

0-797391

На правах рукописи

ФАРХИЕВА Светлана Анатольевна

**ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ
В НАЛОГОВОМ АДМИНИСТРИРОВАНИИ
НА ОСНОВЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ
С БАЙЕСОВСКОЙ РЕГУЛЯРИЗАЦИЕЙ**

**Специальность 05.13.10 – Управление в социальных
и экономических системах**

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Уфа – 2012

Работа выполнена на кафедре математики и информатики
ФГБОУ ВПО «Всероссийский заочный финансово-экономический институт»
Филиал в г. Уфе

Научный руководитель д-р техн. наук, проф.
ГОРБАТКОВ Станислав Анатольевич
Всероссийский заочный финансово-экономический институт, филиал в г. Уфе

Официальные оппоненты д-р техн. наук, проф.
ЧЕРНЯХОВСКАЯ Лилия Рашитовна
Уфимский государственный авиационный
технический университет

д-р физ.-мат. наук, доц.
АСАДУЛЛИН Рамиль Мидхатович
Башкирский государственный педагогиче-
ский университет

Ведущая организация ФГБОУ ВПО «Уфимский государственный
нефтяной технический университет»

Защита диссертации состоится 18 мая 2012 г. в 10 часов
на заседании диссертационного совета Д-212.288.03
при Уфимском государственном авиационном техническом университете
по адресу: 450000, Уфа-центр, ул. К. Маркса, 12

С диссертацией можно ознакомиться в библиотеке института

Автореферат разослан «___» апреля 2012 г.

НАУЧНАЯ БИБЛИОТЕКА КФУ



0000741343

Ученый секретарь
диссертационного совета
д-р техн. наук, проф.

В. Митронов

В. В. Митронов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

В современных условиях развития российской экономики актуальным является максимальное наполнение государственного бюджета доходами и обеспечение их сбалансированности с расходами. В рамках приоритетных направлений совершенствования налоговой системы Российской Федерации, определенных Президентом и Правительством, предусмотрено, что использование эффективных приемов и методов налогового администрирования (НА) и дальнейшее их совершенствование должно способствовать значительному росту налоговых поступлений в бюджет. Поэтому актуальной проблемой исследований представляется совершенствование технологий подсистем налогового администрирования местного уровня на основе новейших инструментариев теории моделирования сложных систем, искусственного интеллекта, информационных технологий и экономико-математического моделирования в условиях существенного искажения данных налоговых деклараций, дефицита наблюдений и сложных условий формализации моделей.

Вопросы совершенствования контрольной работы в налоговых органах и моделей налогового администрирования представлены в работах Г. Г. Брызгалкина, И. Р. Пайзулаева, Ю. Ф. Квапца, А. З. Дадашева, А. В. Лобанова, В. А. Красницкого, Л. Ф. Черкасовой и др. Вопросам управления налогообложением в аспекте моделирования процессов сбора налогов и оценки добросовестности отдельных налогоплательщиков посвящены работы А. Б. Паскачева, Т. Н. Скорика, А. Б. Соколова, Д. Г. Черняка, М. В. Минустина и др. Проблемы интеллектуального управления и нейросетевого моделирования экономических объектов исследованы в трудах отечественных ученых В. И. Васильева, А. А. Ежова, Б. Г. Ильсова, Л. А. Исмагиловой, С. Т. Кусимова, С. А. Терехова, С. А. Шумского, Л. Р. Черняховской, Н. И. Юсуповой, зарубежных ученых И. С. Абу-Мустафы, Д.-Э. Бэстенса, В.-М. ван ден Берга, Д. Вуда. Общие вопросы теории нейронных сетей и нейрокомпьютинга изложены в работах А. И. Галушкина, А. Н. Горбана, В. Л. Дунина-Барковского, Г. Г. Малинецкого, Э. Баррона, А. Г. Ивахненко, Т. Кохонена, Ф. Розенבלата, С. Хайкина и других ученых нашей страны, ближнего и дальнего зарубежья.

Г. И. Букаевым и Н. Д. Бубликом предложена технология оценки финансовых показателей налогоплательщиков. На основе этих предложений С. А. Горбатковым разработаны общие принципы нейросетевого моделирования сложных экономических систем, включая системы налогового администрирования. Построению нейросетевых моделей (НСМ) камеральных проверок торговых предприятий посвящены исследования Н. Т. Габдрахмановой и Д. В. Полупанова. Однако, в этих работах не были рассмотрены вопросы устойчивости (регуляризации) НСМ и не были исследованы подсистемы налогового планирования и регулирования.

Несмотря на существующие многочисленные разработки в области нейросетевого моделирования, для стохастических объектов с сильно зашумленными и даже сознательно искаженными данными, методы и принципы построения эффективных, адекватных и качественных НСМ не разработаны в полном объеме. Поэтому актуальной научной задачей является разработка методов и компьютерных технологий построения НСМ поддержки принятия решений в НА в условиях высокой неопределенности данных.

Объектом исследования в работе являются процессы в системе НА, направленные на повышение собираемости налогов при сохранении финансовой устойчивости предприятий-налогоплательщиков после уплаты налогов.

Предметом исследования в работе являются НСМ, функционирующие в условиях высокой неопределенности, для представления знаний в экспертных системах по принятию решений в системе НА.

Подводя итоги анализа актуальности темы диссертации, следует отметить, что уровень автоматизации налогового контроля, планирования и регулирования в существующих информационных технологиях не соответствует запросам практики, с одной стороны, и потенциальным возможностям современного математического аппарата, в частности нейросетевых методов, с другой стороны. Эти обстоятельства определили цель данной работы и задачи исследования.

