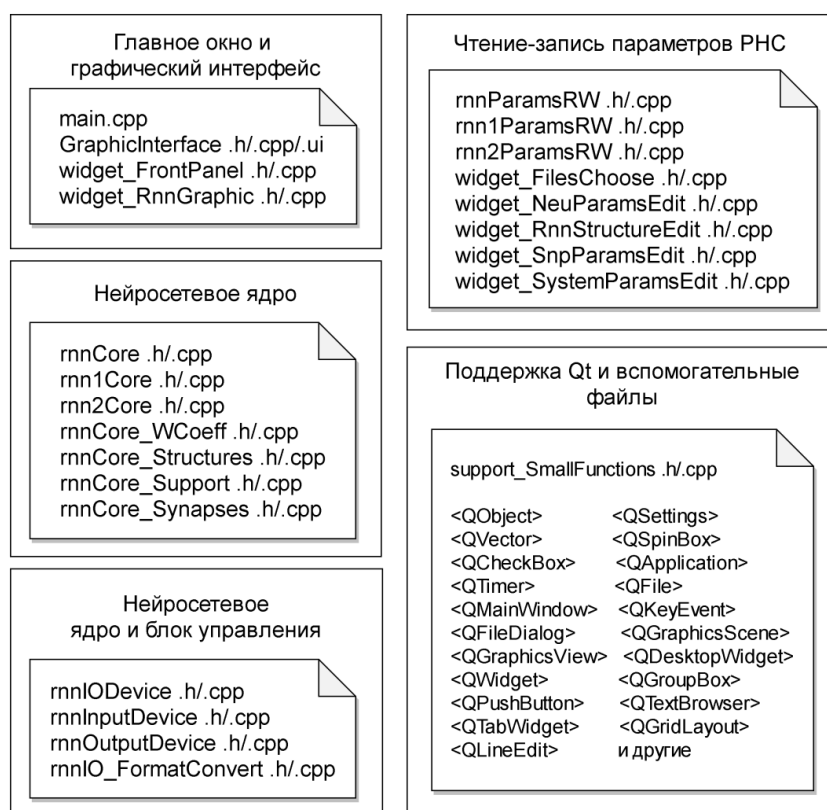


Рисунок 33 — Файловый состав программного обеспечения, реализующего методы нейросетевого прогнозирования с непрерывным обучением

Модули чтения/записи параметров №1 и №2 содержат основные настройки, связанные с функционированием РНС-1 и РНС-2 соответственно, а также общие настройки программной системы. Абстрактная логика, связанная с ними, описана в классе `rnnParamsRW`, а реализации – в классах `rnn1ParamsRW` и `rnn2ParamsRW`. Настройки редактируются через диалоговое окно, вызываемое соответствующей

кнопкой с панели графического интерфейса. Внешний вид диалогового окна настроек РНС-1 представлен на рисунке 34. Его элементы описаны в классах `widget_FilesChoose`, `widget_NeuParamsEdit`, `widget_RnnStructureEdit`, `widget_SnpParamsEdit`, `widget_SustemParamsEdit`.

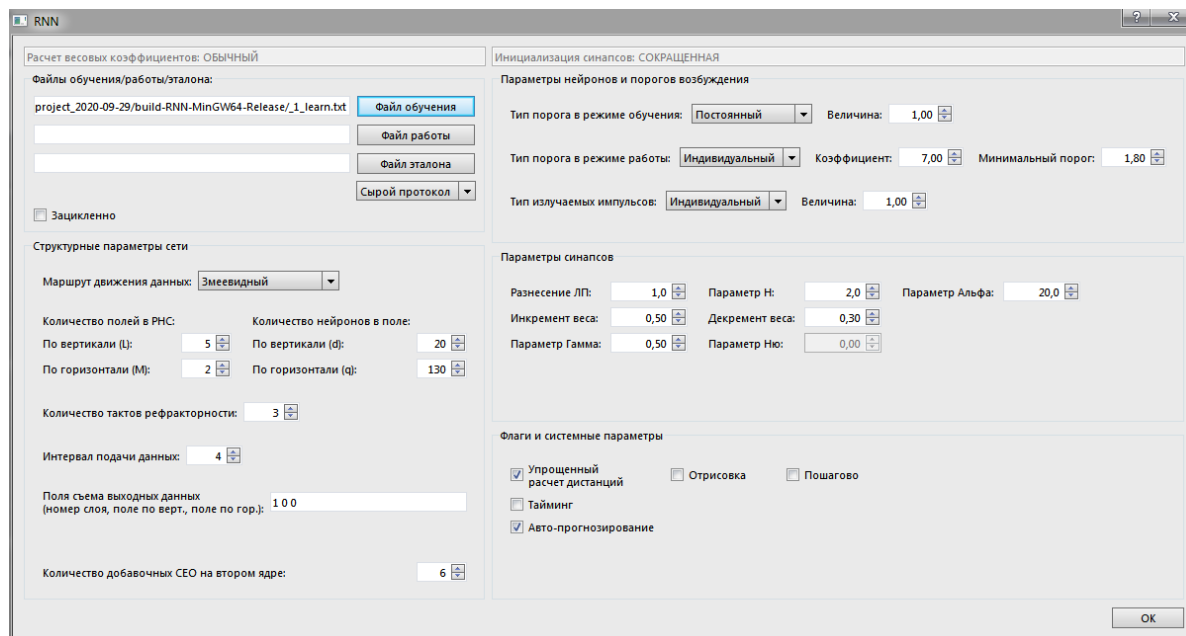


Рисунок 34 — Окно чтения/записи параметров РНС-1

Модуль предобработки входных данных включает в себя эмулятор устройства ввода, описанный в классе `rnnInputDevice`, а модуль постобработки – соответственно устройство вывода `rnnOutputDevice`. Данные классы наследуются от абстрактного класса ввода-вывода `rnnIODevice`, на базе которого могут быть реализованы устройства для различных типов входных данных. В существующем ПО реализовано устройство чтения данных из внешних файлов. Помимо устройства ввода, модули пред- и постобработки содержит функции, ответственные за конвертирование данных из файлов в формат РНС. Эти функции реализованы в классе `rnnIO_FormatConvert`.

Графический интерфейс, реализованный в классе `GraphicInterface`, отражен на рисунках 35-37. На этих скриншотах представлены состояния графического интерфейса в различных режимах работы программной системы. Графический интерфейс (рис. 35) включает в себя кнопочную панель с кнопками настройки параметров, запуска, остановки, очистки памяти РНС, команды на прогноз, а

также переключателя режимов функционирования РНС («обучение» – «прогнозирование»). Кнопочная панель реализована в классе widget_FrontPanel. Окно для служебных сообщений информирует пользователя о статусе процессов, протекающих внутри системы, сообщает об ошибках в случае их возникновения.

Графический модуль отображает нейросетевые слои. Кликнув на соответствующую вкладку, можно получить визуальное отображение состояния первого или второго слоя РНС. Реализация графического модуля находится в классе RnnGraphic.

При выполнении прогноза на интерфейсе отображаются графические элементы, связанные со вторым экземпляром РНС (рис. 36). На рисунке 37 продемонстрировано состояние графического интерфейса во время прогнозирования. Видно, что на слоях РНС-1 и РНС-2 осуществляется обработка данных, закодированных в виде совокупностей возбужденных нейронов.

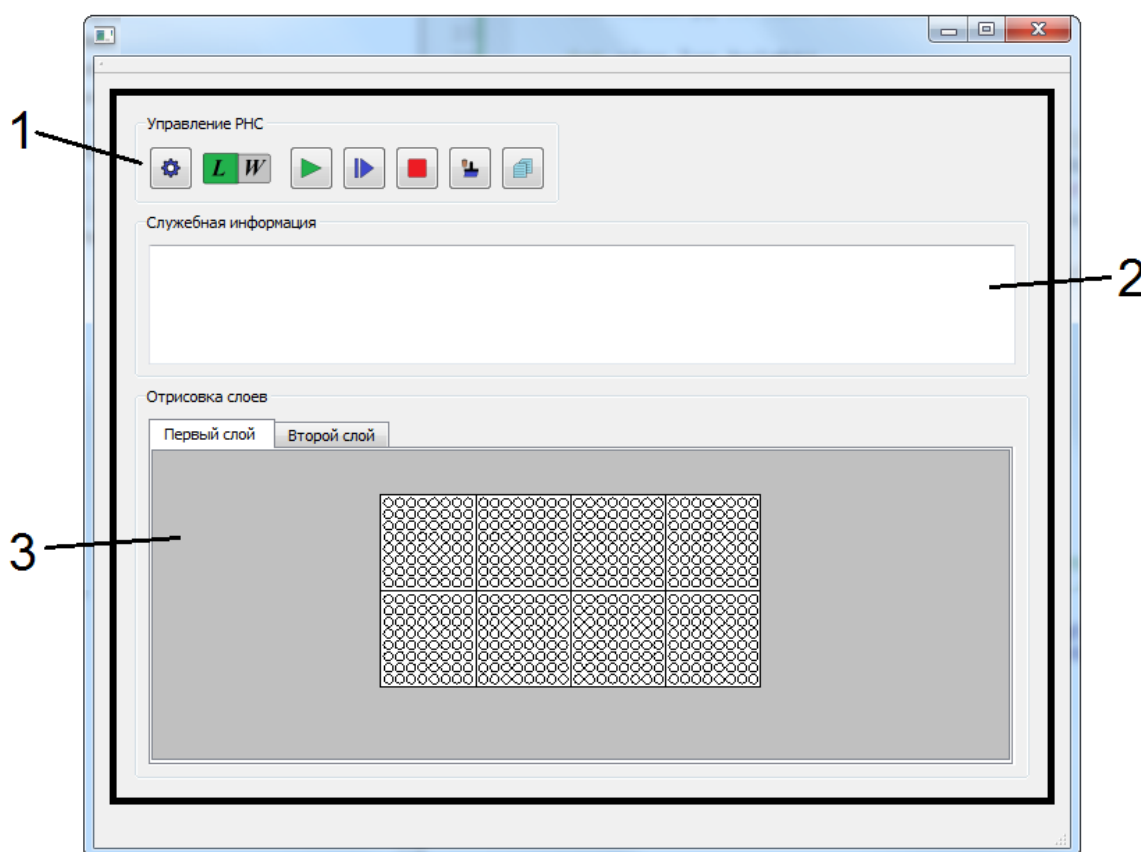


Рисунок 35 — Интерфейс программной системы нейросетевого прогнозирования (в режиме обучения). Здесь 1 – кнопочная панель, 2 – окно для системных сообщений, 3 – графический модуль для отрисовки слоев РНС

