

на правах рукописи

Лукашев Александр Владимирович

**МОДЕЛИ И АЛГОРИТМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ
НА ОСНОВЕ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ**

05.13.01 – Системный анализ, управление и обработка информации (технические системы)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

Санкт-Петербург

2012

Работа выполнена в Санкт-Петербургском государственном электротехническом университете «ЛЭТИ» им. В.И.Ульянова (Ленина) на кафедре Автоматизированных систем обработки информации и управления

Научный руководитель:

кандидат технических наук, доцент, Дубенецкий Владислав Алексеевич

Официальные оппоненты:

доктор технических наук, профессор, Водяхо Александр Иванович, Санкт-Петербургский государственный электротехнический университет "ЛЭТИ", зам. зав. кафедрой вычислительной техники

кандидат технических наук, доцент, Краснова Анастасия Ивановна, Санкт-Петербургский государственный инженерно-экономический университет, доцент кафедры информационных систем в экономике

Ведущая организация – ФБОУ ВПО "Санкт-Петербургский государственный университет водных коммуникаций"

Защита состоится __ мая 2012 года в ____ на заседании диссертационного совета Д 212.238.07 Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И.Ульянова (Ленина) по адресу: 197376, Санкт-Петербург, ул. Проф. Попова, 5.

С диссертацией можно ознакомиться в библиотеке Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И.Ульянова (Ленина).

Автореферат разослан __ апреля 2012 года.

Ученый секретарь

диссертационного совета



Цехановский В.В.

Общая характеристика работы

Актуальность темы исследования. Задача анализа и прогнозирования временных рядов актуальна и востребована во многих развивающихся направлениях, таких как:

- интеллектуальный анализ данных (data mining);
- анализ взаимосвязей экономических данных;
- эконометрика финансовых рынков.

Отдельной актуальной задачей является создание электронных систем, предназначенных для автоматизации анализа факторов, влияющих на изменение временных рядов и подачи сигналов пользователю, либо другим системам при возникновении определенных условий.

В настоящее время в связи с развитием сферы информационных технологий, происходит её интеграция в решениях прикладных задач анализа и прогнозирования временных рядов. Одной из наиболее востребованных областей приложения задачи анализа временных рядов являются системы ведения торгов на биржах капиталов. Переход брокеров и бирж на электронные площадки позволил предоставлять конечным пользователям возможность совершения сделок посредством конечных программных терминалов, программных интерфейсов (API) и протоколов, являющихся отраслевыми стандартами (FIX, FIXML), что позволило использовать ЭВМ для реализации систем поддержки принятия решений на основе анализа рынка в реальном времени. Высокая востребованность в решении задачи в данной области определила её как основную область приложения моделей и алгоритмов, разработанных в ходе диссертационного исследования.

В области анализа временных рядов проводят исследования, как частные аналитики, так и отдельные подразделения крупных компаний, что ведет к значительной дифференциации в качественном уровне разработки и недостаточной освещенности вопроса:

- крупные разработчики не освещают технические и алгоритмические аспекты разрабатываемых моделей анализа;
- частные аналитики обладают недостаточным ресурсом для создания собственных систем анализа и проведения научно-исследовательских разработок.

Сложившаяся ситуация обуславливает необходимость исследования и формализации процессов построения, анализа и оптимизации систем поддержки принятия решений на основе анализа временных рядов с точки зрения технического, алгоритмического и структурного аспекта и подтверждает актуальность и научно-практическую значимость выбранной темы исследования, позволяя сформулировать цели и задачи диссертационной работы.

Цель исследования состоит в разработке моделей и методов реализации систем поддержки принятия оперативных решений на основе трендового анализа временных рядов.

Объектом диссертационного исследования являются системы поддержки принятия оперативных решений на основе трендового анализа временных рядов.

Предметом исследования являются модели технического анализа временных рядов и нейросетевая логическая модель.

В соответствии с поставленной целью работы, определены основные задачи диссертации:

1. разработка модели системы анализа и поддержки принятия решений на основе анализа временных рядов;
2. построение системы, позволяющей реализовать различные модели анализа;
3. разработка алгоритма оценки эффективности и дальнейшей оптимизации системы анализа и поддержки принятия решений;

4. применение аппарата нейронных сетей для анализа временных рядов;

5. экспериментальное исследование разработанных моделей, алгоритмов и методов.

Используемые методы: оптимизации, искусственных нейронных сетей, аппарата математической статистики, системного анализа, вычислительной математики, искусственного интеллекта, финансового менеджмента.