Цель диссертационной работы состоит в повышении эффективности разработки управленческих решений по налоговому администрированию местного уровня, а также повышение уровня автоматизации технологий администрирования на основе НСМ в качестве моделей представления знаний в экспертных системах НА.

Задачи исследования. Для достижения поставленной цели в диссертационной работе сформулированы следующие задачи:

1. Разработка подхода к поддержке принятия решений по НА на основе применения инженерии знаний, технологии экспертных систем и нейросетевых технологий.

2. Разработка концепции построения эффективных, устойчивых и адекватных гибридных нейросетевых математических моделей (ГНСМ) в условиях высокой неопределенности на основе общесистемных закономерностей и байесовского подхода к регуляризации НСМ.

3. Разработка метода регуляризации НСМ, реализующего концепцию по пункту 2 на основе байесовского ансамбля нейросетей.

4. Разработка метода синтеза оптимального плана выездных налоговых проверок на байесовском ансамбле нейросетей.

5. Разработка метода оценки финансовой состоятельности налогоплательщиков.

Методы исследования. В ходе исследования использовались методы системного анализа, теории принятия решений, методы разработки экспертных систем, методы нейроматематики, эконометрики и математической статистики.

На защиту выносятся положения:

1. Подход к поддержке принятия решений в НА с использованием технологии экспертных систем и технологии интеллектуального представления и накопления знаний в виде НСМ.
2. Концепция о необходимости взаимосвязанного управления информативностью данных и качеством аппроксимации в НСМ.
3. Метод вложенных математических моделей (МВММ) для построения обобщенных моделей восстановления многофакторных зависимостей, скрытых в данных, прогнозирования и ранжирования, реализующий концепцию 2.
4. Метод синтеза оптимального плана выездных налоговых проверок на байесовском ансамбле нейросетей, позволяющий повысить достоверность отбора налогоплательщиков для выездных налоговых проверок по налоговому планированию.
5. Гибридный метод оценки кредитоспособности, позволяющий оценить финансовую состоятельность налогоплательщиков и повысить оперативность принятия решений по налоговому регулированию.

Научная новизна работы:

1. Научная новизна подхода к поддержке принятия решений в НА заключается в том, что в экспертной системе НА ранее НСМ не использовались для представления знаний, что позволяет модели представления знаний функционировать в сложных специфических условиях моделирования (сильном зашумлении и искажении данных, характерных для моделируемой системы).
2. Концепция о необходимости взаимосвязанного управления информативностью данных и качеством аппроксимации в НСМ на основе байесовского ансамбля нейросетей, отличается тем, что итерационные процедуры предобработки данных, направленные на повышение их однородности и информативности с использованием вспомогательных НСМ, предлагается проводить взаимосвязано с процедурой регуляризации обучения основной (рабочей) НСМ, направленной на повышение качества и устойчивости нейросетевого отображения, что позволяет обеспечить состоятельность задачи регуляризации в условиях высокой неопределенности данных и приблизить НСМ к реальным условиям разработки управленческих решений в НА.
3. Научная новизна МВММ заключается в том, что предложенный метод, реализующий концепцию по пункту 2, отличается тем, что вводится 6 итерационных оптимизационных процедур предобработки данных на вспомогательных НСМ и 5 процедур обучения рабочей НСМ с ее реализацией, причем критерии предобработки подчинены качеству будущего обучения НС, это обеспечивает разумный компромисс между прогностическими свойствами НСМ и ее устойчивостью по выбранной числовой мере в сложных условиях моделирования при разработке управленческих решений в НА.
4. Научная новизна гибридного нейросетевого метода (ГНМС) синтеза оптимального плана выездных налоговых проверок на байесовском ансамбле нейросетей состоит в том, что принцип построения критерия ранжирования налогоплательщиков при отборе для выездных налоговых проверок оценива-

ет не только вероятность ненулевых доначислений, но и производит упорядочение (ранжирование) предприятий-налогоплательщиков по ожидаемой сумме доначислений, а также оценивает статистическую устойчивость больших отклонений в декларациях по предыстории данных. Это позволяет повысить достоверность синтезированных планов, и как следствие, повысить собираемость налогов.

5. Научная новизна гибридного метода оценки кредитоспособности налогоплательщиков отличается тем, что в его алгоритм впервые введена нейронечеткая процедура оценки кредитоспособности кластеров отдельно взятых (крупных и проблемных) налогоплательщиков. Это позволяет повысить оперативность принимаемых решений по налоговому регулированию, а также повысить собираемость налогов за счет научно обоснованного углубленного анализа финансового состояния крупных налогоплательщиков.

Обоснование достоверности научных положений, выносимых на защиту. Достоверность положений, выносимых на защиту, обосновывалась проведением широких серий вычислительных экспериментов с использованием реальных (закодированных) данных налоговых деклараций в главах 3 и 4, а также сравнением данных синтеза оптимальных планов выездных налоговых проверок с их итогами в производственных условиях для торговых и сельскохозяйственных предприятий Республики Башкортостан. Получена 83-процентная вероятность правильной диагностики нарушений налогового законодательства в декларациях (см. таблицу 2). Кроме того, производилось сравнение результатов расчета оптимального плана выездных проверок в ГНСМ с альтернативной известной модифицированной непараметрической моделью И. И. Голичева в главе 3 (см. таблицу 3).