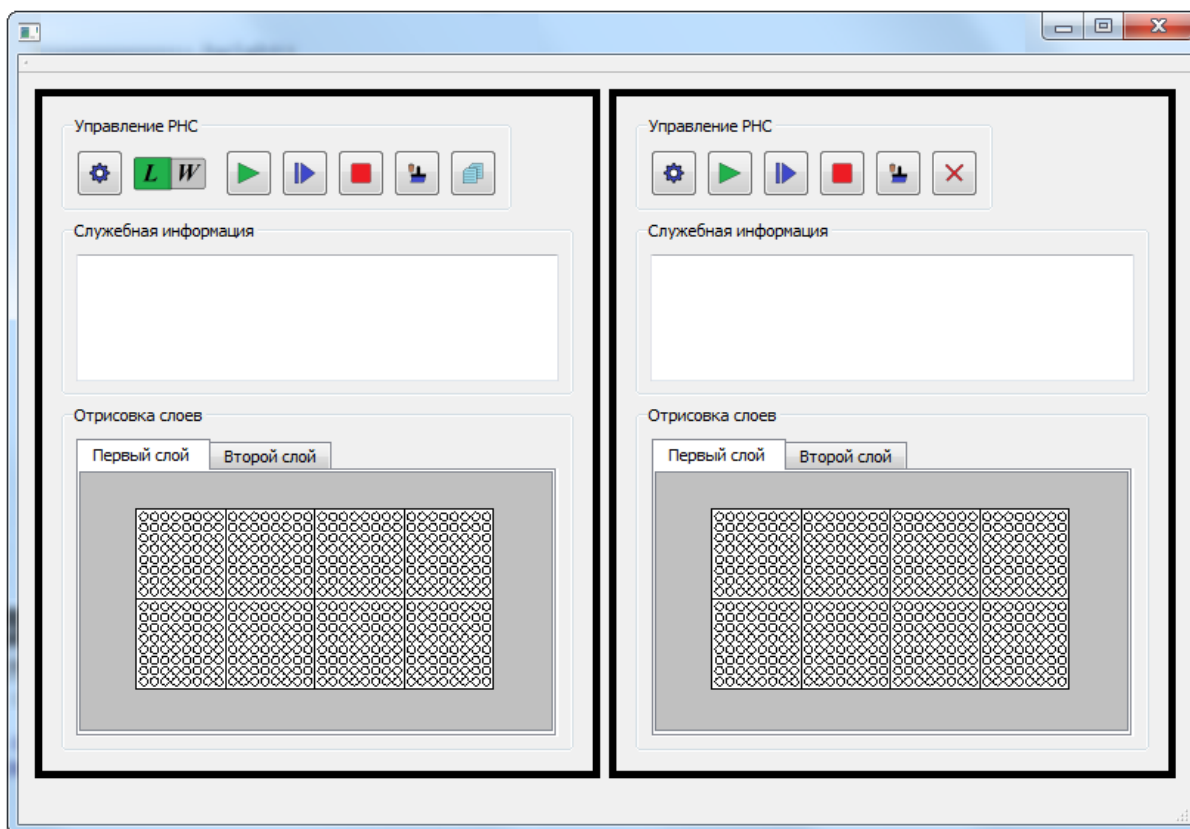


Рисунок 36 — Интерфейс программной системы нейросетевого прогнозирования (в режиме обучения и прогнозирования). Слева – панель РНС-1, справа – РНС-2

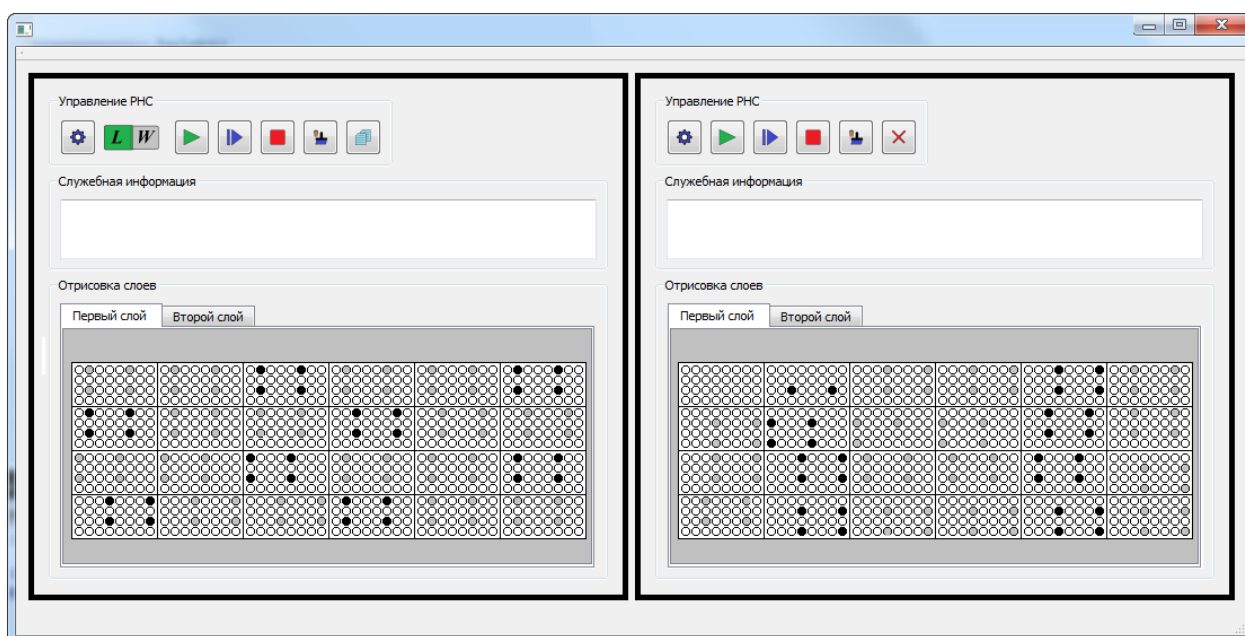


Рисунок 37 — Функционирование программной системы нейросетевого прогнозирования с непрерывным обучением. На слоях РНС белыми кружками обозначены нейроны в состоянии ожидания, серыми – рефракторные нейроны, а черными кружками – возбужденные нейроны

Графический интерфейс, РНС-1 и РНС-2 запущены в трех разных потоках, за счет чего обеспечивается, во-первых, отсутствие «зависаний» интерфейса при выполнении процессов в экземплярах РНС, а во-вторых, независимость и параллельность работы экземпляров РНС. Взаимодействие между потоками обеспечивают сигнально-слотовые соединения.

4.6 Рекомендации по взаимодействию программной системы нейросетевого прогнозирования с другими системами

Предложенные методы и архитектуры могут найти место в программном обеспечении систем, которым требуется получение прогнозов трудноформализуемых событий. Рассмотрим взаимодействие программной системы нейросетевого прогнозирования с другими системами на примере задачи прогнозирования пробок в интересах беспилотного транспортного средства.

Перспективные БТС относятся к классу интеллектуальных автономных роботов и наделены возможностью самостоятельного планирования действий. За эти операции ответственен навигационный модуль. В интересах более эффективного планирования пути в состав этого модуля может быть установлен блок прогнозирования с непрерывным обучением, воплощающий разработанные программные системы (рис. 38). Этот подход является расширением схемы навигационной системы интеллектуального автономного робота [124].

Функционирование навигационного модуля, в состав которого входит блок прогнозирования с непрерывным обучением, осуществляется следующим образом.

На вход модуля навигации из внешней среды поступает информация с различных датчиков и сенсоров, а также информация с внешних интерфейсов. БТС может в реальном времени подключаться посредством радиоканала к информационному сервису, который представляет информацию о загруженности дорог, объеме трафика и средних скоростях движения по городским дорогам, наличие аварий, ремонтных работ, погодных условий и другую информацию, влияющую на состояние транспортных магистралей. Эта информация передается

в блок прогнозирования с непрерывным обучением, и он обучается и формирует в себе пространственно-временную модель событий.

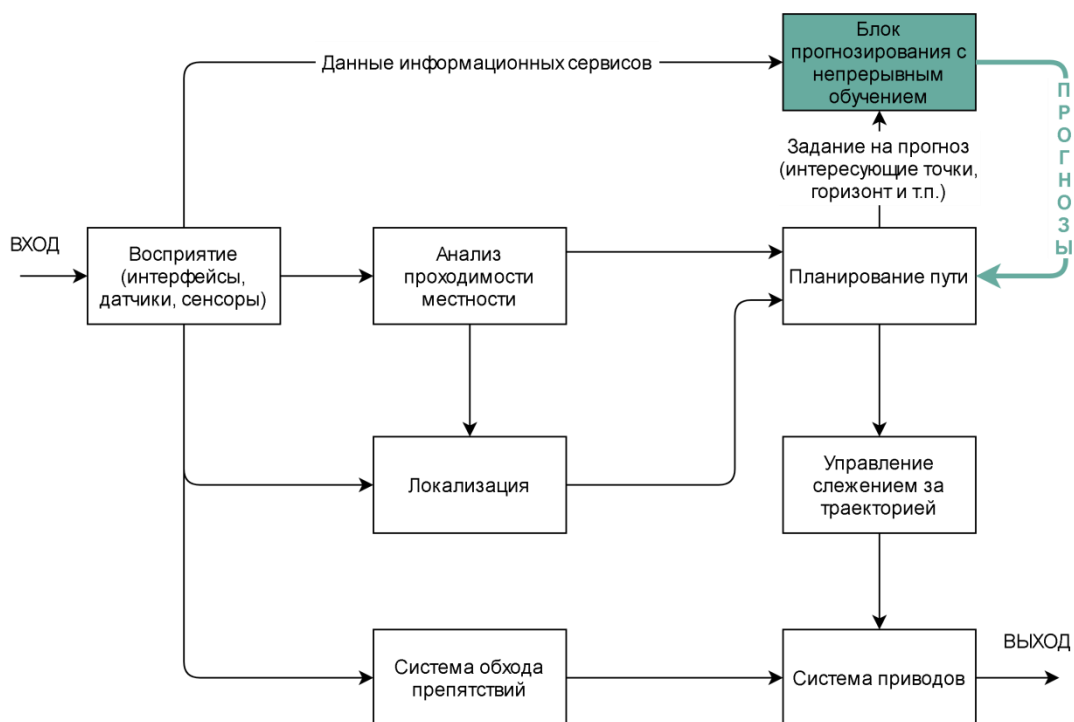


Рисунок 38 — Место системы прогнозирования с непрерывным обучением в навигационном модуле интеллектуального автономного робота класса «беспилотное транспортное средство»

Когда система прогнозирования обучена, блок планирования пути начинает запрашивать прогнозы транспортной ситуации на указанный горизонт и в указанных точках. Блок планирования пути может запрашивать прогнозы при построении маршрута и/или при его обновлении. За счет наличия механизмов непрерывного обучения блок прогнозирования в состоянии осуществлять корректировку прогнозов в реальном времени, учитывать новые данные для корректировки или перестроения выбранного маршрута.

