

На правах рукописи



БЕЛОЛИПЦЕВ Илья Игоревич

**УПРАВЛЕНИЕ В СИСТЕМАХ ФИНАНСОВОГО КОНТРОЛЯ
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 2014

Работа выполнена на кафедре математики и информатики Уфимского филиала
ФГБОУ ВПО «Финансовый университет при Правительстве Российской
Федерации»

Научный руководитель

д-р техн. наук, проф.
Горбатков Станислав Анатольевич

Официальные оппоненты

д-р ф.-м. наук, проф.
Бронштейн Ефим Михайлович
профессор кафедры вычислительной
математики и кибернетики ФГБОУ ВПО
«Уфимский государственный авиационный
технический университет»

д-р техн. наук, проф.
Буренин Владимир Алексеевич
профессор кафедры вычислительной
техники и инженерной кибернетики
ФГБОУ ВПО «Уфимский государственный
нефтяной технический университет»

Ведущая организация

ФГБОУ ВПО «Башкирский
государственный университет»

Защита диссертации состоится «3» октября 2014 г. в 10:00 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 университета

Автореферат разослан «___» _____ 2014 г.

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



В. В. Миронов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования обусловлена следующими соображениями. Наполнение государственного бюджета налоговыми поступлениями является одной из приоритетных задач государственной системы финансового контроля. Административные способы повышения собираемости налогов (увеличение налогового бремени, усиление налогового контроля со стороны государства, увеличение количества инспекторов) практически себя исчерпали. Решение данной проблемы надо искать в рамках повышения эффективности процедур камеральных проверок. Поэтому совершенствование технологий налогового администрирования на основе новейших инструментариев принятия решений, математического моделирования, информационных технологий в условиях искажения данных налоговых деклараций, дефицита наблюдений и сложных условий формализации моделей представляется актуальной задачей.

Вопросы моделирования процессов сбора налогов и оценки добросовестности отдельных налогоплательщиков рассматривались в работах А. Б. Паскачева, Д. Г. Черника, М. В. Мишустина и др. Проблемы интеллектуального управления и нейросетевого моделирования экономических объектов исследованы в трудах отечественных ученых В. И. Васильева, А. А. Ежова, Б. Г. Ильясова, Л. А. Исмагиловой, Е. М. Миркеса, С. А. Терехова, С. А. Шумского, Д. А. Тархова, 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. Горбаня, Т. Кохонена, Ф. Розенблата, С. Хайкина и других ученых нашей страны, дальнего и ближнего зарубежья.

Г. И. Букаевым была предложена технология оценки финансовых показателей налогоплательщиков. С. А. Горбатовым были разработаны общие принципы нейросетевого моделирования сложных экономических систем, включая системы налогового администрирования. Построению нейросетевых моделей камеральных проверок торговых предприятий посвящены исследования Н. Т. Габдрахмановой, Д. В. Полупанова, С. А. Фархиевой.

Степень разработанности темы исследования. Несмотря на существующие разработки в области нейросетевого моделирования для стохастических объектов с сильно зашумленными и даже сознательно искаженными данными, методы и принципы построения эффективных, адекватных и качественных нейросетевых моделей (НСМ) не были разработаны в полном объеме. В частности, отсутствуют работы посвященные оценке адекватности нейросетевых моделей сложных экономических объектов. Поэтому актуальной научной задачей является разработка методов предобработки данных во взаимосвязи с управлением качеством НСМ; методов регуляризации НСМ; разработка гибридных методов на основе НСМ и моделей поддержки принятия решений в финансовых системах в условиях высокой неопределенности данных.

Несмотря на существование множества российских и зарубежных исследований в области диагностики и прогнозирования банкротств, задача построения адекватных и надежных моделей остается востребованной и актуальной. Раннее обнаружение тенденций банкротства и, соответственно упреждающие процедуры, направленные на финансовое оздоровление предприятия, по сути являются условием стабильности предприятия, его экономической безопасно-

сти. Для этого необходим систематический мониторинг финансового состояния предприятия на основе комплексного анализа экономических показателей в их нелинейной взаимосвязи. Для этого нужны достаточно информативные (адекватные) динамические экономико-математические модели диагностики платежеспособности предприятия и его кредитоспособности.

В теоретическом аспекте остается актуальным вопрос выбора оптимальной системы экономических показателей и сокращения размерности факторного пространства внутри выбранной системы показателей для построения модели банкротств в каждой конкретной отрасли. Следует отметить, что уровень автоматизации процедур налогового контроля и процесса оценки кредитоспособности предприятий 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.

Объектом исследования в работе являются процессы управления в системах финансового контроля, на примере задач налогового контроля и диагностики банкротств. Несмотря на различие в деталях, указанные классы задач имеют общую методологическую и теоретическую основу. Методологической основой является концептуальный базис нейросетевого моделирования экономических систем, функционирующих в условиях неопределенности. Теоретической основой являются конкретные методы построения устойчивых нейросетевых моделей (НСМ).

Предметом исследования в работе являются системы поддержки принятия решений на основе нейросетевых экспертных систем, функционирующие в условиях высокой неопределенности.

Цель диссертационной работы состоит в повышении эффективности систем налогового и финансового контроля, финансового менеджмента на основе новейших интеллектуальных методов моделирования. Для достижения этой цели в диссертационной работе были поставлены следующие **задачи**:

1. Разработка метода построения нейросетевых моделей восстановления многомерных экономических зависимостей, скрытых в данных.
2. Разработка метода синтеза оптимальных планов выездных налоговых проверок для максимизации налоговых поступлений в бюджет.
3. Разработка метода оценки адекватности нейросетевых моделей.
4. Разработка нейросетевой модели поддержки принятия решений по ситуационному управлению кредитным процессом.
5. Разработка алгоритмов управления выездными налоговыми проверками и процессом кредитования клиентов в банке.