Практическая значимость работы

Получено свидетельство о регистрации электронного ресурса «Алгоритм структурирования данных с использованием вспомогательных нейросетевых субмоделей и построение рабочей нейросетевой модели с байесовской регуляризацией» в ОФЭРНИО ИНИМ РАО № 16403 от 22.11.2010 г.

В учебном процессе Уфимской государственной академии экономики и сервиса (УГАЭС) и филиала Всероссийского заочного финансово-экономического института в г. Уфе (ВЗФЭИ) внедрено математическое обеспечение экспертной системы НА, а также методика его использования для решения задач налогового контроля, планирования и регулирования.

В работу отдела кредитных продуктов ОАО «Уралсиб» внедрен гибридный метод оценки кредитоспособности заемщиков, который позволяет малозатратно и оперативно получать достаточно достоверную информацию о финансово-экономическом состоянии заемщиков.

Научное исследование выполнено в 2002–2011 гг. во Всероссийском заочном финансово-экономическом институте. И в частности, в соответствии с комплексной темой «Обеспечение устойчивого экономического и социального развития России», по кафедральной подтеме «Разработка теоретических основ и методов нейросетевого моделирования в налогообложении».

Полученные в работе результаты по экспертной системе могут служить научной базой для модернизации аналитического блока электронной системы обработки данных (ЭОД) местного уровня, которая в настоящее время функционирует в налоговой системе России.

Апробация работы

Результаты работы обсуждались и были одобрены на следующих конференциях и семинарах: Всероссийская научно-техническая конференция «Нейроинформатика» (Москва, 2006, 2007, 2009–2012); V Всероссийская научно-практическая конференция «Проблемы и перспективы российской экономики» (Пенза, 2006); Научно-практическая конференция «Современные направления теоретических и практических исследований» (Одесса, 2006); XXII Международная научно-техническая конференция «Математические методы и информационные технологии в экономике, социологии и образовании» (Пенза, 2008); III Всероссийская научно-практическая конференция «Система ценностей современного общества» (Новосибирск, 2008); Всероссийская научно-практическая конференция «Финансовая и актуарная математика» (Уфа, 2009); Международная научно-практическая конференция «Современные проблемы и пути их решения в науке, транспорте, производстве и образовании» (Одесса, 2009); Всероссийская конференция с элементами научной школы молодежи «Проведение научных исследований в области обработки, хранения, передачи и защиты информации» (Ульяновск 2009); Всероссийская научно-практическая конференция «Социальная ответственность бизнеса: теория, методология, практика» (Уфа, 2010, 2011).

Публикации. Список публикаций по теме диссертации содержит 15 работ, в том числе 4 – в рецензируемых журналах ВАК, 1 коллективную монографию, 1 свидетельство о регистрации алгоритма.

Благодарности. Автор благодарит д-ра физ.-мат. наук, проф. И. И. Голичева, д-ра экон. наук, проф. Н. Д. Бублика, канд. техн. наук, доц. Д. В. Попланова за ценные советы при обсуждении работы.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы (109 наименований), 2 приложений и содержит 176 страниц основного текста, включая иллюстрации и таблицы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертационной работы, сформулированы цели и задачи исследования, приведены методы исследования, изложены новизна и практическая ценность выносимых на защиту положений, обоснована достоверность этих положений.

В первой главе рассмотрена актуальная, имеющая важное научно-практическое значение на современном этапе налоговой реформы и развития НА в России, проблема дальнейшей разработки современной эффективной поддержки принятия решений в системе НА, основанной на современных технологиях искусственного интеллекта и достигнутом уровне новейших информационных технологий.

В результате анализа известных подходов для поддержки принятия решений в подсистемах НА и существующих методик налогового контроля, планирования и регулирования, а также особенностей условий моделирования (сильная зашумленность данных) обоснована необходимость исследования управления процессами в подсистемах НА и разработка НСМ для представления знаний в экспертных системах по принятию решений в системе НА.

В ряде научных работ предпринимаются попытки использования ряда производственных факторов, отраженных в налоговых декларациях в виде уравнений регрессии. Однако, такой подход некорректен в силу того, что налоговые данные сильно зашумлены вплоть до сознательного их искажения. Поэтому предпосылки классического метода наименьших квадратов нарушаются. В этом направлении более перспективно применение нейросетевых интеллектуальных информационных технологий и методов нечетких множеств.

Акцент в исследовании сделан на вопросы налогового контроля и планирования, поскольку эта область наименее исследована в аспекте разработки систем поддержки принятия решений и применения математических методов вследствие сложных условий моделирования.

В работе Д. Г. Черника предложена модель сравнения декларированных и расчетных значений выходной величины для отбора налогоплательщиков для выездных проверок, построенная на комбинации уравнения регрессии с логит-пробит анализом и реализованная на модельном примере с искусственно введенными зашумлениями, т. е. строится имитационная модель. В реальную практику указанная модель не внедрена.

Коллективная монография под редакцией А. Н. Романова посвящена вопросам планирования выездных налоговых проверок на основе модели непараметрического сглаживания, которую можно рассматривать как альтернативную компьютерную технологию налогового контроля по отношению к разработанной в диссертации. Вопросы моделирования задач налогового планирования и регулирования в ней не рассматривались. Анализ работ отечественных и зарубежных авторов в области подсистем НА определил круг научных и практических задач, решение которых обеспечит эффективность управления процессом НА за счет применения интеллектуальных экспертных систем с представлением знаний с помощью НСМ, на основе байесовского ансамбля регуляризованных нейросетей.