Предлагаемая схема может быть применена для повышения точности прогнозов в беспилотных транспортных средствах, предназначенных для работы в городской среде.

Выводы по главе 4

Осуществлено моделирование разработанных моделей и методов прогнозирования трудноформализуемых событий применительно к задачам прогнозирования транспортных потоков и лексического содержания новостных лент. Представлены результаты экспериментов по оценке эффективности предлагаемых методов прогнозирования и программных архитектур, их реализующих. Для этого использовалось 4 набора данных. Проведено несколько серий экспериментов, направленных на определение показателей точности прогнозирования предлагаемых методов и моделей и выработку рекомендаций по использованию программных методов и архитектур. Результаты прогнозирования показывают, что предлагаемые методы и модели превышают по точности прогнозирования известные аналоги. Разработанные архитектуры могут немного уступать известным подходам по времени, затрачиваемому на прогнозирование, но выигрывают за счет отсутствия необходимости в переобучении системы, занимающем на порядки больше времени.

Продемонстрирован файловый состав программной системы нейросетевого прогнозирования, реализующей предложенные методы и архитектуры. Раскрыты особенности функционирования ПО, а также представлены графические интерфейсы программной системы на разных этапах работы.

Выработаны практические рекомендации по повышению точности и использованию программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, обеспечивающие повышение точности прогнозов за счет определения наиболее эффективной конфигурации нейросетевых слоев применительно к задаче прогнозирования трудноформализуемых событий и разработки новых правил выбора метода и архитектуры в зависимости от условий, в которых функционирует система прогнозирования. Эти рекомендации формулируются следующим перечнем:

1. Рекомендация по выбору метода прогнозирования: когда точность прогнозирования имеет решающую роль, необходимо использовать метод

прогнозирования с временными сдвигами. Если важнее экономия ресурсов, то предпочтительно использовать метод без временных сдвигов.

2. Рекомендация по выбору архитектуры: при существенных ограничениях на объем памяти, выделяемый системе прогнозирования, необходимо использовать буферную архитектуру, в противных случаях используется архитектура параллельная.

3. Рекомендация по выбору конфигураций слоев РНС: при прогнозировании трудноформализуемых событий предпочтительно использовать конфигурацию слоев РНС в виде спирали.

Представлен обобщенный пример, демонстрирующий возможности использования блока прогнозирования с непрерывным обучением в составе навигационной системы автономного робота. Определен порядок и особенности взаимодействия разработанной системы с другими подсистемами интеллектуальной машины.

В целом, в главе 4 разработаны новые научно обоснованные рекомендации по повышению точности прогнозов и использованию предложенных программных систем. В рамках этих рекомендаций определена эффективность различных конфигураций слоев РНС применительно к задаче прогнозирования событий в различных приложениях с высокой степенью неопределенности анализируемых процессов. Оценены затраты аппаратных ресурсов и скорость прогнозирования с применением различных методов и программных архитектур. Предложены правила выбора метода и архитектуры в зависимости от требований к системе прогнозирования и условий, в которых она действует.

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена задача разработки моделей, методов и архитектур программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, поставленная цель повышения точности получаемых прогнозов трудноформализуемых событий достигнута. Решенная задача имеет важное значение для совершенствования моделей, методов и средств прогнозирования событий в условиях слабо формализуемых процессов с учетом большого числа неявно связанных факторов.

Основные научные результаты, составляющие **итоги** исследования:

1. Предложена модель системы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, отличающаяся своей структурой и правилами обработки сигналов, обеспечивающими оперативное прогнозирование с учетом изменений в законах проявления событий.

Модель содержит блок управления прогнозированием и две идентичные по своей структуре рекуррентные нейронные сети (РНС-1 и РНС-2), объединенные в систему. РНС-1 работает в режиме обучения, блок управления прогнозированием выполняет копирование обученной пространственно-временной модели событий из РНС-1 в РНС-2, а РНС-2 реализует прогнозирование. Предложенная модель обеспечивает непрерывность процесса обучения при прогнозировании. Это позволяет обеспечить работу в реальном времени и возможность постоянного формирования прогнозов с учетом изменяющихся законов поведения временных рядов. Отсутствует необходимость переобучения сети при поступлении новых данных. Исключается искажение пространственно-временной модели РНС из-за смены режимов ее функционирования.

2. Предложены методы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением: с временными сдвигами сигналов и без временных сдвигов сигналов, отличающиеся новыми правилами прогнозирования и управления ассоциативным вызовом информации

из нейросетевой памяти и обеспечивающие высокую точность получаемых прогнозов трудноформализуемых событий.

Согласно методу с временными сдвигами, на вход РНС-1 подаются текущий и задержанный временные ряды. В процессе их прохождения вдоль слоев сети осуществляется ассоциативное пространственно-временное связывание прошлых и будущих событий. Блок управления прогнозированием копирует обученную модель из РНС-1 в РНС-2 и подает текущие входные данные в задержанный канал. В результате в текущем канале за счет вызова сигналов из ассоциативной памяти формируется прогноз будущих событий.

В методе прогнозирования без временных сдвигов на вход РНС-1 подается текущий временной ряд. При прохождении его по сети на ее элементах формируется модель событий, которая постоянно обновляется с учетом вновь поступающих данных. Блок управления прогнозированием копирует состояние РНС-1 в РНС-2 и запускает РНС-2 на формирование прогнозов по новым правилам, предусматривающим управление направленностью ассоциативного вызова сигналов из памяти нейронной сети. Согласно этим правилам, если обрабатываемая выборка признается короткой, то перед прогнозированием предлагается удлинять ее за счет ассоциативного вызова из памяти сети предшествующих значений.

3. Разработаны параллельная и буферная архитектуры программных систем, отличающиеся новой структурой и правилами функционирования программных систем прогнозирования с непрерывным обучением, обеспечивающие программную реализацию предложенных моделей и методов и расширение их функций.

В параллельной архитектуре эмулируется оба экземпляра нейронных сетей (РНС-1 и РНС-2). Буферная архитектура предполагает наличие только одного экземпляра (модуля) нейронной сети (РНС-1), называемого модулем эмуляции РНС-1 и РНС-2, а также входного буфера и модуля памяти для хранения состояний нейронов РНС-2. Новизна буферной архитектуры состоит в отказе от выделения памяти для хранения синапсов РНС-2, в выполнении

квазипараллельного обучения и прогнозирования, а достигаемый эффект заключается в сокращении объемов требуемой памяти в общем случае в два раза.

4. Разработаны практические рекомендации по повышению точности и использованию программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением, обеспечивающие повышение точности прогнозов за счет определения наиболее эффективной конфигурации нейросетевых слоев применительно к задаче прогнозирования трудноформализуемых событий и разработки новых правил выбора метода и архитектуры в зависимости от условий, в которых функционирует система прогнозирования. Разработанные методы и модели позволяют выполнять оперативные и точные прогнозы, устойчивые к зашумленности входных данных и/или их недостатку, учитывающие специфику и условия прогнозирования.

Результаты моделирования показывают, что предлагаемые методы превосходят по точности прогнозирования известные аналоги. Так, применительно к задаче прогнозирования средних скоростей движения по городским дорогам в сравнении с моделями интернет-сервиса «Яндекс.Пробки», ARIMA и LSTM получено преимущество более чем на 10% по показателю MAPE. При прогнозировании лексического содержания новостных лент как на 1-часовом, так и на 4-часовом горизонтах предлагаемый метод превосходит архитектуру LSTM. Так, для словаря из 306 слов общая процентная ошибка типа «0» и типа «1» для LSTM выше на 12.2%, а MAPE выше на 7.6%. Для словаря из 1000 слов и 1-часового горизонта прогноза эти показатели составляют 20.0% и 4.0% соответственно, а для 4-часового горизонта они составляют 10.4% и 2.9%, соответственно.

Разработанные архитектуры могут немного уступать известным подходам по времени, затрачиваемому на прогнозирование, но выигрывают за счет отсутствия необходимости в переобучении системы. В известных методах полный цикл переобучения занимал от 19229 до 28674 мс [83]. Предлагаемый подход отличается гибкостью и может быть настроен для работы с различными задачами и возможностями.

Даны **рекомендации** по использованию результатов исследования для прогнозирования событий в приложениях, требующих прогнозирования временных рядов со сложной динамикой и влиянием большого числа неявно связанных факторов. Они могут быть использованы в перспективных НИР и ОКР, а также в учебном процессе.

В качестве **перспектив дальнейшей разработки темы** можно указать дальнейшее повышение точности прогнозов за счет совершенствования правил управления ассоциативным вызовом из памяти нейронных сетей, внедрение подходов активного обучения, увеличение скорости нейросетевой обработки информации за счет реализации алгоритмов на графических процессорах.