Методы исследования. Для решения поставленных в работе задач применялись методы системного анализа, теории принятия решений, методы нейроматематики, эконометрики, нечетких множеств и математической статистики.

На защиту выносятся:

1. Модифицированный метод вложенных математических моделей для обратных задач восстановления многомерных зависимостей, скрытых в данных.
2. Метод синтеза оптимальных планов выездных налоговых проверок на основе фрактального и логистического подхода к учету неопределенности.

3. Многоступенчатый метод оценки и управления адекватностью нейросетевых моделей финансового контроля.

4. Нейросетевой метод разработки модели поддержки принятия решений по ситуационному управлению кредитным процессом.

5. Алгоритм управления выездными налоговыми проверками и алгоритм управления процессом кредитования клиентов банка.

Научная новизна положений, выносимых на защиту:

1. Разработанный модифицированный метод вложенных математических моделей (МВММ), содержащий алгоритмы предобработки данных и алгоритмы их аппроксимации с помощью байесовского ансамбля нейросетей, отличается тем, что управление информативностью данных и качеством их аппроксимации производится взаимосвязано для обеспечения состоятельности задачи регуляризации НСМ.

2. Разработанный гибридный метод синтеза оптимальных планов отбора налогоплательщиков для выездных налоговых проверок, отличается тем, что учитывается не только величина индикатора нарушения налогового законодательства, но и вероятность этой оценки: а) для «длинных» временных рядов указанного индикатора – через оценку фрактальной размерности; б) для «коротких» временных рядов – через логистическую трансформацию вероятности.

3. Разработанный многоступенчатый метод оценки и управления адекватностью нейросетевых моделей финансового контроля в условиях, когда нарушаются практически все предпосылки классического регрессионного анализа, отличается тем, что механизм оценки и управления адекватностью основан на принципе последовательного сужения класса искомых решений обратной задачи восстановления зависимостей, скрытых в данных. Такой подход обеспечивает адекватность НСМ в сложных условиях моделирования, отягченных дефицитом наблюдений.

4. Нейросетевой метод разработки модели поддержки принятия решений по ситуационному управлению кредитным процессом с логистической трансформацией вероятности риска банкротства, отличается тем, что показатель экспоненты логистической функции агрегируется с помощью байесовского ансамбля нейросетей в виде многомерной нелинейной динамической функции с непрерывным временем, что позволяет прогнозировать во времени различные стадии развития процесса банкротства и, соответственно, разрабатывать упреждающие управленческие решения.

5. Предлагаемый алгоритм управления выездными налоговыми проверками отличается от известных методик тем, что очередность проверки объектов налогового контроля ставится в соответствие с синтезированным оптимальным планом отбора налогоплательщиков, что позволяет повысить объективность и достоверность управленческих решений в системе налогового контроля, и, в конечном итоге, увеличить собираемость налогов.

Предлагаемый алгоритм управления процессом кредитования отличается от известных тем, что в нем, на основе динамической нейросетевой модели, учитывается непрерывная зависимость вероятности риска банкротства заемщика от времени. Это позволяет принимать упреждающие управленческие решения по реструктуризации кредитной задолженности.

Перечисленные элементы научной новизны ранее не были известны и опубликованы в приоритетных публикациях автора.

Теоретическая ценность выносимых на защиту положений заключается в том, что они создают концептуальный базис нейросетевого моделирования в системах финансового контроля для сложных условий моделирования.

Практическая значимость работы заключается в том, что разработанные нейросетевые модели поддержки принятия решений в сфере финансового контроля позволяют обеспечить объективность и научную обоснованность управленческих решений. Автором получено 2 свидетельства о государственной регистрации электронных ресурсов в ОФЭРНиО, а также акт об использовании результатов работы в межрайонной налоговой инспекции №39 Республики Башкортостан. Научное исследование выполнено в 2009 – 2013 гг. в Финансовом университете при Правительстве Российской Федерации на кафедре математики и информатики, в рамках комплексной темы «Социально-экономическая политика России в условиях модернизации и перехода экономики на инновационный путь развития». В дальнейшем полученные в работе результаты могут служить научной базой для модернизации существующих в настоящее время систем финансового контроля.

Апробация работы. Результаты работы обсуждались на следующих конференциях: Всероссийская научно-техническая конференция «Нейроинформатика» (Москва, 2011-2014 гг.); Международная заочная научно-практическая конференция «Научная дискуссия: вопросы экономики и управления» (Москва, 2013 г.); Всероссийская научно-практическая конференция «Финансовая политика инновационного развития России: проблемы и пути решения» (Уфа, 2013 г.); Всероссийская научно-практическая конференция «Социальная ответственность бизнеса: теория, методология, практика» (Уфа, 2010-2011 гг.); Международная научно-практическая конференция «Экономико-математические методы исследования современных проблем экономики и общества» (Уфа, 2013 г.). Соискатель дважды принимал участие в Международном конкурсе научных работ студентов и аспирантов (Москва, 2012-2013 гг.). Получены дипломы I и II степени в номинации «Математические и инструментальные методы в экономике».

Публикации. Основные результаты, представленные в диссертации, опубликованы в 28 печатных работах; из них 6 статей – в научных журналах, рекомендованных ВАК; 18 работ – в трудах международных и всероссийских конференций; получено 2 свидетельства о регистрации алгоритмов. Автор диссертации участвовал в написании отдельных глав и параграфов в 2-х монографиях.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы из 111 наименований и 4 приложений. Работа содержит 175 страниц основного текста, включая иллюстрации и таблицы. В приложениях приводятся выборки данных, на основе которых были построены нейросетевые модели налогового контроля и диагностики банкротств.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность, излагаются цель и задачи исследования, раскрывается научная новизна и практическая ценность работы.