Во второй главе разрабатывается подход к поддержке принятия решений по НА. Разработана когнитивная мультиагентная модель (карта) финансовых и информационных потоков в бюджетно-налоговой системе регионального и муниципального уровней (рисунок 1) в целях системного видения процессов в системе НА. Это позволило декомпозировать общую задачу НА на подзадачи для каждой подсистемы (агента).

В подсистемах НА можно выделить в НСМ общее «ядро» в виде восстанавливаемой параметрической «обобщенной производственной функции» $\hat{Y}(\bar{X}, W)$, где \bar{X} – вектор входных факторов, W – совокупность оцениваемых параметров (в нейросети – это матрица синаптических весов). Моделирование такого «ядра» позволяет разработать общую методологию моделирования в нейросетевом базисе для подсистем НА в виде совокупности методов и алгоритмов.

В предварительных исследованиях был проведен модельный вычислительный эксперимент, который показал, что при повышении некоторого критического значения интенсивности шумов и их объема НСМ теряет устойчивость. Это обосновывает необходимость предобработки данных и регуляризации (сужения класса допустимых решений) НСМ, что обеспечивает повышение достоверности оценок с помощью НСМ в тяжелых условиях моделирования.

Разработана концепция регуляризации НСМ на основе байесовского ансамбля нейросетей, которая дает возможность парировать дисфункции структурируемой системы в НСМ восстановления многомерных нелинейных зависимостей, скрытых в данных.

Концепция о необходимости взаимосвязного управления информативностью данных и качеством их аппроксимации в НСМ. Оценку условного математического ожидания $\hat{Y}(\bar{X}, W)$ при восстановлении многомерной нелинейной зависимости, скрытой в исходных данных, $D = \{ \langle \bar{x}_i, y_i \rangle, i = \overline{1, N} \}$ с помощью нейросетевого отображения целесообразно выполнять, используя не одну НСМ, а байесовский ансамбль из нескольких нейросетей $\{h_q(\bar{X}, W)\}$, где h_q -я априорная гипотеза о порождении эмпирических данных. При этом все априорные гипотезы-нейросети $\{h_q(\bar{X}, W)\}$ принадлежат к одной метагипотезе (классу H). После построения гипотез-нейросетей производится их фильтрация по вероятностному критерию P_q^* равному отношению числа «хороших» точек N_q^* (с допустимой относительной ошибкой вида $S, \leq \xi, \%$, где ξ – экспертно задаваемый уровень ошибки объяснения данных) к общему числу точек N . В итоге априорный ансамбль гипотез $\{h_q(\bar{X}, W)\}$ сужается, что соответствует увеличению наших знаний о моделируемом объекте. Финишная оценка производится путем осреднения на отфильтрованном байесовском ансамбле нейросетей.

Новым в концепции является введение критерия фильтрации $P_q^* = N_q^* / N$ в отсутствие каких-либо предложений о распределении шумов, а также выявление возможности управления критериями P_q^* посредством предпроцессорной обработки данных.

В третьей главе разрабатываются методы поддержки принятия решений в подсистемах НА, модели представления знаний и алгоритмическое обеспечение для последующей их реализации.

Разработан основной метод моделирования (МВММ), реализующий предложенную концепцию в главе 2 (рисунок 3).