Достоверность полученных результатов обеспечивается следованием принципам системного подхода к анализу процесса; корректной интерпретацией рабочих циклов систем анализа; построением детерминированных моделей и алгоритмов анализа временных рядов, работа которых подтверждается экспериментально применительно к различным наборам тестовых данных.

Научные положения, выносимые на защиту:

1. модель системы поддержки принятия решений на основе анализа временных рядов;

2. алгоритм оценки эффективности и оптимизации системы анализа и поддержки принятия решений;

3. способ применения нейронных сетей, обеспечивающий классификацию временного ряда;

4. алгоритм создания обучающих выборок для нейронной сети;

Научная новизна:

1. предложена модель системы поддержки принятия оперативных решений на основе трендового анализа временных рядов, позволяющая строить модели анализа, применяя различные наборы индикаторов. Новизна модели состоит в универсализации блока принятия решений за счет применения настраиваемых интерпретаторов числовых значений индикаторов;

2. разработан алгоритм оценки эффективности и оптимизации системы анализа и поддержки принятия решений, отличающийся использованием интегрального показателя эффективности, получаемого путем обработки сигналов блока принятия решений для поиска наборов параметров, обеспечивающих наилучший результат;

3. разработан новый алгоритм создания выборок путем выделения и классификации подмножеств значений временного ряда для организации процесса обучения нейронной сети;

4. предложен способ применения нейронных сетей, обеспечивающий классификацию временного ряда, отличающийся использованием данных, получаемых от индикаторов.

Практическая ценность полученных результатов заключается в создании набора инструментов, применимого для построения систем поддержки принятия оперативных решений на основе трендового анализа временных рядов, а так же разработке обучаемой системы анализа, использующей механизм нейронных сетей.

Непосредственную практическую значимость имеют следующие полученные результаты:

1. система многокритериального анализа применима для прогнозирования тренда в различных процессах, представляемых в виде временных рядов;

2. инструменты оценки и оптимизации системы принятия решений на основе анализа временных рядов позволяют решить проблему подбора параметров при изменяющихся внешних факторах;

3. алгоритм выделения обучающих выборок и предложенный способ применения нейронной сети позволяет реализовать систему, автоматически регулирующую отзывчивость на различные факторы на основе их влияния в краткосрочной ретроспективе.

Внедрение результатов. Результаты диссертационной работы внедрены в учебный процесс СПбГЭТУ при преподавании дисциплины «Технологии баз данных» на кафедре автоматизированных систем обработки информации и управления.

Апробация. Основные положения и результаты диссертации докладывались на Международной заочной научно-практической конференции «Технические науки: теоретические и прикладные аспекты» (Новосибирск, 2012); Международной заочной научно-практической конференции «теория и практика актуальных исследований» (Краснодар, 2012); Региональной научно-практической конференции «Молодежь, образование и наука XXI века»

Публикации. Основные теоретические и практические результаты диссертации опубликованы в 6 статьях, в том числе 3 публикации в ведущих рецензируемых изданиях, рекомендованных в действующем перечне ВАК. Доклады доложены и получили одобрение на 3 международных научно-практических конференциях.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав с выводами по каждой из них, заключения, содержит 104 страниц машинописного текста, включая 37 рисунков, 8 таблиц и список литературы из 86 источников.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность диссертационной работы, указываются объект, предмет и методы исследования, определяются цель и задачи исследования, формулируются научная новизна, а также результаты и положения, выносимые на защиту. Приводится информация, подтверждающая практическую ценность полученных результатов, их внедрение и апробацию.

Первая глава посвящена рассмотрению текущего состояния, тенденций и перспектив развития методов анализа временных рядов. Поскольку анализ временных рядов является востребованным в задаче поддержки принятия решений при ведении торговой деятельности на рынках капиталов, данная сфера была выбрана в качестве основной для практического приложения разрабатываемых алгоритмов, методов и моделей. Дается оценка методу фундаментального анализа, как способа объяснения причин изменения тенденций во временных рядах, и технического, как способа анализа непосредственно данных временного ряда. Рассмотрены основные инструменты технического анализа – индикаторы, оценивающие статистические характеристики временного ряда. К системе анализа были сформулированы следующие требования:

- возможность построения различных моделей анализа;
- возможность анализа разного рода процессов;
- возможность работы с различными источниками данных.

Так же приведен анализ применимости различных видов нейронных сетей к задаче трендового анализа временных рядов.