Полученные результаты соответствуют паспорту специальности 2.3.5 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

СИМ	Системы и машины
РНС	Рекуррентная нейронная сеть
БПЛА	Беспилотный летательный аппарат
ARIMA	Autoregressive integrated moving average
ИНС	Искусственная нейронная сеть
CPU	Central processing unit
GPU	Graphics processing unit
НС	Нейронная сеть
РБФ	Радиальная базисная функция
LSTM	Long short-term memory
ВАМ	Bidirectional associative memory
ТСН	Temporal Convolutional Network
МСП	Многослойный перцептрон
GRU	Gated recurrent unit
TDNN	Time delay neural network
CNN	Convolutional neural network
SVM	Support vector machine
БТС	Беспилотное транспортное средство
СЕО	Совокупность единичных образов
ПСЕО	Последовательность СЕО
ДС	Динамический синапс
ЕЗ	Единичная задержка
ПВМ	Пространственно-временная модель
ПО	Программное обеспечение

СПИСОК ЛИТЕРАТУРЫ

1. Tealab, A. Time series forecasting using artificial neural networks methodologies: A systematic review / A. Tealab // *Future Computing and Informatics Journal*. – 2018. – Vol. 3. – P. 334-340.
2. Sezer, O. B. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019 / O. B. Sezer, M. U. Gudelek, A. M. Ozbayoglu // *Applied Soft Computing*. – 2020. – Vol. 90. – P. 106181.
3. Osipov, V. Neural network forecasting of news feeds / V. Osipov, S. Kuleshov, A. Zaytseva, D. Levonevskiy, D. Miloserdov // *Expert Systems with Applications* – 2020. – Vol. 169.
4. SAE – международная ассоциация автомобильных инженеров [Электронный ресурс]. – URL: <https://www.sae.org/> (дата обращения: 12.05.2021).
5. «КамАЗ» начал тестирование беспилотника [Электронный ресурс]. – URL: https://kamaz.ru/press/news/kamaz_nachal_testirovanie_bespilotnika/ (дата обращения: 13.05.2021).
6. Терентьев В.В. Внедрение интеллектуальных систем на автомобильном транспорте / В.В. Терентьев // *НИКСС*. – 2018. – №1(21). – С. 117-122.
7. Бабиченко, А. В. Интеллектуальные системы управления беспилотными летательными аппаратами / А. В. Бабиченко, В. А. Пролетарская, М. С. Селезнева, К. А. Неусыпин, А. В. Пролетарский // *Инженерная физика*. – 2018. – № 5. – С. 94-102.
8. Mahmoud, A. A survey on deep learning for time-series forecasting / A. Mahmoud, M. Ammar // *Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges*. – 2021. – P. 365-392.
9. Pascanu, R. On the difficulty of training recurrent neural networks / R. Pascanu, T. Mikolov, Y. Bengio // *30th International Conference on Machine Learning, ICML 2013*. – 2013. – Vol. 28. – P. 1310-1318.
10. Yang, B. Traffic flow prediction using LSTM with feature enhancement / B. Yang, S. Sun, J. Li, X. Lin, Y. Tian // *Neurocomputing*. – 2019. – Vol. 332. – P. 320–327.

11. Osipov, V. Neural Network Forecasting of Traffic Congestion / V. Osipov, D. Miloserdov // Digital Transformation and Global Society, DTGS 2019. – In Communications in Computer and Information Science. – 2019. – Vol. 1038.
12. Osipov, V. Neural Network Associative Forecasting of Demand for Goods / V. Osipov, N. Zhukova, D. Miloserdov // Experimental Economics and Machine Learning, EEML 2019. – 2019. – Vol. 2479.
13. Милосердов, Д. И. Программный комплекс нейросетевого прогнозирования временных рядов / Д.И. Милосердов // 5-я Международная научная конференция «Технологическая перспектива в рамках евразийского пространства: новые рынки и точки экономического роста». – 2019. – С. 166-169.
14. Милосердов Д. И. Нейросетевое прогнозирование событий для интеллектуальных роботов с непрерывным обучением // Технологические тренды и наукоемкая экономика: бизнес, отрасли, регионы. Раздел 1: Цифровые технологии наукоемкой экономики. – 2021. – С. 27-37.
15. Милосердов И.В., Милосердов Д.И. Разработка механизмов обеспечения устойчивости в нейросетевых системах прогнозирования (Материалы конференции «Информационные технологии в управлении», 2020 г.) [Электронный ресурс]. – URL: <https://itc.etu.ru/assets/files/itc-2020/papers/198.pdf> (дата обращения: 13.05.2021)
16. Hewamalage, H. Recurrent Neural Networks for Time Series Forecasting: Current Status and Future Directions / H. Hewamalage, S. Bergmeir, K. Bandara // 2019.
17. Time-series forecasting with deep learning: a survey [Электронный ресурс]. – URL: <https://royalsocietypublishing.org/doi/full/10.1098/rsta.2020.0209> (дата обращения: 28.10.2021).
18. Видмант О. С. Прогнозирование финансовых временных рядов с использованием рекуррентных нейронных сетей LSTM / О. С. Видмант // Общество: политика, экономика, право. – 2018. – № 5.
19. Осипов, В. Ю. Нейросетевое прогнозирование событий для интеллектуальных роботов / В. Ю. Осипов // Мехатроника, автоматизация, управление. – 2015. – Т. 16. – № 12. – С. 836-840.

20. International Federation of Robotics. Industrial Robotics Standardization. [Электронный ресурс]. – URL: <http://www.ifr.org/news/ifr-press-release/iso-robotics-standardisation35/> (дата обращения: 21.12.2020).
21. Wang, T. Current Researches and Future Development Trend of Intelligent Robot: A Review / T. Wang, Y. Tao, H. Liu // International Journal of Automation and Computing. – 2018. – Vol. 15. – P. 525–546 (2018).
22. Савин, С.И. Использование нейронных сетей для прогнозирования нормальных реакций шагающего робота. / С.И. Савин, Л.Ю. Ворочаева // Известия Юго-Западного государственного университета. – 2019. – Т. 23(4). – С. 8-18.
23. Савин, С.И. Способ реализации адаптации робота к изменениям режима контактного взаимодействия с использованием глубоких полносвязных нейронных сетей. / С.И. Савин, Л.Ю. Ворочаева, А.В. Мальчиков, А.М. Салихзянов, Э.М. Заляев // Известия Юго-Западного государственного университета. – 2020. – Т. 24(1). – С. 206-214.
24. Hamza, A. Forecasting battery state of charge for robot missions / A. Hamza, N. Ayanian // Proceedings of the Symposium on Applied Computing, SAC 2017. – In Association for Computing Machinery. – 2017. – P. 249–255.
25. Thurow, K. Transportation robot battery power forecasting based on bidirectional deep-learning method / K. Thurow, C. Chen, S. Junginger, N. Stoll, H. Liu // Transportation Safety and Environment. – 2019. – Vol. 1, Issue 3. – P. 205–211.
26. Diane, S. Forecasting conflicts in multi-robot systems based on intelligent feedback / S. Diane, S. Manko, V. Lokhin // International Siberian Conference on Control and Communications, SIBCON-2016. – 2016. – P. 1-5.
27. Wong, C. Fuzzy tracking method with a switching grey prediction for mobile robot / C. Wong, B. Lin, C. Cheng // 10th IEEE International Conference on Fuzzy Systems. – 2001. – Vol. 1. – P.103-106.
28. Маленков, М. И. Новые проектно-компоновочные решения для повышения подвижности и функциональных возможностей планетоходов / М. И. Маленков и др. // Известия ЮФУ. Технические науки. – 2017. – № 1(186). – С. 42–54.

29. Malenkov, M.I. Increasing the Mobility of Mars Rovers by Improving the Locomotion Systems and Their Control Algorithms / M.I. Malenkov, V.A. Volov, N.K. Guseva, E.A. Lazarev // Russian Engineering Research. Allerton Press, Inc. – 2015. – Vol. 35. – No. 11. – P. 824-831.
30. Gonzalez, R. Slippage prediction for off-road mobile robots via machine learning regression and proprioceptive sensing / R. Gonzalez, M. Fiacchini, K. Iagnemma // Robotics and Autonomous Systems. – 2018. – Vol. 105. – P. 85-93.
31. Ebrahimi, S. A New Sensor for Robotic Mars Rovers in Sandy Terrains Predicting Critical Soil Flow Using the Spiral Soil Flow Model. / S. Ebrahimi, A. Mardani, K. Alipour // Robotica. – 2021. – Vol. 39(2). – P. 346-365.
32. Zhou, R. Predicting Terrain Mechanical Properties in Sight for Planetary Rovers with Semantic Clues. / R. Zhou, W. Feng, H. Yang, H. Gao, N. Li, Z. Deng, L. Ding // 2020.
33. Cunningham, C. Improving slip prediction on Mars using thermal inertia measurements / C. Cunningham, I.A. Nesnas, W.L. Whittaker // Autonomous Robots. – 2019. – Vol. 43. – P. 503–521.
34. Inotsume, H. Slip prediction for exploration rover based on transfer learning / H. Inotsume, T. Kubota // International Symposium on Artificial Intelligence, Robotics and Automation in Space, i-SAIRAS 2020. – 2020.
35. Lopez-Arreguin, A.J.R. Do lunar rover wheels sink equally on Earth and Moon? / A.J.R. Lopez-Arreguin, B. Gundlach, E. Stoll // Results in Physics. – 2019. – Vol. 15.
36. Kamil, F. New robot navigation algorithm for arbitrary unknown dynamic environments based on future prediction and priority behavior / F. Kamil, T. S. Hong, W. Khaksar, M. Y. Moghrabiah, N. Zulkifli, S. A. Ahmad // Expert Systems with Applications. – 2017. – Vol. 86. – P. 274-291.
37. Katyal, K. Uncertainty-Aware Occupancy Map Prediction Using Generative Networks for Robot Navigation / K. Katyal, K. Popek, C. Paxton, P. Burlina, G. D. Hager // International Conference on Robotics and Automation, ICRA 2019. – 2019. – P. 5453-5459.