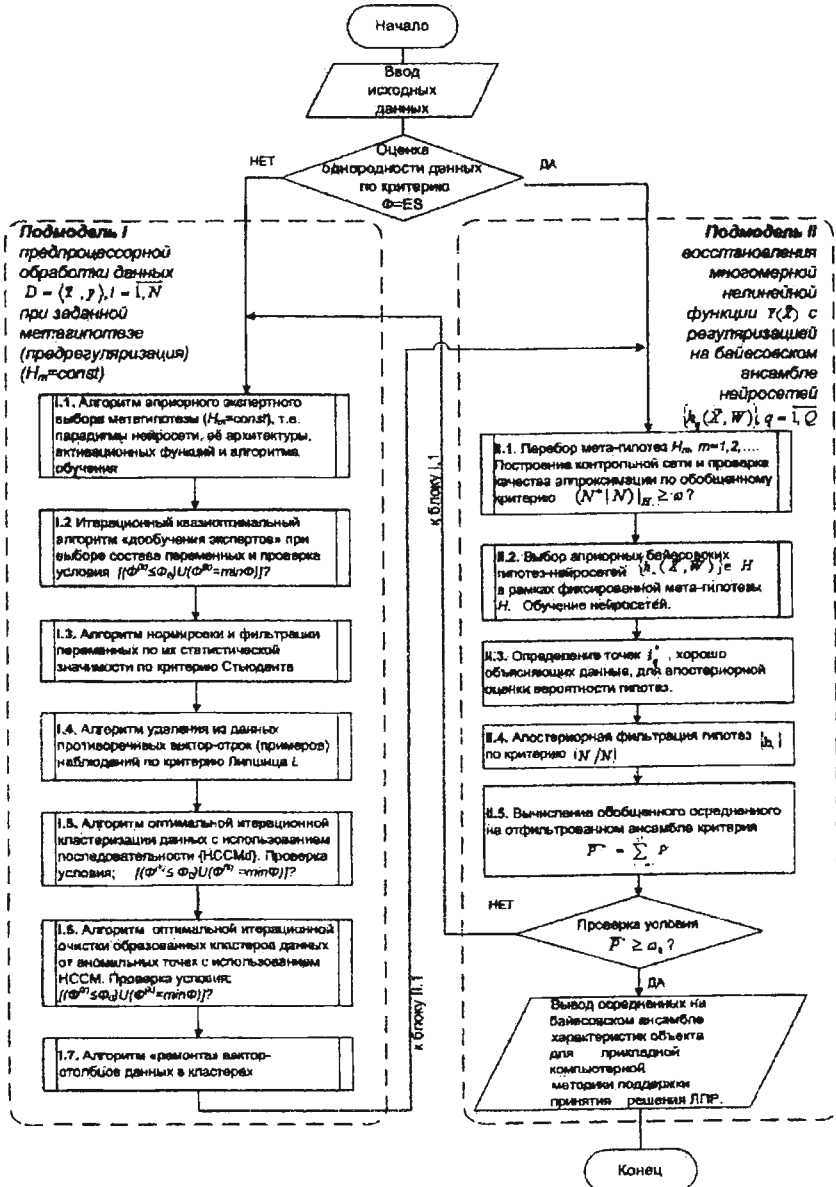


Рисунок 3 – Логическая схема МВММ

Главной целью этапов I.1–I.7. предпроцессорной обработки данных в рамках подмодели I является повышение однородности данных и улучшение их качества (информативности) в аспекте обучения нейросети. Подчинение критерия предобработки $\Phi = E \cdot S$ (где E – ошибка обобщения; S – критерий устойчивости) во вспомогательных нейросетевых субмоделях целям качества будущего обучения рабочей НСМ в подмодели II – это и есть суть нового подхода к предобработке данных. При этом в I.1–I.7 используется априорно выбранная экспертным способом парадигма сети, ее архитектура и вид активационных функций и тем самым фиксируется мета-гипотеза H байесовского подхода к регуляризации нейросети.

В рамках подмодели II считается зафиксированной мета-гипотеза выбранная в П.1. Смысл алгоритмов П.1–П.5 подмодели II состоит в том, что выбранный ансамбль априорных гипотез $\{h_q(\vec{X}, W)\} \in H$, в котором варьируются архитектура сети и вид активационных функций, после обучения и оценки его прогностических свойств – сужается (фильтруется). Согласно теории байесовского подхода, осредненные оценки на байесовском ансамбле всегда лучше оценок одной, даже лучшей нейросети ансамбля.

Необходимость взаимосвязанного управления информативностью данных и качеством их аппроксимации в нейросети для обеспечения состоятельности задачи регуляризации реализуется в МВММ следующим образом. На рисунке 3 между подмоделями I и II имеются перекрестные связи. Таким образом, в подмодель II регуляризации и обучения байесовского ансамбля подаются уже структурированные данные, что создает предпосылки хорошего качества обучения ансамбля и регуляризации всей НСМ на ансамбле. Перекрестная связь от П.5 к I.1 дает возможность организации внешнего итерационного цикла по оптимизации выбора мета-гипотезы H_n .

В третьей главе также проводится оценка адекватности МВММ на основе вычислительных и натуральных экспериментов и сравнение с альтернативным методом. Стратегия и средства реализации этапов метода: оптимальная кластеризация выполнялась с помощью пакета *ОЛИМП:Стат.Эксперт*; построение вспомогательных инверсных НССМ в алгоритме I.7 производилось с помощью пакета *NeuroSolution 5* (демо-версия); для каждого этапа предобработки данных по МВММ оценки находились усреднением по байесовскому ансамблю нейросетей.

На рисунке 4 представлены значения критерия Липшица по этапам предобработки данных (1 – исходная база данных; 2 – нормировка данных очищенных кластеров I.4; 3 – оптимальная итерационная кластеризация I.5; 4 – устранение аномальных точек в образованных кластерах I.6; 5 – «ремонт» вектор-столбца I.7).

Каждый этап предобработки данных вносит свой последовательный вклад в повышение однородности данных, углубляя их структурирование (кривая $L(k)$ монотонно снижается с ростом номера этапа k). Суммарный вклад всех этапов в повышение однородности данных весьма ощутим.

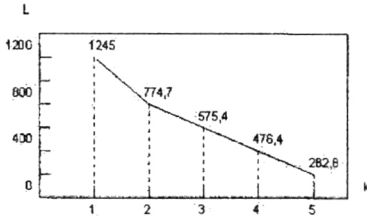


Рисунок 4 – Значения критерия Липшица по всем этапам алгоритма предобработки данных

Таким образом, вычислительный эксперимент подтвердил идею использования НССМ в блоках I.4–I.7.

Оценки байесовской регуляризации П.1–П.6, приведены в таблице 1.

Таблица 1 – Оценка совместных процедур предобработки данных и регуляризации обучения гипотез-нейросетей $\{h_q\}$

	h_1	h_2	h_3	h_4	h_6	h_7	h_9	h_{10}	Среднее	h_7 на "сырых" данных
E	0.049	0.053	0.041	0.057	0.055	0.054	0.071	0.057	0.057	0.449
S	0.530	0.363	0.856	0.720	0.520	0.460	0.540	0.650	0.582	5.249
Φ	0.026	0.019	0.035	0.040	0.020	0.020	0.039	0.037	0.036	2,36

Из таблицы 1 видно: процедуру фильтрации прошли 8 гипотез-нейросетей из 10 (таким образом, адекватность ГНСМ установлена с вероятностью $\bar{P}^* = 95\%$); предобработка данных совместно с байесовской регуляризацией в составе МВММ позволили более чем на порядок улучшить качество НСМ.