Вторая глава посвящена разработке системы поддержки принятия решений на основе анализа временных рядов средствами технического анализа. В начале главы рассматривается процесс формирования сигнала. Для обеспечения работы системы был разработан набор сущностей:

- слушатель данных – получает новые значения для формирования временного ряда от внешних систем;
- элемент ряда – сущность объекта анализа;
- источник данных – предоставляет возможность хранения и получения наборов данных;
- конвертер данных – преобразует входящие данные в соответствии с требованиями (например, преобразует котировки в ценовые «свечи»);
- индикатор – вычисляет производные показатели временного ряда, отражающие его статистические характеристики;
- система анализа – производит принятие решения (формирует сигнал) на основе анализа временного ряда;
- сигнал – объект, содержащий данные, необходимые для интерпретации результата анализа;
- результат индикатора – объект, позволяющий передавать набор значений, вычисляемых индикатором;
- интерпретатор результата – компонент, позволяющий интерпретировать числовые значения индикатора в сигналы системы;
- слушатель системы анализа – интерфейс передачи сигнала окончания анализа;
- внешняя система – интерфейс передачи данных для внешних систем;
- блок принятия решений – реализует логику принятия решения.

Представленный набор сущностей позволяет применить различные модели анализа временных рядов посредством задания наборов индикаторов. Для каждого из индикаторов существует свой способ интерпретации результата. Так, для индикатора «стохастический осциллятор», разработан Джорджем С. Лэйном, применяемого в анализе временных рядов цен, применяют следующую интерпретацию: сигнал на покупку поступает, когда стохастический осциллятор падает ниже линии 20, а затем проходит эту линию снизу вверх. Сигнал на продажу поступает, когда стохастиче-

ский осциллятор поднимается выше линии 80, а затем пробивает эту линию вниз. Таким образом, результатом работы пары индикатор-интерпретатор для анализа ценовых рядов является сигнал на покупку или продажу (другими словами, прогноз восходящего или нисходящего тренда). Диаграмма сущностей представлена на рисунке 1:

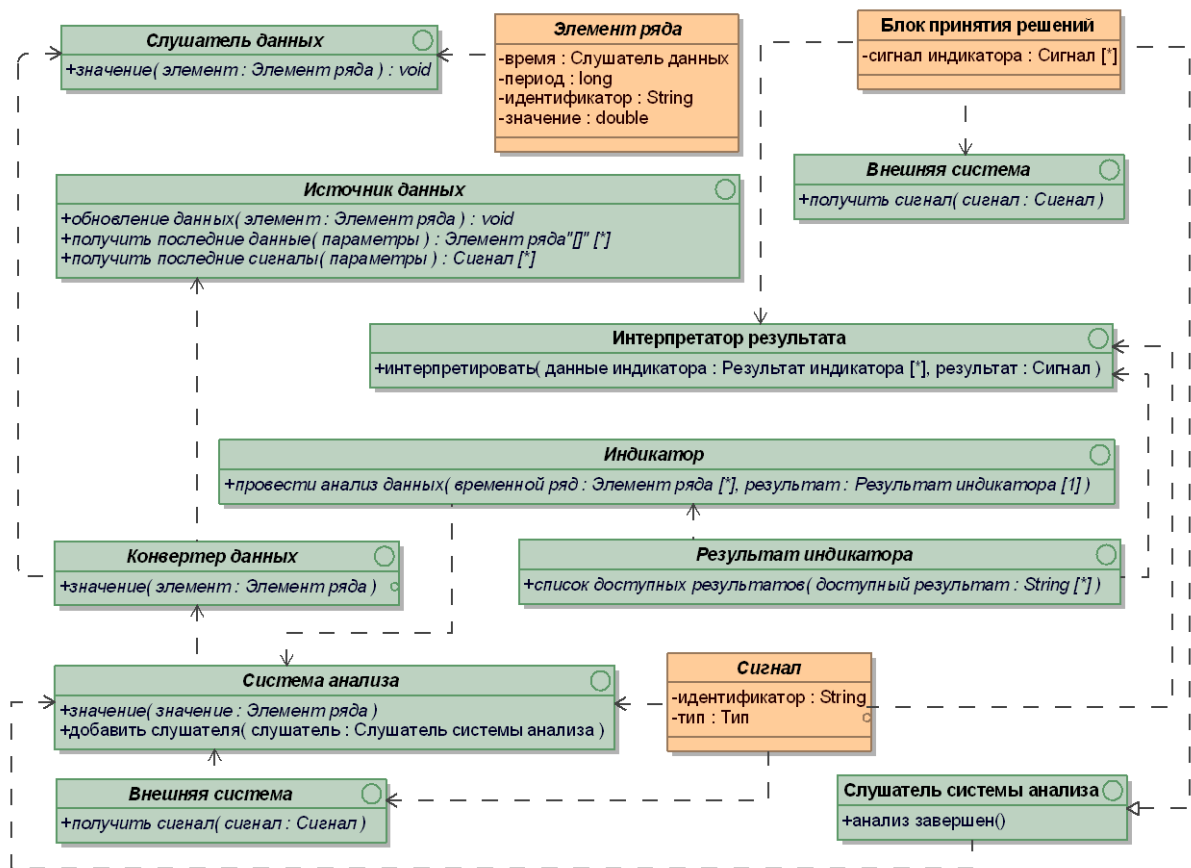


Рисунок 1. Статическая модель системы анализа

Данная модель системы показала свою состоятельность в ходе дальнейших экспериментальных исследований, определив гибкость системы анализа. Для применения данной модели к различным типам данных и областям исследований, необходимо изменение следующих сущностей:

- сигнал – необходимо определить набор типов сигналов (например: восходящее, нисходящее или боковое движение);

- индикатор/результат индикатора – необходимо определить набор инструментов анализа конкретной области исследования;
- конвертер данных – необходимо определить правила формирования элементов временного ряда на основе поступающих сигналов.

На рисунке 2 представлена временная диаграмма взаимодействий сущностей системы.