В первой главе проведен краткий анализ состояния проблемы применения нейрокомпьютеров в финансовой деятельности в России и за рубежом, выявлены сильные и слабые стороны нейросетевого инструментария при решении обратных задач восстановления скрытых в данных закономерностей. Изложена концепция «квазистатистики» и описаны особенности развиваемого системного подхода к нейросетевому моделированию сложных экономических систем. Раскрыта суть понятия «триада НЕ-факторов». Рассмотрена парадигма регуляризации обратных задач на основе теорий А. Н. Тихонова и Байеса и проведен анализ возможности ее применения при решении конкретных прикладных задач. В результате анализа были выявлены ряд вопросов, требующих углубленного изучения. Рассмотрен разработанный на основе общесистемных законов концептуальный базис нейросетевого моделирования сложных экономических объектов, функционирующих в условиях неопределенности.

Во второй главе проведен анализ устойчивости нейросетевых моделей, по результатам которого был сделан вывод о максимально допустимом уровне шума в данных и необходимости предобработки данных (предрегуляризации) при построении нейросетевых моделей. На модельном примере была показана эффективность и состоятельность процедур предрегуляризации модели в целях повышения однородности и информативности данных и повышения качества обучения нейросетей. Проведен анализ существующих методов регуляризации нейросетевых моделей при решении обратных задач восстановления функциональных зависимостей, скрытых в данных и дана их подробная классификация (рисунок 1).



Рисунок 1 - Классификация методов регуляризации нейросетевых моделей

Разработан оригинальный квазибайесовский метод регуляризации обучения нейросети при сильном зашумлении данных с неизвестной функцией закона распределения шума. Предложенный метод отличается от известного байесовского подхода тем, что он не использует априорную информацию о законе

распределения шумов в данных. В задаче налогового контроля для оценки апостериорных вероятностей гипотез - нейросетевых моделей о порождении данных $P(h_q(w, \bar{x}) | D | H)$ была предложена числовая мера «хорошего» объяснения данных:

$$\left. \begin{aligned} P(h_q(W, \bar{x}) | D | H) \rightarrow P_q^* = N_q^* / N, q = \overline{1, Q} \in \Omega^{test}; \\ N_q^* = \{i_q^*\}: \delta_{q, i_q^*} \leq \xi; q = \overline{1, Q}; i = \overline{1, N}, \end{aligned} \right\} \quad (1)$$

где: N_q^* – число «хороших» точек, в которых остатки $\{\delta_{q,i}\}$ имеют приемлемый уровень ошибки ξ для q -ой гипотезы – нейросети $\{h_q(\bar{X}, W)\}$; N – общее число примеров данных, одинаковое для всех гипотез-нейросетей; Ω^{test} – тестовое множество примеров; i_q^* – номер «хорошей» точки; ξ – экспертно задаваемый параметр.

Для задач диагностики банкротств в качестве критерия фильтрации априорных нейросетей-гипотез был предложен критерий, учитывающий качество идентификации банкротов нейросетью, а также количество ошибок 1-го и 2-го рода:

$$K_q = N_q^{(1)} r_1 + N_q^{(2)} r_2, \quad (3)$$

где: $N_q^{(1)}$ - количество ошибок 1-го рода; $N_q^{(2)}$ - количество ошибок 2-го рода; r_1, r_2 - удельный вес каждого показателя (коэффициенты Фишберна).

На основе сформулированных в главе 1 диссертации концепций был разработан модифицированный метод вложенных математических моделей (МВММ), отличительной особенностью которого является взаимосвязанное управление информативностью данных и качеством их аппроксимации в байесовском ансамбле нейросетей.

В третьей главе был проведен анализ существующих в настоящее время методов и моделей диагностики банкротств. Проведенный анализ позволил сделать следующие выводы:

1. Наиболее распространенные дискриминантные модели банкротств мало пригодны для российской экономики и нуждаются в коренной переработке и адаптации в рамках каждой отрасли.

2. Перспективным направлением усовершенствования моделей банкротства являются logit-модели. Перспективными методами оценки риска банкротств являются методы с нечетким описанием риска, нейросетевые, нейронечёткие, комплексные (агрегированные).

3. Малоисследованными вопросами при построении моделей банкротств остаются: формализованный выбор и количественное сравнение систем показателей (спецификаций) для построения моделей; разработка методов и алгоритмов сокращения размерности данных; разработка методов предобработки данных; регуляризация нейросетевых моделей в условиях сильного зашумления данных при неизвестном законе распределения шумов; учет непрерывного времени в динамике развития процесса кризиса на предприятии.

Был предложен нейросетевой метод оценки риска банкротств, в котором числовой мерой риска банкротства выбрана вероятность, получаемая по формуле logit-трансформации факторов \bar{X} .

$$P = [1 + e^{-\hat{Y}(\bar{X})}]^{-1} \quad (4)$$

Предложенный метод отличается от общеизвестных logit-моделей тем, что:

- 1) показатель экспоненты $\hat{Y}(\vec{X})$ вычисляется как нелинейное нейросетевое отображения факторов $\vec{X} = (X_1, X_2, \dots, X_j, \dots, X_n)$ в значение \hat{Y} на выходе нейросети:

$$\vec{X} \in \mathbb{R}^{(n)} \xrightarrow{F} \hat{Y} \in \mathbb{R}^1 \quad (5)$$

- 2) Значение $\hat{Y}(\vec{X})$ вычисляется как усредненное на отфильтрованном байесовском ансамбле нейросетей.

Далее был разработан формализованный метод выбора и сравнения между собой сложившихся систем показателей, основанный на байесовской методологии, суть которого заключается в следующем:

1. На основе данных бухгалтерской отчетности формируется несколько альтернативных систем показателей H_k (СП). Системы показателей $\{H_1, H_2, \dots, H_k, \dots, H_J\}$ отличаются между собой набором факторов: $H_k = \langle \vec{X}_j, Y \rangle, j = \overline{1, N_{H_k}}, k = \overline{1, G}$, где N_{H_k} - количество факторов (независимых переменных) x_j в системе показателей H_k ; G - количество альтернативных систем показателей.