Для подсистем *налогового контроля и планирования* разработан метод синтеза оптимального плана выездных налоговых проверок на байесовском ансамбле нейросетей. Идея метода:

1) относительные отклонения $\delta_i^* = (\sum_{q=1}^{Q^*} \delta_{q,i}) / Q^*$ ($Q^* \leq Q; i = \overline{1, N_{Q^*}}$) усредняются на отфильтрованном байесовском ансамбле нейросетей (Q^* - количество гипотез-нейросетей, оставшихся после фильтрации).

2) налогоплательщики ранжируются по критерию отбора $\{\psi_g\}$: $\psi_g = |\delta_g| \cdot P(\delta_g > \bar{\delta}_g) \cdot M_g$; $\bar{\delta}_g = M[\delta_{g,t}] + U_g$ (g – номер предприятия; $P(\delta_g > \bar{\delta}_g)$ – вероятность того, что текущее (по времени t) значение отклонения δ_g для g -го налогоплательщика будет больше его математического ожидания, смещенного вверх на полуширину доверительного интервала U_g ; $M[\cdot]$ – оператор математического ожидания временного ряда $\{\delta_{g,t}\}$, который считается стационарным процессом; M_g – коэффициент масштаба предприятия).

В таблице 2 представлены оптимальный план выездных проверок (составленный по критерию отбора на основе байесовского ансамбля нейросетей) и величина налоговых доначислений, полученных по результатам выездных проверок (указаны первые 12 номеров предприятий согласно оптимальному плану и с наибольшими доначислениями).

Таблица 2 – Сравнение результатов моделирования с итогами выездных налоговых проверок

№ предприятия	11	3	22	9	14	21	8	24	2	12	30	10
Значение \bar{y}_i	0.20276	0.09776	0.07272	0.06346	0.04393	0.02258	0.02081	0.01875	0.01612	0.01565	0.01513	0.01316
№ предприятия	11	9	8	24	21	22	16	10	13	3	14	2
Доначисления, руб.	7002076	3140252	2780893	1911636	1700531	1520399	1427865	1384382	1371761	1366995	1041774	839696

Для теоретической оценки адекватности НСМ предложен специальный двухступенчатый метод:

I шаг. Оценка адекватности (ее механизм) скрыта в самом алгоритме регуляризации. То есть, в процедуре фильтрации по апостериорному критерию \bar{P}^* и последующему осреднению всех расчетных характеристик на отфильтрованном ансамбле нейросетей.

II шаг. Обоснование адекватности по финишным характеристикам, полученным с помощью НСМ. То есть, сравниваются синтезированные оптимальные планы выездных налоговых проверок, полученные на каждой нейросети из отфильтрованного ансамбля. Совпадение этих планов с определенной требуемой вероятностью и служат гарантией оценки адекватности модели.

Независимым подтверждением адекватности ГНСМ служат результаты сравнения полученного на ее основе *плана отбора налогоплательщиков* для 17 проверяющих бригад с планом отбора, полученным по альтернативной модифицированной непараметрической модели (И. И. Голичев). В таблице 3 приведены коды налогоплательщиков, включенных в планы отбора по обеим моделям.

Таблица 3 – Сравнение ГНСМ с альтернативной моделью

ГНСМ	41	35	26	43	66	68	11	27	40	73	15	25	71	52	46	57	16
Модифицированная непараметрическая модель	41	60	21	35	11	26	43	18	40	75	68	27	71	46	25	66	78

Из таблицы 3 следует, что по каждой модели совпадают 12 объектов налогового контроля из 17 (совпадения обозначены цветом), т. е. модели взаимно подтверждают друг друга на 71 %. Таким образом, рассмотренная модель может быть использована как модуль экспертной системы подсистемы *налогового планирования*.

Для подсистемы *налогового регулирования* предлагается методика, основанная на общесистемном законе декомпозиции-агрегирования. Методика о введения групповой оценки финансово-экономического состояния налогоплательщиков и индивидуальной (углубленной) оценки крупных и проблемных налогоплательщиков. Это позволяет повысить оперативность и понизить затратность разработок управленческих решений.

Предлагается декомпозировать задачу оценки финансово-экономического состояния налогоплательщиков на две подзадачи:

I. Групповой оценки в виде кластеризации объектов (внутри образуемых кластеров предприятия считаются приближенно однородными в аспекте их кредитоспособности).

II. Индивидуального углубленного анализа по итогам кластеризации некоторых предприятий, либо предприятий, выбираемых из кластеров экспертно.

Методика реализуется с помощью гибридного метода оценки кредитоспособности для разработки модели поддержки принятия решений *по налоговому регулированию* с использованием нечетких правил вывода. Алгоритм метода состоит из двух этапов: 1) оценка прогнозных значений основных экзогенных (объясняющих) переменных с помощью НСМ одномерных временных рядов; 2) применение правила нечеткого вывода с использованием нечеткой пенташкалы.

В таблице 4 приведен фрагмент интервального анализа попадания одного из предприятий образованных кластеров в диапазон значений финансовых показателей, соответствующих каждому из уровней финансового состояния предприятия, определенных экспертно (интервалы попадания выделены цветом). Как видно из таблицы 4 предприятие принадлежит к классу низкого риска банкротства, что подтверждается экспертами.