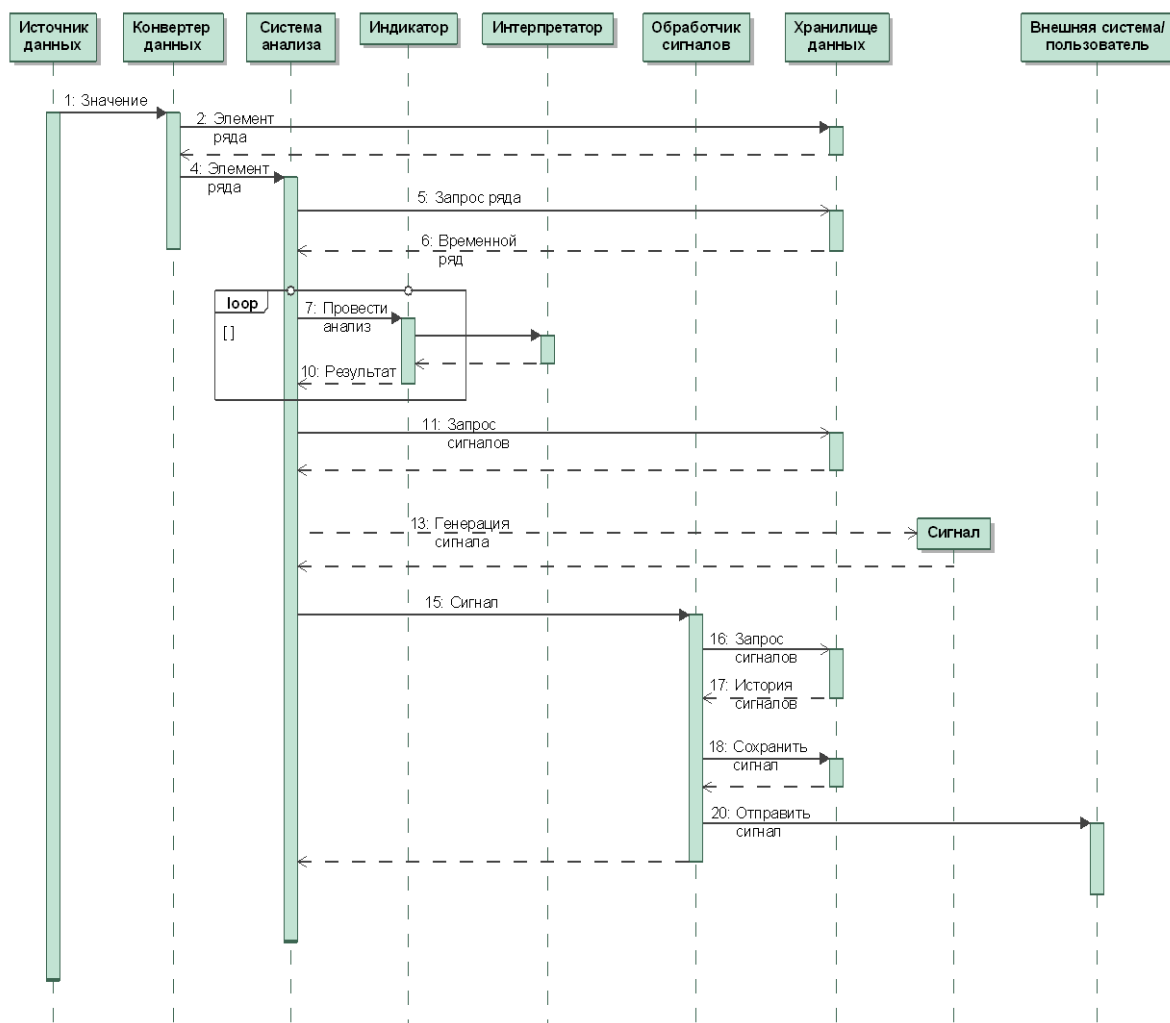


Рисунок 2. Временная диаграмма взаимодействия сущностей

Система анализа представляет собой функцию многих переменных, входом которой является подмножество числового ряда, размерность которого определяется применяемыми индикаторами. Выходом является сигнал, применяемый для дальнейшей оценки, либо непосредственно в качестве решения. Функциональная диаграмма представлена на рисунке 3.

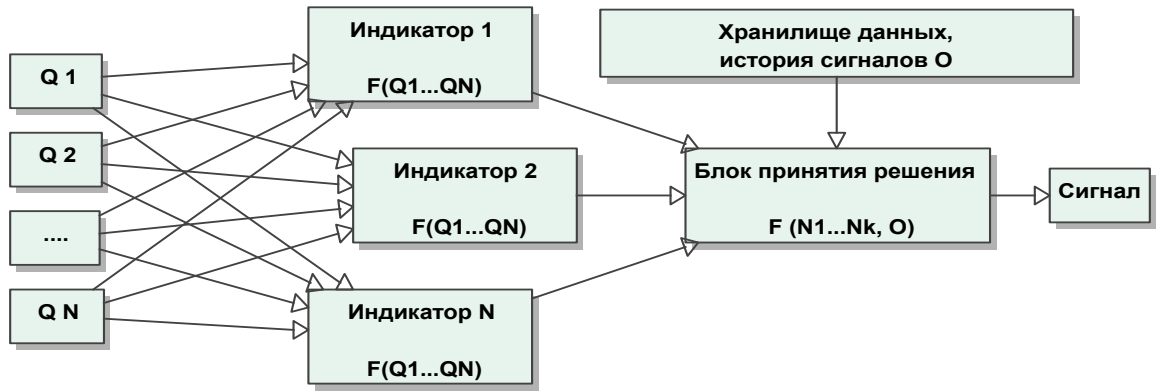


Рисунок 3. Функциональная диаграмма системы поддержки принятия решений

Далее рассматривается оценка эффективности и оптимизации системы. Для автоматизации процесса, был разработан алгоритм, позволяющий произвести оценку по сигналам системы анализа. Для симуляции процесса применяется база исторических данных. На рисунке 4 представлена функциональная схема.

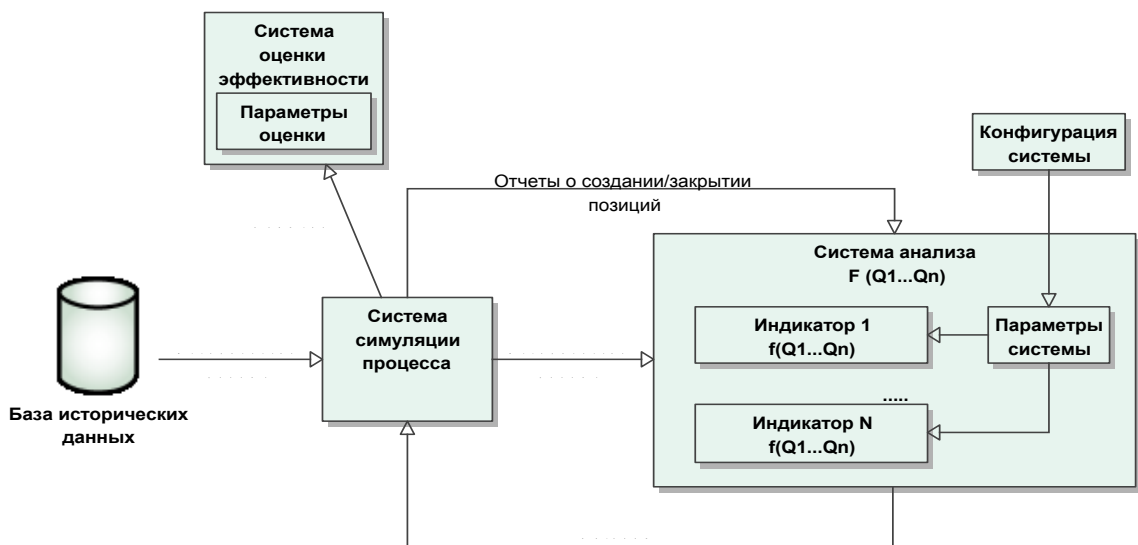


Рисунок 4. Оценка эффективности системы анализа

Для определения эффективности системы поддержки принятия решений, необходимо разработать набор критериев оценки. Так для системы поддержки принятия оперативных решений при ведении торговой деятельности был разработан следующий набор критериев:

- **Оценка риска** $R = (A_{\max} - A_{\min})/P_{\text{avg}}$, где A_{\max} и A_{\min} – локальные экстремумы, достигнутые последовательно и имеющие наибольшую разницу, P_{avg} – среднее значение временного ряда.
- **Оценка прибыли** $\varepsilon = \sum_i |P_{i+1} - P_i| / \sum_j ((P_{j+1} - P_j) * \tau_j)$, где P_i – цена, P_j – цена актива в момент j -й сделки, τ_j – функция, принимающая значение (1) сигнала восходящего тренда и (-1) нисходящего.
- **Оценка стабильности** $\sigma = N_p/N$, где N_p – количество верных сигналов, N – общее число сигналов.

Таблица 1. Результаты тестирования различных моделей

Инструмент	E	R	σ
Стратегия 1			
EUR/USD	0,0055	0,0311	0,3759
GBP/USD	0,0082	0,0257	0,4100
USD/JPY	0,0099	0,0500	0,3187
Стратегия 2			
EUR/USD	0,0074	0,0297	0,3923
GBP/USD	0,0105	0,0245	0,4167
USD/JPY	0,0125	0,0340	0,3195
Фильтрация разрывов			
EUR/USD	0,0083	0,0293	0,3973
GBP/USD	0,0119	0,0241	0,4187
USD/JPY	0,0131	0,0326	0,3288

В таблице 1 представлены результаты тестирования системы принятия решений при ведении торгов валютными парами. При условиях использования исторических данных за 2011 год при точности котировок $1e - 4$ и периоде агрегации 15 минут;

В тестировании участвовали три модели анализа:

1. индикаторы пересечения средних скользящих и среднего направления движения;
2. как и в пп.1, с использованием индикатора роста объемов сделок;
3. как и в пп.2, с использованием фильтра разрывов ряда.

Далее была рассмотрена задача автоматизации процесса оптимизации системы поддержки принятия решений. Был разработан алгоритм оптимизации системы, использующий интегральный показатель эффективности системы (для модели, рассмотренной ранее, данный показатель – модуль вектора $[R, \varepsilon, \sigma]$). Функциональная диаграмма системы оптимизации представлена на рисунке 5.

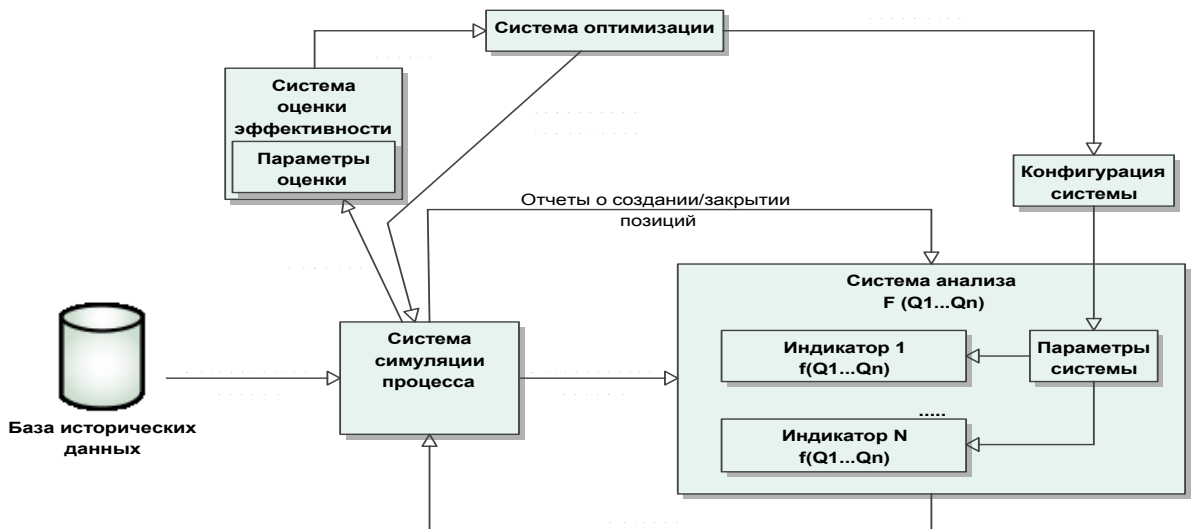


Рисунок 5. Оптимизация системы поддержки принятия решений

Так же, в результате исследований, представленных во второй главе, было выявлено, что оптимизация системы под ограниченный набор данных может не дать ожидаемых результатов на участке вне тестового набора. Отсутствие повторяемости результата на различных наборах данных определило задачу разработки обучаемой системы анализа.