2. Для выбора оптимальной СП строится вспомогательная нейросеть, которая обучается на разных наборах данных H_k . При этом используется один и тот же алгоритм обучения, одинаковое количество эпох обучения, фиксируется количество точек в обучающем и тестовом множестве.

3. Оценка качества полученных нейросетевых моделей проводится по результатам тестирования на данных, не участвовавших в обучении. Для окончательной оценки качества нейросетевых моделей предлагается использовать критерий, аналогичный критерию (3). Окончательный выбор системы показателей осуществляется по правилу:

$$H_{k_{\text{opt}}} : K|H_k = \min\{K\}, k \in \overline{1, G}. \quad (6)$$

То есть наилучшей признается СП, для которой величина (6) будет наименьшей.

Для обеспечения лучшего качества обучения нейросетей необходимо обеспечить достаточную однородность и информативность данных. Для выявления противоречивых наблюдений предлагается следующее правило: вектор-строка $\langle \vec{x}_i, y_i \rangle$ признается противоречивой, если хотя бы один компонент вектора \vec{x}_i нарушает правило:

$$x_{ij} \in [\bar{x}_j - l \cdot \sigma_{x_j}; \bar{x}_j + l \cdot \sigma_{x_j}], \quad (7)$$

где \bar{x}_j - среднее по выборке значение фактора x_j ; σ_{x_j} - выборочное среднеквадратическое отклонение фактора x_j ; l - экспертно задаваемое значение ($l = 1..4$). Противоречивые наблюдения удаляются из базы данных и не участвуют в обучении нейросети, но включаются в тестовое множество и используются для оценки качества обучения.

Также был разработан пошаговый метод сокращения размерности пространства факторов внутри выбранной системы факторов, который можно условно назвать «нейросетевым байесовским алгоритмом включения».

Предложенные идеи и методы были апробированы в вычислительных экспериментах. На основе данных бухгалтерской отчетности 136 российских строительных предприятий, были сформированы 4 альтернативные системы

показателей, которые используются в различных моделях диагностики банкротств: в пятифакторной модели Альтмана, моделях А. О. Недосекина, О. П. Зайцевой и Г. З. Рахимкуловой.

Известно, что из 136 предприятий 68 являются банкротами, а 68 – стабильно работающие предприятия. Количество записей в обучающих и тестовых множествах составило – 95 и 41 строк соответственно (70% и 30% от общего объема выборки). Перед обучением все данные были нормированы стандартным образом. Для выбора наилучшей системы показателей была создана и обучена вспомогательная нейросеть. Выбор наилучшей системы показателей осуществлялся по правилу (6) (см. таблицу 1).

Для отнесения предприятий к той или иной группе использовалось простейшее правило: если оценка вероятности банкротства была больше 0,5, то предприятие классифицировалось как банкрот, в противном случае – считалось платежеспособным. Наилучший результат (наименьшее количество ошибок и наименьшее значение $K_{H_k} = 1,33$) был достигнут для СП А. О. Недосекина.

Таблица 1. Результаты тестирования нейросети для различных СП

Система показателей	Доля верно распознанных предприятий	Ошибок 1-го рода	Ошибок 2-го рода	K_{H_k}
СП Недосекина	91,67%	1	2	1,33
СП Рахимкуловой	88,89%	2	2	2,0
СП Зайцевой	75,00%	3	6	4,0
СП Альтмана	86,11%	3	2	2,66

Далее, в соответствии с байесовским подходом, был создан ансамбль из 10 нейросетей, отличающихся количеством скрытых слоев, количеством нейронов в скрытых слоях и видом активационных функций. Для обучения нейросетей использовалась выбранная ранее система показателей. Результаты тестирования нейросетей ансамбля приведены в таблице 2.

Фильтрация нейросетевых моделей байесовского ансамбля проводилась по правилу: $K_{H_k} \leq \eta$. Был выбран уровень отбраковки $\eta = 4,5$. Таким образом, процедуру фильтрации прошли только 5 из 10 сетей байесовского ансамбля: НСМ1, НСМ3, НСМ4, НСМ7 и НСМ9. Для каждого предприятия, вошедшего в тестовое множество, было рассчитано среднее на отфильтрованном ансамбле значение отклика \hat{Y} . После этого были получены окончательные оценки вероятности банкротства для каждого предприятия по (4).

Таблица 2. Результаты тестирования нейросетей, обучавшихся на системе показателей А. О. Недосекина

№ сети	Доля верно распознанных предприятий	Количество ошибок 1-го рода	Количество ошибок 2-го рода	K_q
НСМ1	78,05%	4	5	4,33
НСМ2	80,49%	6	2	4,67
НСМ3	85,37%	4	2	3,33
НСМ4	85,37%	3	3	3,00
НСМ5	75,61%	5	5	5,00
НСМ6	53,66%	0	19	6,33
НСМ7	85,37%	5	1	3,67
НСМ8	78,05%	6	3	5,00
НСМ9	80,49%	2	6	3,33
НСМ10	75,61%	4	6	4,67

Доля верно распознанных предприятий составила 87,8% (36 из 41 предприятия тестового множества). Этот результат превосходит показатели каждой из нейросетей байесовского ансамбля в отдельности. Таким образом, мы получили убедительное свидетельство эффективности байесовского подхода.

В диссертации также была рассмотрена динамическая нейросетевая модель диагностики банкротств:

$$P = [1 + e^{-\hat{Y}(\bar{x}, t)}]^{-1}. \quad (8)$$

Вычислительные эксперименты по построению динамической модели (8) проводились на той же выборке данных 136 строительных предприятий. При этом точность выявления предприятий-банкротов оказалась чуть ниже, чем для модели (4) – 85,9% предприятий, вошедших в тестовое множество были распознаны верно.