Таблица 4 – Значения экономических показателей по группам уровня риска

Группа уровня риска	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	...	
Очень высокий риск	0,1-0,4	2,70-2,50	0,25-0,31	-2,00...-1,66	1,60-1,51	...
Высокий риск	0,4-0,8	2,50-1,70	0,31-0,54	-1,66...-0,33	1,51-1,18	...
Средний риск	0,8-1,5	1,70-1,17	0,54-0,83	-0,33...0,92	1,18-0,83	...
Низкий риск	1,5-2,4	1,17-0,90	0,83-0,93	0,92-1,84	0,83-0,56	...
Очень низкий риск	2,4-2,8	0,90-0,50	0,93-0,95	1,84-3,30	0,56-0,20	...

В четвертой главе рассмотрена реализация модулей экспертной системы НА с помощью прикладных ГНСМ подсистем НА, которые служат основой для поддержки принятия управленческих решений.

Модули экспертной системы реализованы с помощью нейросетевого аналитического программного продукта *BrainMaker Pro 3.11* и *NeuroSolution 5*, статистического пакета *СтатЭксперт (ОЛИМП)*.

Вычислительные эксперименты показали, что предлагаемые модули экспертной системы по принятию решений в подсистемах НА работоспособны и могут функционировать в сложных специфических условиях, что позволяет использовать ее в качестве аналитического блока программного комплекса ЭОД.

В работе разработана структурно-функциональная схема информационно-аналитической системы поддержки принятия решений по НА (рисунок 5). В ней наименьший структурный элемент – налоговая декларация, содержащая весь набор компонентов вектора \bar{X} и взаимно однозначно соответствующие этим переменным значения выходной случайной величины Y .

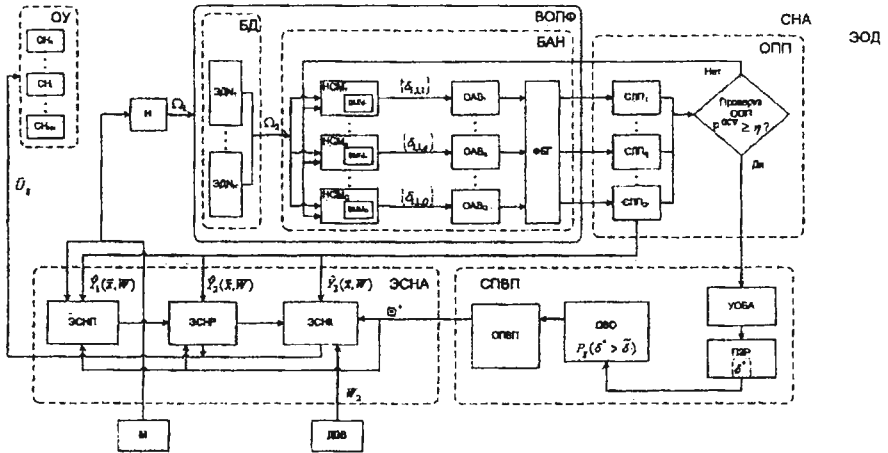


Рисунок 5 – Структурно-функциональная схема информационно-аналитической системы поддержки принятия решений в СНА

В **заключении** подводятся основные итоги выполнения исследований в аспекте логической взаимосвязи и завершенности всех задач, рассмотренных в работе.

В **приложениях** приводятся данные, использованные для построения ГНСМ в рамках экспертной системы НА.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

В диссертационной работе содержится новое решение научной задачи разработки поддержки принятия решений в НА. В ходе исследований получены следующие результаты:

1. Предложен подход к поддержке принятия решений в НА на основе технологии экспертных систем и технологии накопления и представления знаний в виде НСМ.
2. Предложена концепция о необходимости взаимосвязанного управления информативностью данных и качеством аппроксимации в НСМ на байесовском ансамбле нейросетей в условиях отсутствия априорных сведений о шумах в данных.
3. Разработан МВММ для построения обобщенных моделей восстановления многофакторных зависимостей, скрытых в данных, реализующий концепцию по пункту 2.
4. Разработан метод синтеза оптимального плана выездных проверок на основе байесовского ансамбля НСМ.
5. Разработан гибридный метод оценки кредитоспособности предприятий-налогоплательщиков.

В ходе исследования эффективности поддержки принятия решений в технологии НА показана работоспособность и эффективность предложенных концепций, методов и алгоритмов. Разработанный МВММ, алгоритмы и вы-

двинутые идеи, положенные в их основу, первоначально не были очевидны, поэтому уделено особое внимание проверке каждой идеи на широких сериях вычислительных экспериментов. Рассмотрены вопросы, связанные с прикладными моделями. Предложения по принятию решений в системе НА позволяют организовать аналитический блок, встраиваемый в программный комплекс ЭОД местного уровня, что позволит усовершенствовать технологию НА.

СПИСОК ПУБЛИКАЦИЙ

В рецензируемых журналах из списка ВАК:

1. Регуляризация нейросетевых математических моделей: константа Липшица обучающей выборки. Обзорные прикладной и промышленной математики / С. А. Горбатков, С. А. Фархиева // Обзорные прикладной и промышленной математики. Девятый Всероссийский симпозиум по прикладной и промышленной математике. М.: Редакция журнала «ОПиПМ», 2008. Т. 15, вып. 6. С. 1062–1063.

2. Метод вложенных математических моделей для регуляризации нейросетевых задач финансового контроля / С. А. Горбатков, С. А. Фархиева, М. В. Коротнева, А. А. Чапкович // Обзорные прикладной и промышленной математики. Десятый Всероссийский симпозиум по прикладной и промышленной математике. М.: Редакция журнала «ОПиПМ», 2009. Т. 16, вып. 6. С. 1050–1051.

3. Регуляризация по Тихонову нейросетевых моделей финансового контроля / С. А. Горбатков, С. А. Фархиева // Вестник Ижевского государственного технического университета. 2009. № 4 (44). С. 62–65.