Третья глава исследования посвящена решению задачи разработки системы анализа временных рядов, обучаемой на ретроспективных выборках. В качестве основы разработки был рассмотрен механизм искусственных нейронных сетей. Для определения возможности применения нейронных сетей был произведен эксперимент по классификации нейронной сетью набора точек по признаку принадлежности к классу математических функций. В ходе эксперимента обученная сеть была способна в 100% случаев определить класс входного сигнала, что позволило судить о возмож-

ности применения нейронных сетей в задаче классификации временных рядов. Для обеспечения нейронной сети обучающими выборками данных был разработан алгоритм непрерывной кластеризации и классификации подмножеств. Выделение и определение принадлежности участка к тому или иному классу производится с использованием средних скользящих значений по временному ряду. С их помощью определяется приблизительный участок тренда и направление движения на нем. Каждое пересечение говорит о предположительной смене направления изменения временного ряда. После получения такого сигнала необходимо уточнить границы участка и направление тренда внутри него. Для определения принадлежности к классу тренда, введен коэффициент отношения абсолютного изменения значения к периоду изменения $K = \frac{\sum_i^N \Delta P_i}{N}$, где N – количество периодов в отрезке, ΔP_i – изменение значения на i -м периоде. В случае превышения данного коэффициента по модулю заранее заданного значения (определяется на основе желаемой чувствительности системы) можно судить о том, что на участке имел место сильный тренд. В противном случае можно говорить о незначительном изменении значения или скачке.

Алгоритм непрерывной кластеризации данных позволил разработать систему непрерывного обучения нейронной сети. В ходе обработки поступающих данных, учитель (алгоритм кластеризации) производит поиск подмножеств, характеризующихся принадлежностью к одному из видов тренда. При нахождении такого подмножества формируется обучающая выборка, передаваемая ученику (нейронная сеть) для проведения цикла обучения, производимого в отдельном вычислительном потоке. Диаграмма последовательности представлена на рисунке 6.

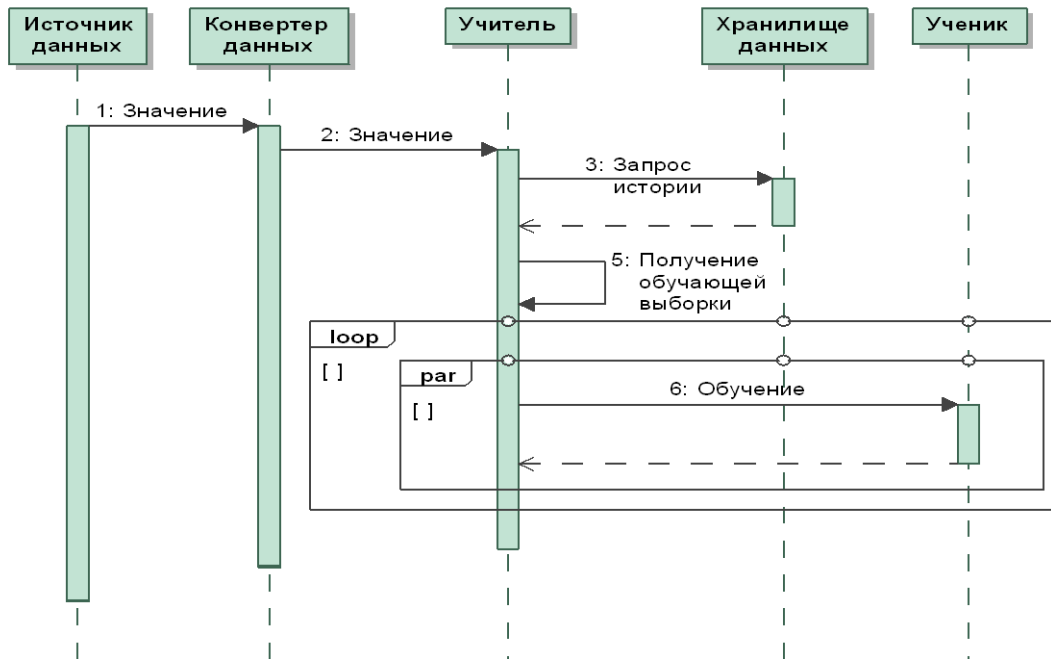


Рисунок 6. Временная диаграмма процесса обучения

В результате была разработана система поддержки принятия решений, функциональная диаграмма которой представлена на рисунке 7.