Рисунок 2 - Алгоритм принятия решений при управлении кредитным процессом

Разработанные нейросетевые модели диагностики банкротств могут использоваться в кредитных организациях при управлении кредитным процессом.

В четвертой главе рассмотрены методы синтеза плана выездных налоговых проверок. Основой метода построения НСМ налогового контроля является модифицированный метод вложенных математических моделей (МВММ), подробно описанный в главе 2 диссертации. Этот метод позволяет восстановить скрытую в данных зависимость моделируемого индикатора нарушения налогового законодательства Y , от ряда экономических показателей. При составлении плана проверок должны соблюдаться следующие принципы:

1) план проверок формируется после подачи налогоплательщиками последней налоговой декларации (квартального отчета) в момент времени t_0 . Появление значительных относительных отклонений вида

$$\delta_i = |(y_i - \hat{y}_i) / y_i| \cdot 100\%, \quad i = \overline{1, N} \quad (9)$$

может свидетельствовать о нарушении налогоплательщиком налогового законодательства и является основанием для включения его в план проверок;

2) большие отклонения δ_i на момент составления плана проверок могут быть вызваны и объективными причинами (неудачная сделка, повлекшая за собой убытки, форс-мажор и т.д.). Поэтому необходимо учитывать предысторию появления у данного налогоплательщика таких отклонений и оценить вероятность нарушения данным налогоплательщиком налогового законодательства;

3) проведение выездных проверок в первую очередь преследует цель получения максимальных налоговых доначислений. Очевидно, что величина возможных доначислений зависит от масштаба предприятия, который нужно учитывать при составлении плана проверок.

Для синтеза оптимального плана выездных проверок использовался метод отбора, основанный на ранжировании предприятий-налогоплательщиков по Ψ -критерию, учитывающему косвенно величину ожидаемых доначислений и их вероятность. При этом следует учитывать риск неверной идентификации налогоплательщика, т.е. ошибки I и II рода.

Было предложено 2 способа вычисления Ψ -критерия, отличающиеся подходом к оценке вероятности нарушения налогового законодательства.

Модификация 1. Оценка вероятности нарушения осуществляется через оценку фрактальной размерности временного ряда относительных отклонений $\delta_g(t)$, где номер предприятия-налогоплательщика g фиксирован. Критерий ранжирования ψ будет вычисляться следующим образом:

$$\Psi_g = |\delta_g|_{t=t_0} \cdot D^{-1} \cdot M_g; \quad g = \overline{1, G}. \quad (10)$$

Величина $|\delta_g|_{t=t_0}$ в формуле (10) – это абсолютная величина погрешности, вычисленная по (9) на момент подачи налогоплательщиком последней декларации t_0 ; D - величина фрактальной размерности временного ряда $\{\delta_{g,t}\}$; M_g - коэффициент масштаба g -го налогоплательщика; G - количество налогоплательщиков. Таким образом, критерий (10) учитывает все принципы составления плана проверок, указанные выше.

Оценивать вероятность нарушения налогового законодательства предлагается на основе фрактальной размерности D временных рядов $\{\delta_{g,t}\}$. Если уров-

ни ряда $\{\delta_{g,t}\}$ независимы, то величина D будет стремиться к величине топологической размерности плоскости, то есть $D \rightarrow 2$. В этом случае вероятность систематического нарушения налогового законодательства крайне мала, т.к. уровни ряда $\{\delta_{g,t}\}$ не коррелированы. Если же значения ряда $\{\delta_{g,t}\}$ не являются независимыми, то величина D будет значительно меньше 2, и это является сигналом о возможном систематическом искажении данных налоговых деклараций. Отношение $1/D$ условно трактуется как оценка вероятности нарушения налогового законодательства налогоплательщиком.

Существует несколько способов определения фрактальной размерности временного ряда. В диссертации рассмотрены два из них: метод Херста и метод, основанный на вычислении площади минимального покрытия. Однако, использовать критерий (10) можно только в случае достаточно длинных временных рядов $\{\delta_{g,t}\}$ (несколько десятков наблюдений), что на практике встречается редко.

Модификация 2. Оценка вероятности нарушения осуществляется по формуле логит-трансформации:

$$P_g = [1 + e^{-\delta_g}]^{-1}, \quad (11)$$

где δ_g - относительное отклонение декларируемой величины Y от моделируемой величины \hat{Y} для g -го налогоплательщика; регрессионная зависимость $Y(\vec{X})$ от вектора объясняющих переменных \vec{X} строится как нейросетевое отображение $\vec{X} \in R^n \rightarrow Y \in R^1$.

В данном случае ψ критерий вычисляется по (10), где вместо D^{-1} следует подставить P_g из (11). Данная модификация метода отбора является более общей по сравнению с модификацией 1. Этот метод не стеснен практическими никакими серьезными ограничениями, относительно закона распределения или длины временных рядов $\{\delta_{g,t}\}$.

Приведем результаты вычислительных экспериментов по синтезу оптимальных планов выездных налоговых проверок. По данным квартальных налоговых деклараций 24 сельскохозяйственных предприятий за период с 2006 по 2009 гг. была сформирована исходная база данных $D = \{Y_i, \vec{X}_i\}_{i=1}^{i=N}$, включающая в себя 9 факторов x_j и 276 вектор-строк наблюдений. Был создан ансамбль из 12 гипотез-нейросетей. После обучения сетей, согласно байесовскому подходу к регуляризации, была проведена их фильтрация. Процедуру фильтрации прошли 6 сетей. На основе 6 нейросетей, прошедших процедуру фильтрации были составлены планы выездных проверок по правилу (10) (две модификации) и правилу (11). Окончательные планы выездных проверок получили путем ранжирования всех предприятий по осредненному критерию $\bar{\psi}_g$ на отфильтрованном байесовском ансамбле нейросетей. В таблице 3 приведены оптимальные планы выездных проверок, составленные тремя вышеуказанными способами и величина налоговых доначислений, полученных по результатам реальных выездных проверок для анализируемой группы предприятий. В таблице указаны первые 10 номеров предприятий согласно оптимальным планам и 10 номеров предприятий с наибольшими доначислениями.