38. Victerpaul, P. Path planning of autonomous mobile robots: A survey and comparison / P. Victerpaul, S. Devaraj, J. Subbiah, P. Jayabala // *Journal of Advanced Research in Dynamical and Control Systems*. – 2017. – Vol. 9.
39. Solea, R. Online path planner for mobile robots using particle swarm optimization / R. Solea, D. Cernega // *20th International Conference on System Theory, Control and Computing, ICSTCC-2016*. – 2016. – P. 222–227.
40. Tusi, Y. Using ABC and RRT algorithms to improve mobile robot path planning with danger degree / Y. Tusi, H.-Y. Chung // *Fifth International Conference on Future Communication Technologies, FGCT-2016*. – 2016. – P. 21–26.
41. Буслаев, С. П. Общая методология совместного применения операторного и автоматического управления движением лунохода в «ночных» кратерах на полюсах Луны / С. П. Буслаев, В. А. Воронцов, О. С. Графодатский, А. М. Крайнов // *Инженерный журнал: наука и инновации*. – 2017. – № 8(68).
42. Byrne, C. J. *Travels with Curiosity. Exploring Mars by Rover* / C. J. Byrne // Springer International Publishing, Springer Nature Switzerland AG. – 2020. – 217 p.
43. Покусаев О.Н. Мировой рынок автономных (беспилотных) автомобилей / О.Н. Покусаев, А.С. Мишарин, В.П. Куприяновский, А.А. Климов // *Digital Transformation of Transport*. – 2018. – Vol. 14, no 3.
44. Barbaglia, L. and Consoli, S. and Manzan, S. Forecasting with Economic News [Электронный ресурс]. – URL: <https://ssrn.com/abstract=3698121> (дата обращения: 19.12.2021)
45. Park, E. Tourism demand forecasting with online news data mining / E. Park, J. Park, Mingming Hu // *Annals of Tourism Research*. – 2021. – Vol. 90. – P. 103273.
46. Остапенко, А. Г. Программное обеспечение для мониторинга процессов восприятия и распространения деструктивных контентов в социальных сетях / А. Г. Остапенко, Е. Ю. Чапурин, Е. С. Соколова, А. Г. Зимницкий, И. А. Боков, С. В. Лихобабин, А. О. Ткаченко, А. И. Дегтярев // *Информация и безопасность*. – 2019. – Т. 22, № 2. – С. 188-205.
47. Mars Curiosity Rover [Электронный ресурс]. – URL: <http://mars.jpl.nasa.gov/msl/> (дата обращения: 19.05.2021)

48. Кудрина, Л.В. Лексические особенности новостных текстов интернет-изданий / Л.В. Кудрина // Пушкинские чтения. Художественные стратегии классической и новой словесности: жанр, автор, текст. – 2018. – С. 399-408.
49. Ping, T. Studying on path planning and dynamic obstacle avoiding of soccer robot / T. Ping, Z. Qi, Y. Y. Min // Proceedings of the 3rd World Congress on Intelligent Control and Automation (Cat. No.00EX393). – 2000. – Vol. 2. – P. 1244-1247.
50. Brinkmann, W. Space Robotic Systems and Artificial Intelligence in the Context of the European Space Technology Roadmap / W. Brinkmann, F. Cordes, C. Koch, M. Wirkus, R. Dominguez, A. Dettmann, T. Vögele, F. Kirchner // Space Tech and Smallsats Conferences. – 2019.
51. Шихамова, Р. Р. Лингвостилистические особенности французских новостных текстов в сфере образования / Р. Р. Шихамова, Н. В. Тихонова // Terra linguae. – 2020. – С. 125-130.
52. Процессор марсохода Perseverance стоит 200 тысяч долларов, но слабее наручных часов [Электронный ресурс]. – URL: <https://hi-news-ru.turbopages.org/hi-news.ru/s/technology/processor-marsoxoda-perseverance-stoit-200-tysyach-dollarov-no-slabee-naruchnyx-chasov.html> (дата обращения: 13.05.2021)
53. Playter, R. BigDog / R. Playter, M. Buehler, M. Raibert // Proc. SPIE 6230, Unmanned Systems Technology VIII, 62302O. – 2006.
54. Boubin, J. Poster Abstract: Characterizing Computational Workloads in UAV Applications / J. Boubin, S. Zhang, V. Mandadapu, C. Stewart, // IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation, IoTDI 2018. – 2018. P. 275-276.
55. Xing, F.Z. Natural language based financial forecasting: a survey / F.Z. Xing, E. Cambria, R.E. Welsch // Artificial Intelligence Review. – 2018. – Vol. 50. – P. 49–73.
56. Wang, H. Hat: Hardware-aware transformers for efficient natural language processing / H. Wang et al. // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. – 2020.
57. Afshar, M. Price Development and application of a high throughput natural language processing architecture to convert all clinical documents in a clinical data

warehouse into standardized medical vocabularies / M. Afshar, D. Dligach, B. Sharma, X. Cai, J. Boyda, S. Birch, D. Valdez, S. Zelisko, C. Joyce, F. Modave, R. Price // Journal of the American Medical Informatics Association. – 2019. – Vol. 26, Issue 11. – P. 1364–1369.

58. Findi, A.H.M. Collision Prediction Based Genetic Network Programming-Reinforcement Learning for Mobile Robot Navigation in Unknown Dynamic Environments / A.H.M. Findi, H. Mohammad, R. K. Marhaban, K. H. Mohd // Journal of Electrical Engineering and Technology. – 2017. – Vol. 12. – No. 2. – P. 890–903.

59. Rösmann, C. Online trajectory prediction and planning for social robot navigation / C. Rösmann, M. Oeljeklaus, F. Hoffmann, T. Bertram // IEEE International Conference on Advanced Intelligent Mechatronics, AIM 2017. – 2017. P. 1255-1260.

60. Swapna, R. H., Soft Computing in Robotics: A Decade Perspective / R. H. Swapna, J. Nayak, B. Naik, D. Pelusi // Applications of Robotics in Industry Using Advanced Mechanisms, ARIAM 2019. – In Learning and Analytics in Intelligent Systems. – 2020. – Vol. 5.

61. Бестужев-Лада, И.В. Рабочая книга по прогнозированию / И.В. Бестужев-Лада. // М.: МЫСЛЬ. – 1982. – 426 с.

62. Palit, A. K. Computational Intelligence in Time Series Forecasting. Theory and Engineering Applications / A. K. Palit, D. Popovic // Springer-Verlag London. – 2005. – 372 p.

63. Timmermann, A. Forecasting methods in finance / A. Timmermann // Annual Review of Financial Economics. – 2018. – Vol. 10(1). – P. 449-479.

64. Brockwell, P. Introduction to Time Series and Forecasting / P. Brockwell, R. Davis // Springer International Publishing. – 2016. – 425 p.

65. Nagy, A. Survey on traffic prediction in smart cities. / A. Nagy, V. Simon // Pervasive and Mobile Computing. – 2018. – Vol. 50. – P. 148-163.

66. Bok, B. Macroeconomic Nowcasting and Forecasting with Big Data. / B. Bok, D. Caratelli, D. Giannone, A. Sbordone, A. Tambalotti // Annual Review of Economics. – 2018. – Vol. 10. – P. 615-643.

67. Livera, A. M. Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing / A. M. Livera, R. J. Hyndman, R. D. Snyder // *Journal of the American Statistical Association*. – 2011. – Vol. 106:496. – P. 1513-1527.
68. Кувайскова, Ю. Е. Статистические методы прогнозирования: учебное пособие / Ю. Е. Кувайскова, В. Н. Клячкин // Ульяновск : УлГТУ. – 2019. – 197 с.
69. Yarushev, S. Review of studies on time series forecasting based on hybrid methods, neural networks and multiple regression / S. Yarushev, A. Averkin // *Программные продукты и системы*. – 2016. – Т. 31. – С. 75-82.
70. Webby, R. Judgemental and statistical time series forecasting: a review of the literature / R. Webby, M. O'Connor // *International Journal of Forecasting*. – 1996. – Vol. 12, Issue 1. – P. 91-118.
71. Hyndman, R.J. Forecasting: principles and practice, 3rd edn. / R.J. Hyndman, G. Athanasopoulos // OTexts, Melbourne. – 2018.
72. Amini, M. H. ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation / M. H. Amini, A. Kargarian, O. Karabasoglu // *Electric Power Systems Research*. – 2016. – Vol. 140. – P. 378-390.
73. Слепых, А.А. Навигация на марсе / А.А. Слепых // М : МГТУ им. Н.Э. Баумана. – 2020. – 44 с.
74. Ono, M. MAARS: Machine learning-based Analytics for Automated Rover Systems / M. Ono et al. // *IEEE Aerospace Conference, Big Sky*. – 2020. P. 1-17.
75. Исмагилов, И.И. Нечеткие прогнозы: классификация и метод их разработки на основе процедуры групповой экспертизы / И.И. Исмагилов, Р.В. Бичурин // *Фундаментальные исследования*. – 2014. – № 11-6. – С. 1240-1247.
76. Ярушкина, Н. Г. Интеллектуальный анализ временных рядов: учебное пособие / Н. Г. Ярушкина, Т. В. Афанасьева, И. Г. Перфильева // Ульяновск: УлГТУ. – 2010. – 320 с.
77. Fajardo-Toro, C.H. Adaptive and Hybrid Forecasting Models—A Review / C.H. Fajardo-Toro, J. Mula, R. Poler // *Engineering Digital Transformation. Lecture Notes in Management and Industrial Engineering*. – 2019. – P. 315-322.

78. Malik, S. Heuristic and Statistical Prediction Algorithms Survey for Smart Environments / S. Malik, I. Ullah, D. Kim, K. Lee // *Journal of Information Processing Systems*. – 2020. – Vol. 16(5). – P. 1196-1213.
79. Chang, W.-Y. A Literature Review of Wind Forecasting Methods / W.-Y. Chang // *Journal of Power and Energy Engineering*. – 2014. – Vol. 2. – P. 161-168.
80. Lu, H. Advances in neural networks – ISNN 2019 / H. Lu, H. Tang, Z. Wang (eds) // *LNCS 11554*, Springer, Berlin. – 2019. – 483 p.
81. Saadaoui, F. Hybrid feedforward ANN with NLS-based regression curve fitting for US air traffic forecasting / F. Saadaoui, H. Saadaoui, H. Rabbouch // *Neural Computing and Applications*. – 2019. – Vol. 32(14). – P. 10073-10085.
82. Gairaa, K. Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: a combined approach / K. Gairaa, A. Khellaf, Y. Messlem, F. Chellali // *Renewable and Sustainable Energy Reviews*. – 2016. – Vol. 57. – P. 238-249.
83. Shen, Z. A novel learning method for multi-intersections aware traffic flow forecasting / Z. Shen, W. Wang, Q. Shen et al // *Neurocomputing*. – 2019. – Vol. 398(4).
84. Haykin, S. *Neural Networks and Learning Machines (3rd Edition)* / S. Haykin // McMaster University, Ontario. – 2009.
85. Гудфеллоу, Я., Бенджио И., Курвилль А. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль // М. : ДМК-Пресс. – 2018. – 652 с.
86. Redmon, J. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2016. – P. 779-788.
87. Sui, S. Neural Network Filtering Control Design for Nontriangular Structure Switched Nonlinear Systems in Finite Time / S. Sui, C. L. P. Chen, S. Tong // *IEEE Transactions on Neural Networks and Learning Systems*. – 2019. – Vol. 30, no. 7. – P. 2153-2162.
88. Yang, D. BM3D-Net: A Convolutional Neural Network for Transform-Domain Collaborative Filtering / D. Yang, J. Sun // *IEEE Signal Processing Letters*. – 2018. – Vol. 25, no. 1. – P. 55-59.