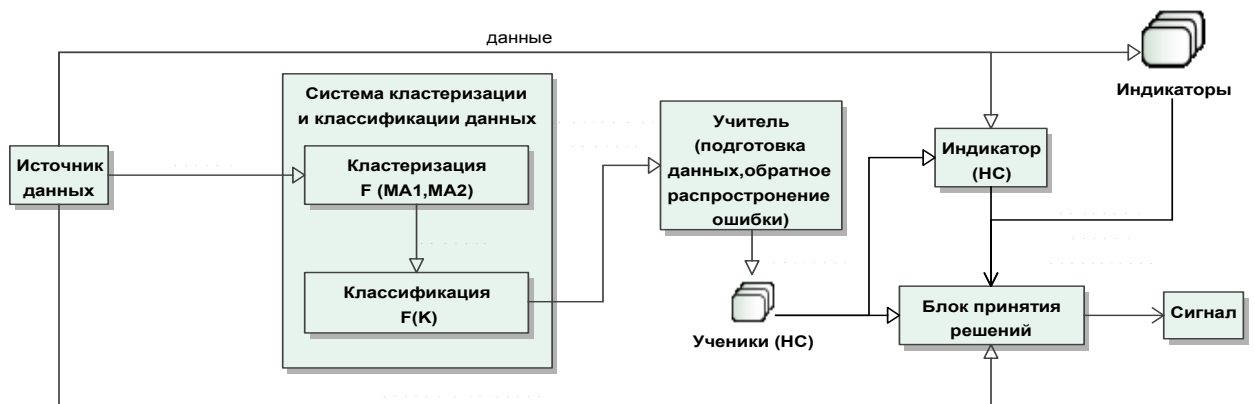


Рисунок 7. Диаграмма системы поддержки принятия решений

Четвертая глава посвящена обзору примененных технологий и особенностям реализации. Приведен обзор возможностей конфигурирования и развертывания системы.

В ходе разработки были применены наиболее современные и гибкие каркасы разработки и технологии: Apache maven2, Spring framework, Hibernate ORM;

В заключении подводятся итоги работы, делаются выводы об эффективности и применимости результатов, описываются дальнейшие пути развития работы.

Основные результаты работы

Итогом проведенных в диссертационной работе исследований явились следующие научные и практические результаты:

1. Разработана модель и реализации системы поддержки принятия оперативных решений на основе анализа временных рядов, испытанная в области ведения торговой деятельности.
2. Предложен алгоритм оценки эффективности и оптимизации системы поддержки принятия решений, получивший отражение в программной реализации.
3. Разработан алгоритм создания выборок путем выделения и классификации подмножеств значений временного ряда, позволяющий обучать сеть на краткосрочных ретроспективных данных.
4. Разработана система анализа, использующая нейронную сеть для классификации временного ряда на основании данных индикаторов, работающая в контуре с системой непрерывного обучения.

Список публикаций по теме диссертации**Публикации в изданиях, рекомендованных ВАК**

1. Лукашев А.В. Применение технического анализа для построения алгоритма работы торговой системы // Информационные системы и технологии. - 2011. - № 4 (66). - с. 94-98.
2. Лукашев А.В. Архитектура программной реализации автоматизированной торговой системы, использующей технический анализ // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. - 2011. - №06 (76). - С. 128-131.
3. Недосекин А.О., Бессонов Д.Н., Лукашев А.В. Сводный финансовый анализ российских предприятий за 2000-2003 гг. // Аудит и финансовый анализ. - 2005. - №1. - с. 53-60.

Публикации в сборниках трудов научных конференций

4. Лукашев А.В. Автоматическое непрерывное обучение нейронных сетей к распознаванию ценовых рядов // Теория и практика актуальных исследований. - 2012.
5. Лукашев А.В. Автоматизация процесса оптимизации торговых систем // Технические науки: теоретические и прикладные аспекты. - 2012. – с. 58-63.
6. Лукашев А.В. Оценка эффективности торговых систем и стратегий // Молодежь, образование и наука XXI века. – 2012. – с. 31-33.