Из таблицы 3 видно, что из 10 предприятий, по которым были получены наибольшие доначисления, в окончательные планы проверок №1 и №2 попало 7 предприятий. В план №3 попало 6 предприятий из числа 10 самых крупных нарушителей. Таким образом, ГНСМ налогового контроля достаточно точно идентифицирует нарушителей налогового законодательства и может существенно повысить эффективность работы налоговой службы в части налогового контроля.

Таблица 3. Сравнение результатов моделирования с итогами выездных налоговых проверок

Очередность проведения проверки	План проверок №1 на основе фрактальной размерности Херста D_H		План проверок №2 на основе размерности минимального покрытия D_μ		План проверок №3 на основе логит-трансформации вероятности нарушения		Результаты реальных налоговых проверок	
	№ предпр.	$\bar{\psi}_g$	№ предпр.	$\bar{\psi}_g$	№ предпр.	$\bar{\psi}_g$	№ предпр.	Доначисления, руб.
1	11	0,2081	11	0,1903	11	0,1208	11	7002076
2	22	0,0646	9	0,0589	22	0,0556	9	3140252
3	9	0,0629	22	0,0578	9	0,0270	8	2780893
4	14	0,0378	14	0,0349	14	0,0226	24	1911636
5	2	0,0263	2	0,0253	2	0,0176	21	1700531
6	3	0,0240	21	0,0228	21	0,0174	22	1520399
7	21	0,0225	3	0,0210	3	0,0145	16	1427865
8	8	0,0198	8	0,0193	30	0,0109	10	1384382
9	24	0,0178	24	0,0164	8	0,0103	13	1371761
10	30	0,0141	30	0,0129	12	0,0071	3	1366995

Сумма налоговых доначислений 10-ти крупнейших нарушителей (сумма первых 10 строк в последнем столбце таблицы 3) составляет 23606790 руб. Если предположить, что выездные проверки проводились бы по сформированным планам, то сумма полученных доначислений составила бы:

- Для планов №1 и №2 – 21715626 руб., что составляет 92% от максимально возможной величины доначислений.
- Для плана №3 – 20300226 руб., что составляет 86% от максимально возможной величины доначислений.

Таким образом, можно сделать вывод, что НСМ налогового контроля позволяет достаточно достоверно выявлять нарушителей налогового законодательства. Результаты моделирования хорошо согласуются с результатами реальных налоговых проверок.

В настоящее время в территориальных налоговых инспекциях используется система электронной обработки данных (ЭОД), предназначенная для автоматизации налогового администрирования. Она представляет собой единую интегрированную систему, объединяющую все направления деятельности инспекции (регистрация физических и юридических лиц, учет налогоплательщиков, камеральные и выездные проверки и т.д.).