89. Furber, S.B. Overview of the SpiNNaker System Architecture / S.B. Furber et al. // IEEE Transactions on Computers. – 2013. - Vol. 62, no. 12. – P. 2454-2467
90. Furber, S.B. The SpiNNaker Project / S.B. Furber, F. Galluppi, S. Temple, L. A. Plana // Proceedings of the IEEE. – 2014. – Vol. 102, no. 5. – P. 652-665.
91. Introducing a Brain-inspired Computer. TrueNorth's neurons to revolutionize system architecture [Электронный ресурс]. – URL: <http://www.research.ibm.com/articles/brain-chip.shtml> (дата обращения: 19.05.2021)
92. Esser, S. K. Convolutional networks for fast, energy-efficient neuromorphic computing // S. K. Esser, P. A. Merolla, J. V. Arthur, etc. // Proceedings of the National Academy of Sciences. – 2016. – Vol. 113(41).
93. Koprinska, I. Convolutional Neural Networks for Energy Time Series Forecasting / I. Koprinska, D. Wu, Z. Wang // International Joint Conference on Neural Networks, IJCNN. – 2018. P. 1-8.
94. Pisarev, A. D. A biomorphic neuroprocessor based on a composite memristor-diode crossbar / A. D. Pisarev, A. N. Busygin, S. Yu. Udovichenko, O. V. Maevsky // Microelectronics Journal. – 2020. – Vol. 102.
95. Skorobogatchenko, D. A. Application of artificial neural networks with short-term prediction of the level of loading of urban automobile roads. / D. A. Skorobogatchenko, P. Yu. Stepanova // Izvestiya VolgGTU. – 2017. – No. 4 (209). – P. 24-30.
96. Zhou, T. δ -agree AdaBoost stacked autoencoder for short-term traffic flow forecasting / T. Zhou, G. Han, X. Xu, Z. Lin, C. Han, Y. Huang, J. Qin // Neurocomputing. – 2017. – Vol. 247. – P. 31–38.
97. Egrioglu, E. Median-Pi artificial neural network for forecasting / E. Egrioglu, U. Yolcu, E. Bas, A. Z. Dalar // Neural Computing and Applications. – 2019. – Vol. 31. – P. 307-316.
98. Wu, W. The Coal Demand Prediction Based on the Grey Neural Network Model / W. Wu, X. Wang // LISS 2014. – 2015.
99. Srinivasan, D. Energy demand prediction using GMDH networks / D. Srinivasan // Neurocomputing. – 2008. – Vol. 72(1–3). – P. 625-629.

100. Ghaderi, A. Deep Forecast: Deep Learning-based Spatio-Temporal Forecasting / A. Ghaderi, B. M. Sanandaji, F. Ghaderi // Proceedings of the ICML. – 2017. – P. 264-271.
101. Dai, X. DeepTrend: A Deep Hierarchical Neural Network for Traffic Flow Prediction. / X. Dai, R. Fu, Y. Lin, F. -Y Wang, L. Li // 20 th International Conference on Intelligent Transportation Systems (ITSC). – 2017. – P. 394-399.
102. Krichene, E. Forecasting Using Elman Recurrent Neural Network. / E. Krichene, Y. Masmoudi, A.M. Alimi, A. Abraham, H. Chabchoub // Intelligent Systems Design and Applications, ISDA 2016. – In Advances in Intelligent Systems and Computing. – 2017. – Vol. 557.
103. Goudarzi, S. Self-organizing traffic flow prediction with an optimized deep belief network for Internet of vehicles / S. Goudarzi, M. N. Kama, M. H. Anisi, S. A. Soleymani, F. Doctor // Sensors. – 2018. – Vol. 18. – P. 3459.
104. Fernandez-Navarro, M. Time series forecasting by recurrent product unit neural networks / M. Fernandez-Navarro, M. A. de la Cruz, P. A. Gutierrez, A. Castano, C. Hervás-Martínez // Neural Computing and Applications. – 2018. – Vol. 29. – P. 779-791.
105. Faloutsos, C. Forecasting Big Time Series: Theory and Practice / C. Faloutsos, V. Flunkert, J. Gasthaus, T. Januschowski, Y. Wang // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). – In Association for Computing Machinery. – 2019. – P. 3209–3210.
106. Wu, Y. A hybrid deep learning based traffic flow prediction method and its understanding / Y. Wu, H. Tan, L. Qin, B. Ran, Z. Jiang // Transportation Research Part C. Emerging Technologies. – 2018. – Vol. 90. – P. 166–180.
107. Polson, N. G. Deep learning for short-term traffic flow prediction / N. G. Polson, V. O. Sokolov // Transportation research Part C. – 2017. – Vol. 79. – P. 1-17.
108. Haitao, L. Research on prediction of traffic flow based on dynamic fuzzy neural networks / L. Haitao // Neural Computing and Applications. – 2016. – Vol. 27. – P. 1969-1980.

109. Sarıca, B. A new hybrid method for time series forecasting: AR–ANFIS / B. Sarıca, E. Egrioglu, B. Asıkgil // *Neural Computing and Applications*. – 2018. – Vol. 29. – P. 749-760.
110. Moretti, F. Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling / F. Moretti, S. Pizzuti, S. Panzieri, M. Annunziato // *Neurocomputing*. – 2015. – Vol. 167. – P. 3-7.
111. Kewat, P. Support vector machines through financial time series forecasting / P. Kewat, R. Sharma, U. Singh, R. Itare // *International conference of Electronics, Communication and Aerospace Technology, ICECA-2017*. – 2017. – P. 471-477.
112. Jaramillo, J. Research in Financial Time Series Forecasting with SVM: Contributions from Literature / J. Jaramillo, J. D. Velasquez, C. J. Franco // *IEEE Latin America Transactions*. – 2017. – Vol. 15, no. 1. – P. 145-153.
113. Yujun, Y. Research on financial time series forecasting based on SVM / Y. Yujun, Y. Yimei, L. Jianping // *13th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP-2016*. – 2016. – P. 346-349.
114. Herrera, I. Artificial Neural Networks for Bottled Water Demand Forecasting: A Small Business Case Study / I. Herrera, J. Chicaiza, E. Herrera, L. Lorente-Leyva, J. Caraguay-Procel, I. García-Santillán, D. Peluffo // *Advances in Computational Intelligence, IWANN 2019*. – In *LNCS*. – 2019. – P. 362-373.
115. Chawla, A. Demand Forecasting Using Artificial Neural Networks—A Case Study of American Retail Corporation / A. Chawla, A. Singh, A. Lamba, N. Gangwani, U. Soni // *Applications of Artificial Intelligence Techniques in Engineering*. – In *Advances in Intelligent Systems and Computing*. – 2019. – Vol. 697.
116. Galeshchuk, S. Neural networks performance in exchange rate prediction / S. Galeshchuk // *Neurocomputing*. – 2016. – Vol. 172. – P. 446-452.
117. Christopher, J. Convolutional Neural Network Deep-Learning Models for Prediction of Shared Bicycle Demand / J. Christopher, J. Mou, D. Yin // *International Conference on Applications and Techniques in Cyber Security and Intelligence, ATCI-2018*. – In *Advances in Intelligent Systems and Computing*. – 2019. – Vol. 842.

118. Qin, M. Red tide time series forecasting by combining ARIMA and deep belief network / M. Qin, Z. Li, Z. Du // Knowledge-Based Systems. – 2017. – Vol. 125. – P. 39-52
119. Hu, H. Predicting the direction of stock markets using optimized neural networks with Google Trends / H. Hu, L. Tang, S. Zhang, H. Wang // Neurocomputing. – 2018. – Vol. 285. – P. 188-195.
120. Essien A. A Deep Learning Model for Smart Manufacturing Using Convolutional LSTM Neural Network Autoencoders / A. Essien, C. Giannetti // IEEE Transactions on Industrial Informatics. – 2020. – Vol. 16, no. 9. – P. 6069-6078.
121. Awad, M. Forecasting of chaotic time series using RBF neural networks optimized by genetic algorithms / M. Awad // Int. Arab J. Inf. Technol. – 2017. – Vol. 14, no. 6. – P. 826-834.
122. Wang, L. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting / L. Wang, Y. Zeng, T. Chen // Expert Systems with Applications. – 2015. – Vol. 42, Issue 2. – P. 855-863.
123. Vadlamani, R. Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms / R. Vadlamani, P. Dadabada, D. Kalyanmoy // Swarm and Evolutionary Computation. – 2017. – Vol. 36. – P. 136-149.
124. Wong, C. Adaptive and intelligent navigation of autonomous planetary rovers – A survey / C. Wong, E. Yang, X. Yan, D. Gu // NASA/ESA Conference on Adaptive Hardware and Systems, AHS 2017. – 2017. P. 237-244.
125. Shi, X. Machine Learning for Spatiotemporal Sequence Forecasting: A Survey / X. Shi, D. Yeung // IEEE. – 2018.
126. Faloutsos, C. Forecasting big time series: old and new / C. Faloutsos, J. Gasthaus, T. Januschowski, Y. Wang // Proceedings of the VLDB Endowment. – 2018. – Vol. 11(12). – P. 2102-2105.
127. Mori, U. A review of travel time estimation and forecasting for Advanced Traveller Information Systems / U. Mori, A. Mendiburu, M. Álvarez, J. A. Lozan // Transportmetrica A: Transport Science. – 2015. – Vol. 11, Issue 2.