4. Об одном подходе к принятию решений в налоговом администрировании на основе нейросетевых моделей с байесовской регуляризацией / С. А. Фархиева // Управление экономическими системами: электронный научный журнал. 2012. № 4 (40). 13 с. Рег. № 0421200034. (Режим доступа: http://www.uces.ru/index.php?option=com_flexicontent&view=items&id=1217)

В монографии:

5. Нейросетевое математическое моделирование в задачах ранжирования и кластеризации в бюджетно-налоговой системе регионального и муниципального уровней: монография / С. А. Горбатков, Д. В. Полупанов, А. М. Солнцев, С. А. Фархиева, И. И. Белолипцев, М. В. Коротнева, О. Б. Рашитова. Уфа: РИЦ БашГУ, 2011. 224 с.

В других изданиях:

6. Об одном алгоритме предобработки сильнозашумленных данных при построении нейросетевой модели налогового контроля / С. А. Горбатков, Д. В. Полупанов, С. А. Фархиева // Научная сессия МИФИ-2006. 8-я Всеросс. науч.-техн. конф. «Нейроинформатика-2006»: сб. науч. тр. М.: МИФИ, 2006. Ч. 3. С. 104–112.

7. Регуляризация нейросетевых математических моделей налогового контроля на основе идеи вложенных математических моделей / С. А. Горбатков,

С. А. Фархиева // Научная сессия МИФИ-2007. 9-я Всеросс. науч.-техн. конф. «Нейроинформатика-2007»: сб. науч. тр. М.: МИФИ, 2007. Ч. 3. С. 172–177.

8. Об одном алгоритме предобработки сильнозашумленных входных данных в аспекте обеспечения состоятельности задачи регуляризации для нейросетевых моделей / С. А. Горбатков, Г. А. Бесхлебнова, С. А. Фархиева // Научная сессия МИФИ-2009. 11-я Всеросс. науч.-техн. конф. «Нейроинформатика-2009»: сб. науч. тр. М.: МИФИ, 2009. Ч. 3. С. 18–26.

9. Обобщение метода вложенных математических моделей на основе байесовского подхода к регуляризации задач нейросетевого моделирования налогового и финансового контроля / С. А. Горбатков, Д. В. Полупанов, С. А. Фархиева // Научная сессия НИЯУ МИФИ-2010. 12-я Всеросс. науч.-техн. конф. «Нейроинформатика-2010»: сб. науч. тр. М.: НИЯУ МИФИ, 2010. Ч. 2. С. 228–237.

10. Свид. о рег. электронного ресурса № 16403. Алгоритм структурирования данных с использованием вспомогательных нейросетевых субмоделей и построение рабочей нейросетевой модели с байесовской регуляризацией / С. А. Горбатков, С. А. Фархиева. Зарег. в ОФЭРНяО ИНИМ РАО 22.11.2010. (Номер гос. регистрации в Национальном информационном фонде неопубликованных документов: 50201050147).

11. Приближенный метод байесовской регуляризации и двухступенчатая оценка адекватности гибридной нейросетевой модели налогового контроля / С. А. Горбатков, И. И. Белолипец, С. А. Фархиева // 13-я Всеросс. науч.-техн. конф. «Нейроинформатика-2011»: сб. науч. тр. М.: НИЯУ МИФИ, 2011. Ч. 2. С. 144–154.

12. Восстановление многомерных функциональных зависимостей с помощью нейросетевых моделей в задачах налогового администрирования / С. А. Фархиева // Научная перспектива: науч.-аналит. журнал. Уфа: Инфинити, 2011. № 1. С. 27–29.

13. Приближенный метод байесовской регуляризации нейросетевой модели налогового контроля / С. А. Фархиева // Альманах современной науки и образования. Тамбов: Грамота, 2011. № 1 (44). С. 171–173.

14. Об одном методе оценки кредитоспособности российских предприятий / С. А. Фархиева // Социальная ответственность бизнеса: теория, методология, практика: матер. II Всеросс. науч.-практ. конф., 21 ноября 2011 г., ВЗФЭИ. Уфа: Мир печати, 2011. С. 278–282.

15. Оценка эффективности алгоритмов предрегуляризации и байесовской регуляризации нейросетей для камеральной налоговой проверки / С. А. Горбатков, И. И. Белолипец, С. А. Фархиева, Д. В. Полупанов // Научная сессия НИЯУ МИФИ-2012. 14-я Всеросс. науч.-техн. конф. «Нейроинформатика-2012»: сб. науч. тр. М.: НИЯУ МИФИ, 2012. Ч. 3. С. 38–48.

Диссертант



С. А. Фархиева

ФАРХИЕВА Светлана Анатольевна

**ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ
В НАЛОГОВОМ АДМИНИСТРИРОВАНИИ
НА ОСНОВЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ
С БАЙЕСОВСКОЙ РЕГУЛЯРИЗАЦИЕЙ**

**Специальность 05.13.10 – Управление в социальных
и экономических системах**

**АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук**

**Подписано к печати 12.04.2012. Формат 60x84 1/16.
Бумага офсетная. Печать плоская. Гарнитура Таймс.
Усл. печ. л. 1,0. Уч. – изд. л. 1,0.
Тираж 100 экз. Заказ № 707**

**ФГБОУ ВПО Уфимский государственный авиационный
технический университет
Центр оперативной полиграфии
450000, Уфа-центр, ул.К. Маркса, 12**