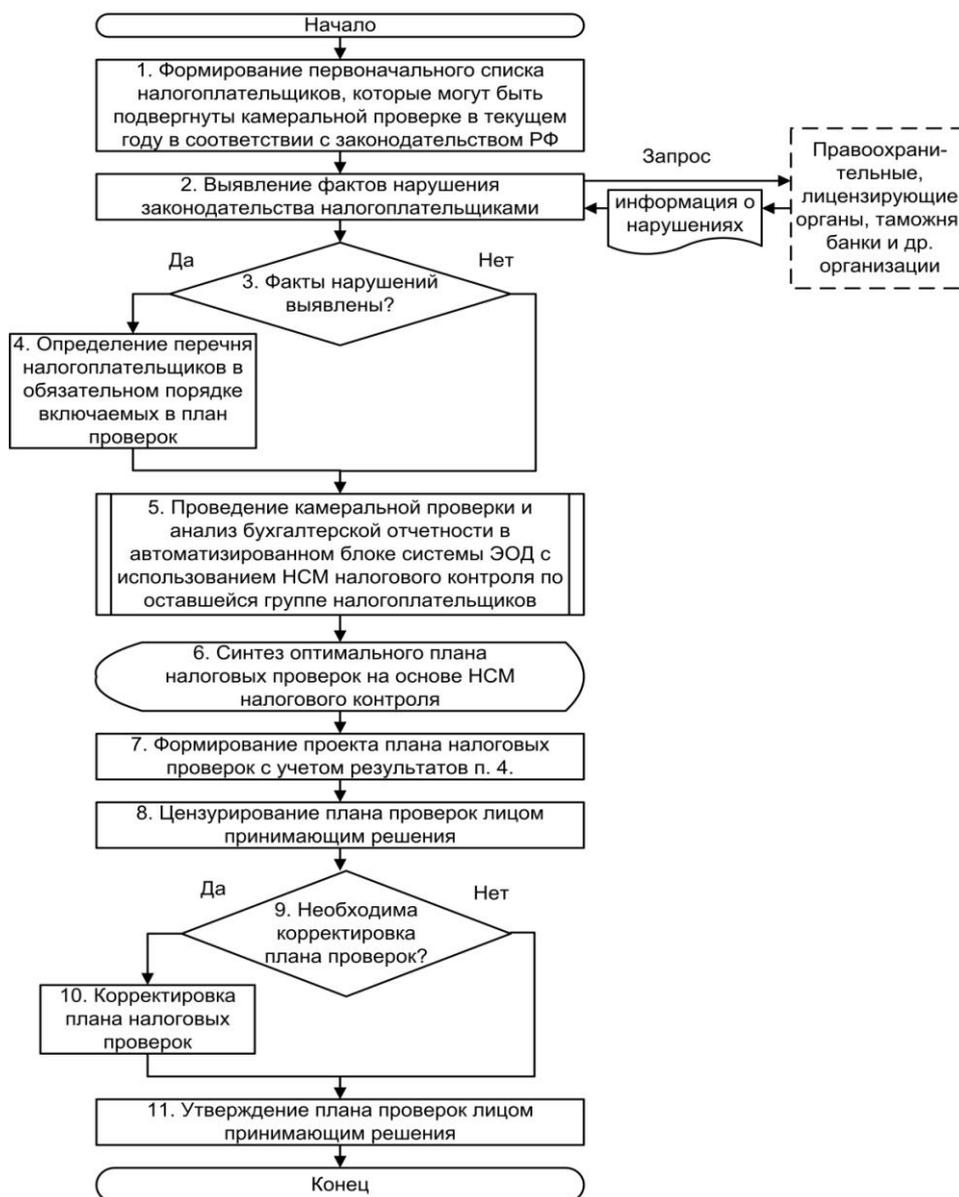


Рисунок 3 - Алгоритм принятия решений при составлении плана налоговых проверок

Разработанная нейросетевая модель налогового контроля может быть встроена в ЭОД в качестве автоматизированного аналитического блока (рисунок 3). Это позволит автоматизировать процесс составления плана налоговых проверок и тем самым сократить время принятия решений.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Разработан модифицированный метод вложенных математических моделей (МВММ), содержащий алгоритмы предобработки данных и алгоритмы их аппроксимации с помощью байесовского ансамбля нейросетей, отличающийся тем, что управление информативностью данных и качеством их аппроксимации производится взаимосвязано для обеспечения состоятельности задачи регуляризации НСМ. В итоге, в НСМ достигается оптимальный компромисс между ошибкой обобщения и устойчивостью в условиях триады НЕ-факторов (неточность, неполнота, неопределенность в данных). Достоверность МВММ обоснована вычислительными экспериментами, проведенными на реальных данных и данными производственных экспериментов.

2. Разработан гибридный метод синтеза оптимальных планов отбора налогоплательщиков для выездных налоговых проверок, основу которого составляет МВММ. Метод отличается тем, что учитывается не только величина индикатора нарушения налогового законодательства, но и вероятность этого нарушения: а) для «длинных» временных рядов указанного индикатора – через оценку фрактальной размерности; б) для «коротких» временных рядов – через логистическую трансформацию вероятности. Метод позволяет создать предпосылки для повышения объективности отбора налогоплательщиков для выездных проверок, а также снижения затрат на налоговый контроль. Достоверность метода подтверждается вычислительными экспериментами, а также сравнением результатов моделирования с результатами реальных налоговых проверок.

3. Разработан многоступенчатый метод оценки и управления адекватностью нейросетевых моделей финансового контроля в условиях триады «НЕ-факторов». Предложенный метод основан на принципе последовательного сужения класса искомых решений обратной задачи восстановления скрытых в данных зависимостей:

1) при априорном задании гипотез-нейросетей $\{h_q\}$, $q = \overline{1, Q}$ с учетом эвристической информации об архитектуре нейросетей и передаточных функциях промежуточных слоев (оптимизируется длина описания данных);

2) при апостериорной фильтрации гипотез-нейросетей $\{h_q\}$ по критерию качества объяснения данных;

3) при осреднении (прямом или весовом) расчетных выходных характеристик моделируемой системы финансового контроля на отфильтрованном ансамбле нейросетевых моделей;

4) при сравнении выходных характеристик с данными натуральных (производственных) экспериментов.

Такой подход позволяет добиться адекватности НСМ в сложных условиях моделирования, отягченных дефицитом наблюдений.

4. Предложен метод разработки модели поддержки принятия решений по ситуационному управлению кредитным процессом с логистической трансформацией вероятности риска банкротства. Метод отличается тем, что показатель экспоненты логистической функции агрегируется с помощью байесовского ансамбля нейросетей в виде многомерной нелинейной динамической функции с непрерывным временем, что позволяет прогнозировать во времени различные стадии развития процесса банкротства и, соответственно, разрабатывать упреждающие управленческие решения. Достоверность метода подтверждается вычислительными экспериментами: для выборки из 136 предприятий точность идентификации банкротов составила 88%.

5. Предложены алгоритмы управления выездными налоговыми проверками и процессом кредитования клиентов банков, отличающиеся тем, что принятие управленческих решений основывается на результатах нейросетевого моделирования объектов управления. Внедрение предложенных алгоритмов на практике позволит повысить объективность и достоверность принимаемых решений, а значит, повысить эффективность соответствующих систем финансового контроля.

Перспективы дальнейшей разработки темы: комбинация разработанных подходов с методами нечетких множеств и нечетко-интервальной матема-

тики; определение границ применимости разработанных методов; обобщение аппарата описания риска на случай обработки качественной информации, а также агрегирования информации различной природы; более широкая апробация разработанных методов на натуральных (производственных) экспериментах; математическое обоснование предложенных итерационных процедур предобработки данных; встраивание в существующие экспертные системы (ЭС) в системах налогового администрирования, банковского кредитования, корпоративного мониторинга разработанных на базе НСМ интеллектуальных ЭС.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Всего по теме диссертации опубликовано 28 работ, основные из них:

В рецензируемых журналах из списка ВАК

1. О состоятельности процедур регуляризации нейросетевых моделей налогового и финансового контроля / С. А. Горбатков, И. И. Белолипец, А. О. Солнцев // Информационные технологии: теоретический и прикладной научно-практический журнал. 2013. №11(207). С. 64-68.