128. Schmidhuber, J. Deep learning in neural networks: An overview / J. Schmidhuber // *Neural Networks* – 2015. – Vol. 61. – P. 85–117.
129. Tian, Y. LSTM-based traffic flow prediction with missing data / Y. Tian, K. Zhang, J. Li, X. Lin, B. Yang // *Neurocomputing*. – 2018. – Vol. 318. – P. 297–305.
130. Prokoptsev, N. G. Traffic flow speed prediction on transportation graph with convolutional neural networks / N. G. Prokoptsev, A. E. Alekseenko, Ya. A. Kholodov // *Computer research and modeling*. – 2018. – Vol. 10, No. 3. – P. 359–367.
131. Lv, Y. Traffic flow prediction with Big Data: a deep learning approach. / Y. Lv, Y. Duan, W. Kang, Z. Li, F. Wang // *IEEE Transactions on Intelligent Transportation Systems*. – 2014. – Vol. 16(2). – P. 865–873.
132. Sagheer, A. Time series forecasting of petroleum production using deep LSTM recurrent networks / A. Sagheer, M. Kotb // *Neurocomputing*. – 2019. – Vol. 323. – P. 203-213.
133. Bai, S. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling / S. Bai, J. Kolter, V. Koltun // 2018.
134. Injarapu, A. S. H. H. V. A survey of autonomous mobile robot path planning approaches / A. S. H. H. V. Injarapu, S. K. Gawre // *International Conference on Recent Innovations in Signal Processing and Embedded Systems, RISE-2017*. – 2017. P. 624-628.
135. Torres, J. Deep Learning for Time Series Forecasting: A Survey / J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, A. Troncoso // *Published Online*. – 2021.
136. Osipov, V. Space–time signal binding in recurrent neural networks with controlled elements / V. Osipov, M. Osipova // *Neurocomputing*. – 2018. – Vol. 308. – P. 194–204.
137. Osipov, V. Space–time structures of recurrent neural networks with controlled synapses / V. Osipov // *International Symposium on Neural Networks*. – In *Advances in neural networks*. – 2016. – P. 177–184.
138. Osipov, V. Formal aspects of streaming recurrent neural networks / V. Osipov, V. Nikiforov // *International Symposium on Neural Networks*. – In *Advances in neural networks*. – 2018. – P. 29–36.

139. Патент РФ RU 2553074 C1. Способ интеллектуальной обработки информации в нейронной сети // Патент России № RU 2553074 C1. 2015. / Осипов В.Ю.
140. Osipov, V. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers / V. Osipov, V. Nikiforov, N. Zhukova, D. Miloserdov // Neural Computing and Applications. – 2020. – Vol. 32(209).
141. Hall, F. L. Traffic Flow Theory. Chapter 2: Traffic Stream Characteristics [Электронный ресурс]. – URL: <https://www.fhwa.dot.gov/publications/research/operations/tft/chap2.pdf> (дата обращения: 19.05.2021)
142. Осипов, В. Ю. Ассоциативная интеллектуальная машина с тремя сигнальными системами / В. Ю. Осипов // Информационно-управляющие системы. – 2014. – № 5 (72). – С. 12-17.
143. Осипов, В. Ю. Ассоциативная интеллектуальная машина / В. Ю. Осипов // Информационные технологии и вычислительные системы. – 2010. – Т. 2.
144. Бабич, Т. Н. Прогнозирование и планирование в условиях рынка: учебное пособие / Т. Н. Бабич, И. А. Козьева, Ю. В. Вертакова, Э. Н. Кузьбожев // Москва : Инфра-М. – 2018. — 336 с.
145. Критерий точности и надежности прогнозов [Электронный ресурс]. – URL: https://studbooks.net/2397612/matematika_himiya_fizika/kriteriy_tochnosti_nadezhnosti_prognozov (дата обращения: 19.05.2021)
146. Михайлов, Б. Б. Автономные мобильные роботы – навигация и управление / Б. Б. Михайлов, А. В. Назарова, А. С. Ющенко // Известия ЮФУ. Технические науки. – 2016. – № 2 (175). – С. 48-67.
147. Usha, M. N. S. Military Reconnaissance Robot / M. N. S. Usha, S. Priyadharshini, K. R. Shree, P. Devi, G. Sangeetha // International Journal of Advanced Engineering Research and Science. – 2017. – Vol. 4. – P. 49-56.
148. Ghouse, Z. Military Robot for Reconnaissance and Surveillance using Image Processing / Z. Ghouse, N. Hiwrale, N. Ranjan // International Research Journal of Engineering and Technology (IRJET). – 2017. – Vol. 4, Issue 5.

149. Springer, P. J. Outsourcing War to Machines: The Military Robotics Revolution / P. J. Springer // Praeger Security International. – 2018. – 280 p.
150. Патент РФ RU 2446463 C1. Способ и устройство интеллектуальной обработки информации в нейронной сети // Патент России № RU 2446463 C1. 2012. / Осипов В.Ю.
151. Осипов, В. Ю. Нейросетевое прогнозирование событий для роботов с непрерывным обучением / В. Ю. Осипов, Д. И. Милосердов // Информационно-управляющие системы. – 2020. – №5(108). – С. 33-42.
152. Патент РФ RU 2737227 C1. Способ интеллектуальной многоуровневой обработки информации в нейронной сети // Патент России № RU 2737227 C1. 2020. / Осипов В.Ю.
153. Свидетельство о государственной регистрации программы для ЭВМ №2019662053. Осипов В. Ю., Милосердов Д. И. Программа прогнозирования событий на основе рекуррентных нейронных сетей с управляемыми элементами. 2019.
154. Свидетельство о государственной регистрации программы для ЭВМ №2020616182. Милосердов Д. И. Программа прогнозирования событий с непрерывным обучением на основе рекуррентной нейронной сети с управляемыми элементами. 2020.
155. Милосердов Д. И. Архитектурные особенности программных систем нейросетевого прогнозирования с непрерывным обучением / Д. И. Милосердов // Информационные технологии. – 2020. – Т. 26, № 11. – С. 641-647.
156. Яндекс.Карты [Электронный ресурс]. – URL: <http://www.yandex.ru/maps> (дата обращения: 19.05.2021)
157. RTMC *.traffic Data (Archived Loop Data) Downloads Provided by Dr. Taek M. Kwon [Электронный ресурс]. – URL: <http://www.d.umn.edu/~tkwon/TMCdata/TMCarchive.html> (дата обращения: 19.05.2021)

СПИСОК ИЛЛЮСТРАТИВНОГО МАТЕРИАЛА

Список рисунков

1. Классификация методов прогнозирования событий по степени формализации.....	22
2. Классификация методов прогнозирования временных рядов.....	23
3. Классификация способов реализации нейросетевых интеллектуальных систем.....	28
4. Сверточная нейронная сеть, реализованная на базе нейрочипа TrueNorth.....	29
5. Классификация систем нейросетевого прогнозирования временных рядов.....	30
6. Обобщенная модель системы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением.....	39
7. Структурная схема рекуррентной нейронной сети с управляемыми элементами.....	48
8. Рекуррентная нейронная сеть с линейной структурой слоев.....	49
9. Блок-схема, описывающая выполнение метода прогнозирования с непрерывным обучением системы с временными сдвигами сигналов.....	52
10. Структурная схема системы нейросетевого прогнозирования с временными сдвигами.....	53
11. Сущность процесса формирования пространственно-временной модели в РНС с управляемыми элементами и последующего ассоциативного вызова информации из памяти сети.....	54
12. Канальная структура РНС-1 при обучении с временными сдвигами сигналов.....	54
13. Блок-схема, описывающая выполнение метода прогнозирования временных рядов без временных сдвигов с управлением направленностью вызова сигналов из ассоциативной памяти.....	57
14. Структурная схема системы нейросетевого прогнозирования без временных сдвигов.....	58
15. Сущность процесса формирования пространственно-временной модели	

в РНС с управляемыми элементами и последующего ассоциативного вызова информации из памяти сети.....	58
16. Канальная структура РНС-1 при обучении в интересах прогнозирования без временных сдвигов.....	59
17. Схема управления прогнозированием в условиях коротких выборок, шумов и пропусков.....	60
18. Архитектура программной системы нейросетевого прогнозирования с непрерывным обучением.....	66
19. Диаграмма классов UML для параллельной архитектуры программной системы нейросетевого прогнозирования с непрерывным обучением.....	67
20. Блок-схема, демонстрирующая процесс функционирования программной архитектуры.....	71
21. Архитектура программной системы нейросетевого прогнозирования с рациональным использованием памяти.....	76
22. Диаграмма классов UML для буферной архитектуры программной системы нейросетевого прогнозирования с непрерывным обучением.....	77
23. Блок-схема, демонстрирующая функционирование программной архитектуры при сокращенных затратах памяти.....	79
24. Функция гиперболического тангенса и ее область условной Чувствительности.....	81
25. Пример состояния первого слоя РНС-1 при обучении.....	87
26. Пример состояния первого слоя РНС-2 во время прогнозирования.....	87
27. Состояние первого слоя РНС-2 при прогнозировании методом без временных сдвигов.....	88
28. Результаты прогнозирования методом с временными сдвигами.....	90
29. Результаты прогнозирования методом без временных сдвигов.....	91
30. Примеры конфигураций слоев РНС с управляемыми элементами.....	95
31. Пример последовательности СЕО, несущей информацию о лексическом содержании новостных текстов, для словаря размером 306 слов.....	97
32. Спектральное представление радиосигнала в полосе 970...990 МГц	

в обычной обстановке и при появлении аномального сигнала (канал управления БПЛА «Орлан-10») на частоте 980 МГц.....	104
33. Файловый состав программного обеспечения, реализующего методы нейросетевого прогнозирования с непрерывным обучением.....	106
34. Окно чтения/записи параметров РНС-1.....	107
35. Интерфейс программной системы нейросетевого прогнозирования (в режиме обучения).....	108
36. Интерфейс программной системы нейросетевого прогнозирования (в режиме обучения и прогнозирования).....	109
37. Функционирование программной системы нейросетевого прогнозирования с непрерывным обучением.....	109
38. Место системы прогнозирования с непрерывным обучением в навигационном модуле интеллектуального автономного робота класса «беспилотное транспортное средство».....	111

Список таблиц

1. Каналы передачи информации в программной архитектуре нейросетевого прогнозирования с непрерывным обучением.....	72
2. Каналы передачи информации в программной архитектуре нейросетевого прогнозирования с сокращенными затратами памяти.....	78
3. Кодирование поддиапазонов скоростей движения на логических полях размером 8×4 нейронов.....	85
4. Показатели эффективности предлагаемых методов.....	89
5. Сравнение показателей эффективности прогнозирования транспортных потоков по предлагаемым методам и известным аналогам.....	92
6. Результаты оценки времени прогнозирования по методу с временными сдвигами и без сдвигов.....	92
7. Показатели точности прогнозирования по предлагаемым методам на втором наборе данных.....	94
8. Результаты оценивания точности прогнозирования.....	95

9. Результаты прогнозирования лексического содержания новостных лент для горизонта 1 час и словаря 306 слов.....	99
10. Результаты прогнозирования лексического содержания новостных лент для горизонта 1 час и словаря 1000 слов.....	99
11. Результаты прогнозирования лексического содержания новостных лент для горизонта 4 часа и словаря 1000 слов.....	100
12. Сравнительные показатели эффективности прогнозирования нейронными сетями LSTM и РНС с управляемыми элементами.....	101
13. Значения показателей эффективности ПО (компилятор MinGW).....	103
14. Значения показателей эффективности ПО (компилятор MSVC).....	103

Приложение А. Список публикаций соискателя по теме диссертации

Публикации в журналах из перечня рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёной степени кандидата наук, на соискание учёной степени доктора наук:

1. Осипов В. Ю., Милосердов Д. И. Нейросетевое прогнозирование событий для роботов с непрерывным обучением // Информационно-управляющие системы. – 2020. – №5(108). – С. 33-42. DOI: 10.31799/1684-8853-2020-5-33-42
2. Милосердов Д. И. Архитектурные особенности программных систем нейросетевого прогнозирования с непрерывным обучением // Информационные технологии. – 2020. – Т. 26, № 11. – С. 641-647. DOI: 10.17587/it.26.641-647

В зарубежных изданиях, индексируемых в WoS/Scopus:

3. Osipov V., Nikiforov V., Zhukova N., Miloserdov D. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers // Neural Computing and Applications. – 2020. – Vol. 32(209). DOI: 10.1007/s00521-020-04843-5
4. Osipov V., Kuleshov S., Zaytseva A., Levonevskiy D., Miloserdov D. Neural network forecasting of news feeds // Expert Systems with Applications – 2020. – Vol. 169. DOI: 10.1016/j.eswa.2020.114521
5. Osipov V., Miloserdov D. Neural Network Forecasting of Traffic Congestion // Digital Transformation and Global Society, DTGS 2019. – In Communications in Computer and Information Science. – 2019. – Vol. 1038. DOI: 10.1007/978-3-030-37858-5_20
6. Osipov V., Zhukova N., Miloserdov D. Neural Network Associative Forecasting of Demand for Goods // Experimental Economics and Machine Learning, EEML 2019. – 2019. – Vol. 2479.
7. Miloserdov I., Miloserdov D. Development of Stability Control Mechanisms in Neural Network Forecasting Systems // Journal of Physics: Conference Series, 2021. DOI: 10.1088/1742-6596/1864/1/012105

В других изданиях:

8. Милосердов Д. И. Программный комплекс нейросетевого прогнозирования временных рядов // 5-я Международная научная конференция «Технологическая перспектива в рамках евразийского пространства: новые рынки и точки экономического роста». – 2019. – С. 166-169.
9. Милосердов И.В., Милосердов Д.И. Разработка механизмов обеспечения устойчивости в нейросетевых системах прогнозирования (Материалы конференции «Информационные технологии в управлении», 2020 г.) URL: <https://itc.etu.ru/assets/files/itc-2020/papers/198.pdf>
10. Милосердов Д. И. Нейросетевое прогнозирование событий для интеллектуальных роботов с непрерывным обучением // Технологические тренды и наукоемкая экономика: бизнес, отрасли, регионы. Раздел 1: Цифровые технологии наукоемкой экономики. – 2021. – С. 27-37. DOI: 10.53115/9785001880134

Интеллектуальная собственность:

11. Свидетельство о государственной регистрации программы для ЭВМ №2019662053. Осипов В. Ю., Милосердов Д. И. Программа прогнозирования событий на основе рекуррентных нейронных сетей с управляемыми элементами. 2019.
12. Свидетельство о государственной регистрации программы для ЭВМ №2020616182. Милосердов Д. И. Программа прогнозирования событий с непрерывным обучением на основе рекуррентной нейронной сети с управляемыми элементами. 2020.

Приложение Б. Полученные свидетельства об интеллектуальной
собственности

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ
№ 2019662053

**Программа прогнозирования событий на основе
рекуррентных нейронных сетей с управляемыми
элементами**

Правообладатель: *Федеральное государственное бюджетное
учреждение науки Санкт-Петербургский институт
информатики и автоматизации Российской академии наук (RU)*

Авторы: *Осипов Василий Юрьевич (RU),
Милосердов Дмитрий Игоревич (RU)*

Заявка № **2019660896**
Дата поступления **09 сентября 2019 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **16 сентября 2019 г.**



*Руководитель Федеральной службы
по интеллектуальной собственности*

 **Г.П. Изrael**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2020616182

Программа прогнозирования событий с непрерывным обучением на основе рекуррентной нейронной сети с управляемыми элементами

Правообладатель: *Милосердов Дмитрий Игоревич (RU)*

Автор: *Милосердов Дмитрий Игоревич (RU)*



Заявка № 2020614706

Дата поступления 25 мая 2020 г.

Дата государственной регистрации
в Реестре программ для ЭВМ 11 июня 2020 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Г.П. Илиев Г.П. Илиев

Приложение В. Акты внедрения результатов диссертационной работы



Центр системного анализа и моделирования -
филиал АО «Научно-технический центр
радиоэлектронной борьбы»
(ЦСАМ АО «НТЦ РЭБ»)

Большой Смоленский пр-т, д. 2, оф 518
г. Санкт-Петербург, 192029,
т./ф.: (812) 412-77-88,
e-mail: csam@ntc-reb.ru

№ С-531 от «31» мая 20 21 г.

УТВЕРЖДАЮ
Директор
Центра системного анализа и
моделирования - филиала
АО «Научно-технический центр
радиоэлектронной борьбы»
доктор технических наук, доцент

Л. А. Овчаренко

«31» мая 2021 г.



АКТ

об использовании результатов диссертационной работы на соискание
ученой степени кандидата технических наук «Модели, методы и
архитектуры программных систем нейросетевого прогнозирования
трудноформализуемых событий с непрерывным обучением»
Милосердова Дмитрия Игоревича в ЦСАМ АО «НТЦ РЭБ» при
проведении исследований по обнаружению и траекторному
сопровождению малоразмерных беспилотных летательных аппаратов



Научно-техническая комиссия в составе: председателя – начальника отдела,
кандидата технических наук Никитченко А. Н., членов комиссии – старшего научного
сотрудника, кандидата технических наук Горохова Н. Л., ведущего инженера-
программиста, кандидата физико-математических наук Ананьева А.В. составила
настоящий акт о том, что основные научные результаты диссертационной работы
Милосердова Дмитрия Игоревича, а именно:

- методы нейросетевого прогнозирования событий с непрерывным обучением;
- программный продукт, реализующий методы прогнозирования, выполненный с
использованием фреймворка «Qt»,
использованы в ЦСАМ АО «НТЦ РЭБ» при проведении научных исследований по
обнаружению и траекторному сопровождению малоразмерных беспилотных летательных
аппаратов для прогнозирования радиолокационной обстановки и выявления аномальных
радиосигналов.

Председатель комиссии:

 А. Никитченко

Члены комиссии:

 Н. Горохов
 А. Ананьев

«31» мая 2021 года

МИНОБРНАУКИ РОССИИ
**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
«САНКТ-ПЕТЕРБУРГСКИЙ ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР РОССИЙСКОЙ
АКАДЕМИИ НАУК» (СПб ФИЦ РАН)**

14 линия В.О., д. 39, Санкт-Петербург, 199178
Телефон: (812) 328-34-11, факс: (812) 328-44-50, E-mail: info@spcras.ru, https://spcras.ru/
ОКПО 04683303, ОГРН 1027800514411, ИНН/КПП 7801003920/780101001

УТВЕРЖДАЮ

заместитель директора
по научной работе СПб ФИЦ РАН



С.В. Кулешов

«28» февраля 2022 г.

А К Т

**об использовании результатов диссертационной работы
Милосердова Дмитрия Игоревича
«Модели, методы и архитектуры программных систем
нейросетевого прогнозирования трудноформализуемых событий
с непрерывным обучением»
в научно-исследовательской работе СПб ФИЦ РАН**

Комиссия в составе: председателя – к.т.н. Савельева Антона Игоревича, членов комиссии – к.т.н. Левоневского Дмитрия Константиновича, к.т.н. Мотиенко Анны Игоревны, составила настоящий акт о том, что научные результаты, полученные Милосердовым Дмитрием Игоревичем в рамках выполнения диссертационной работы «Модели, методы и архитектуры программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением», а именно:

- методы нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением;
- архитектуры программных систем нейросетевого прогнозирования трудноформализуемых событий с непрерывным обучением;

были внедрены в научно-исследовательской работе СПб ФИЦ РАН № 0073-2019-0001 «Теоретические основы и алгоритмические модели когнитивного управления, взаимодействия и анализа состояния групп гетерогенных робототехнических комплексов».

Реализация положений, выносимых на защиту в диссертационной работе, позволила повысить точность получаемых прогнозов трудноформализуемых событий в интересах робототехнических комплексов за счет использования новых правил управления ассоциативной обработкой

информации в потоковых рекуррентных нейронных сетях. Комиссия отмечает теоретическую и практическую значимость полученных в диссертационной работе научных результатов.

Председатель комиссии:
Старший научный сотрудник,
руководитель лаборатории
автономных робототехнических
систем, к.т.н.



(подпись)

А.И. Савельев

Члены комиссии:
Старший научный сотрудник,
руководитель лаборатории
технологий больших данных
социокиберфизических систем, к.т.н.



(подпись)

Д.К. Левоневский

Старший научный сотрудник лаборатории
технологий больших данных
социокиберфизических систем, к.т.н.



(подпись)

А.И. Мотиенко