2. Выбор системы экономических показателей для диагностики и прогнозирования банкротств на основе байесовского подхода и интеллектуальных информационных технологий / С. А. Горбатков, И. И. Белолипец, Е. Ю. Макеева // Вестник Финансового Университета при Правительстве РФ. 2013. №4(76). С. 50-62.

3. Удаление противоречивых наблюдений как процедура предрегуляризации нейросетевой модели налогового контроля / С. А. Горбатков, И. И. Белолипец, Е. А. Мурзина // Вестник УГАТУ. Сер. Управление, вычислительная техника и информатика. 2013. №5 (58). С. 110-114.

4. О моделях диагностики банкротств организаций / Е. Ю. Макеева, С. А. Горбатков, И. И. Белолипец // Менеджмент и бизнес-администрирование. 2014. №1 . С. 151-172.

5. Гибридный метод оценки риска банкротства на базе байесовского ансамбля нейросетей и logit-модели / С. А. Горбатков, И. И. Белолипец // Интернет журнал «Наукоедение». 2013. №6 (19). URL: <http://naukovedenie.ru/PDF/25EVN613.pdf> (дата обращения 22.04.2014).

6. Построение оптимального плана выездных проверок на основе гибридной нейросетевой модели налогового контроля / И. И. Белолипец // Управление экономическими системами: электронный научный журнал. 2013. №4 (52). URL: <http://www.uecs.ru/uecs-52-522013/item/2143-2013-05-17-07-42-26> (дата обращения 22.04.2014).

В других изданиях

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

8. С. А. Горбатков, Д. В. Полупанов, Е. Ю. Макеева, А. Н. Бирюков Методологические основы разработки нейросетевых моделей экономических объектов в условиях неопределенности. / Под ред. д.т.н., профессора Горбаткова С. А.: Монография. – М.: Издательский дом «Экономическая газета», 2012. – 494

с. (Белолипец И. И. принимал участие в написании главы II и параграфов 3.5-3.9)

9. Приближенный метод байесовской регуляризации и двухступенчатая оценка адекватности гибридной нейросетевой модели налогового контроля / Горбатков С. А., Белолипец И. И., Фархиева С. А. // Нейроинформатика: сб. науч. тр. XIII всеросс. конф. М.:МИФИ, 2011, Т. 3. С. 144-154.

10. Оценка эффективности алгоритмов предрегуляризации и байесовской регуляризации нейросетей для камеральной налоговой проверки / Горбатков С. А., Белолипец И. И., Фархиева С. А., Полупанов Д. В. // Нейроинформатика: сб. науч. тр. XIV всеросс. конф. М.:МИФИ, 2012, Т. 3. С. 38-48.

11. Использование методов фрактальной теории при ранжировании объектов налогового контроля / Горбатков С. А., Белолипец И. И., Фархиева С. А. // Нейроинформатика: сб. науч. тр. XV всеросс. конф. М.:МИФИ, 2013, Т. 3. С. 184-192.

12. Обобщенная нейросетевая модель риска банкротства предприятия / Горбатков С. А., Фархиева С. А., Белолипец И. И. // Нейроинформатика: сб. науч. тр. XVI всерос. конф. М.:МИФИ, 2014, Т. 3. С. 210-221.

13. Использование моделей бинарного выбора в задачах диагностики банкротств / Горбатков С. А., Белолипец И. И. // Финансовая политика инновационного развития России. Проблемы и пути решения: сб. науч. тр. всеросс. науч.- практ. конф. Уфа, 2013, С. 161-166.

14. Гибридный нейросетевой метод динамической оценки банкротства предприятий / И. И. Белолипец, С. А. Горбатков, С. А. Фархиева // Экономико-математические методы исследования современных проблем экономики и общества: сб. мат. Межд. науч.-практ. конф. Уфа, 2013, С. 292-300.

15. Очистка данных наблюдений как процедура предрегуляризации нейросетевой модели налогового контроля / Горбатков С. А., Белолипец И. И. // Социальная ответственность бизнеса: теория, методология, практика: мат. II всеросс. науч.- практ. конф. Уфа: ВЗФЭИ, 2011, С. 114-118.

16. Использование методов фрактальной теории при проведении выездных проверок / Горбатков С. А., Белолипец И. И. // Социальная ответственность бизнеса: теория, методология, практика: мат. всеросс. науч.- практ. конф. Уфа: ВЗФЭИ, 2011, С. 89-92.

17. К вопросу об устойчивости нейросетевой модели налогового контроля / Белолипец И. И. // Управление. Экономический анализ. Финансы: сб. науч. тр. / под. общ. ред. Н. К. Зайнашева. Уфа: УГАТУ, 2012. С. 233-239.

18. Свид. о гос. рег. электронного ресурса № 19722. Алгоритм построения гибридной нейросетевой модели восстановления скрытых в данных закономерностей в технических и экономических системах / С. А. Горбатков, И. И. Белолипец, М.: Зарег. в ИНИПИ РАО ОФЭРНиО 2 декабря 2013.

19. Свид. о гос. рег. электронного ресурса № 19720. Алгоритм поиска и удаления из данных противоречивых наблюдений / С. А. Горбатков, И. И. Белолипец, М.: Зарег. в ИНИПИ РАО ОФЭРНиО 2 декабря 2013.

Диссертант



Белолипец И. И.

БЕЛОЛИПЦЕВ Илья Игоревич

УПРАВЛЕНИЕ В СИСТЕМАХ ФИНАНСОВОГО КОНТРОЛЯ
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 в печать 07.07.14 г. Формат 60x84 1/16.
Бумага офсетная. Печать ризографическая. Тираж 100 экз.
Гарнитура «TimesNewRoman». Отпечатано в типографии
«ПЕЧАТНЫЙ ДОМЪ» ИП ВЕРКО.
Объем 1 п.л. Уфа, Карла Маркса 12 корп. 5,
т/ф: 27-27-600, 27-29-123