



**С.А. Горбатков, Д.В. Полупанов,  
Е.Ю. Макеева, А.Н. Бирюков**

**Методологические основы  
разработки нейросетевых моделей  
экономических объектов  
в условиях неопределенности**

**С.А. Горбатков, Д.В. Полупанов,  
Е.Ю. Макеева, А.Н. Бирюков**

**Методологические основы  
разработки нейросетевых моделей  
экономических объектов  
в условиях неопределенности**

**Москва 2012**

УДК 336.1  
ББК 65.261.3  
ISBN 978-5-905735-36-3  
2012. 494с.

### **Рецензенты:**

СПИВАК С.И. доктор физ.-мат. наук, профессор, зав. кафедрой  
математического моделирования Башкирского государственного университета

БЛАЖЕНКОВА Н.Н., д.э.н., профессор

### **Горбатков С.А., Полупанов Д.В., Макеева Е.Ю., Бирюков А.Н.**

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. – 494с.

### **Аннотация**

Монография посвящена общим принципам разработки эффективных нейросетевых математических моделей экономических объектов, характеризующихся сложными условиями моделирования: сильным зашумлением данных вплоть до их сознательного искажения, отягченным дефицитом наблюдений; неизвестным априори законом распределения шумов; стремлением авторов к достаточно общим постановкам задач восстановления многомерных зависимостей, скрытых в данных, ранжирования объектов и их кластеризации, многофакторного прогноза. Учитывается триада “НЕ-факторов”: неполнота, неточность, неопределенность. Изложение ведется в основном на примере задач налогового администрирования и оценки кредитоспособности заемщиков, где указанные условия моделирования проявляются особенно ярко. Однако разработанный на основе общесистемных законов концептуальный базис нейромоделирования применим для более широкого круга экономических объектов. Особое внимание уделено регуляризации нейромоделей на основе байесовского подхода и управления их адекватностью в сложных условиях моделирования. Рассмотрены также модели прогнозирования вероятности банкротства экономических объектов. Монография обобщает опыт многолетних исследований авторов по указанной проблематике.

Издание предназначено для научных работников, преподавателей, аспирантов и магистрантов экономических вузов.

© Горбатков С.А., Полупанов Д.В.,  
Макеева Е.Ю., Бирюков А.Н., 2012

## СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ.....	8
ВВЕДЕНИЕ.....	11
1. Актуальность темы исследований, их цель, решаемые задачи.....	11
2. Методологические особенности развиваемого системного подхода к разработке концептуального базиса моделирования.....	14
3. Краткое содержание глав.....	16
ГЛАВА I. СИСТЕМНЫЙ ПОДХОД К МОДЕЛИРОВАНИЮ СИСТЕМЫ НАЛОГОВОГО АДМИНИСТРИРОВАНИЯ.....	19
1.1. Описание системы налогового администрирования и взаимосвязь моделей ее подсистем.....	19
1.2. Комбинированный системно-синергетический-информационный подход к исследованию сложных математико-информационных систем.....	30
1.2.1. Условия моделирования.....	30
1.2.2. Сущность системного синергетического информационного подхода.....	32
ГЛАВА II. РАЗРАБОТКА КОНЦЕПТУАЛЬНОГО БАЗИСА МОДЕЛИРОВАНИЯ НА ОСНОВЕ ОБЩЕСИСТЕМНЫХ ЗАКОНОВ КИБЕРНЕТИКИ.....	37
2.1. Концепция построения модели «обобщенной производственной функции» для кластера налогоплательщиков на основе закона энтропийного равновесия и фонового закона.....	37
2.1.1. Закон энтропийного равновесия в открытой информационной системе.....	37
2.1.2. Фоновая закономерность.....	40
2.1.3. Формулировка концепции сравнения показателей состояния предприятий-налогоплательщиков с «эталонным» фоном.....	45
2.2. Концепция 3 байесовской регуляризации обучения нейросетей на основе закона неполного подавления дисфункций структурируемой системы.....	49
2.2.1. Закон появления дисфункций при структурировании информационной системы и необходимость их парирования.....	49
2.2.2. Концепция регуляризации нейросетевой модели на основе байесовского ансамбля нейросетей.....	51
2.3. Концепция взаимосвязанного управления информативностью данных и качеством их аппроксимации в нейросети для обеспечения состоятельности задачи регуляризации на основе закона роста и убывания энтропии.....	54
2.3.1. Проблема обеспечения состоятельности задачи регуляризации.....	54
2.4. Концепция 5 учета неопределенности в оценке отклонений показателей налогоплательщиков от эталонной поверхности на основе вероятностного и фрактального подходов.....	59

2.4.1. Двойственная природа больших отклонений как результат неполного подавления дисфункций структурируемой математико-информационной системы.....	59
2.4.2. Вероятностный подход к оценке неопределенности в задаче ранжирования налогоплательщиков .....	61
2.4.3. Фрактальный подход к оценке неопределенности в задаче ранжирования налогоплательщиков .....	62
2.4.4. Формулировка концепции синтеза оптимального плана отбора налогоплательщиков для выездных проверок .....	65
<b>ГЛАВА III. МЕТОДЫ ПОСТРОЕНИЯ НЕЙРОСЕТЕВОЙ МОДЕЛИ ВОССТАНОВЛЕНИЯ РЕГУЛЯРНОЙ ЧАСТИ МНОГОМЕРНЫХ ФУНКЦИЙ И ЕЕ РЕГУЛЯРИЗАЦИИ, РЕАЛИЗУЮЩЕЙ КОНЦЕПТУАЛЬНЫЙ БАЗИС ИЗ ГЛАВЫ II .....</b>	<b>68</b>
3.1. Вычислительные модельные эксперименты по анализу устойчивости нейросетевых моделей.....	68
3.1.1. Методика вычислительных экспериментов по устойчивости НСМ.....	68
3.1.2. Обсуждение вычислительного эксперимента. Обобщения. Парадигма парирования больших возмущений в данных.....	72
3.2. Корректная постановка обратных задач восстановления многомерных функций (задач интерпретации) и понятия регуляризации по А.Н.Тихонову.....	79
3.2.1. Основные понятия и определения.....	79
3.2.2. Введение обратных задач интерпретации в класс условно корректных по Тихонову .....	83
3.2.3. Понятие квазирешения.....	87
3.2.4. Принцип регуляризации для задач интерпретации.....	88
3.2.5. Задача о квазиминимизации и регуляризирующий алгоритм для задач интерпретации.....	89
3.2.6. Понятие стабилизатора А.Н. Тихонова и связанные с ним обобщенно-корректные постановки обратных задач.....	91
3.3. Использование сглаживающего функционала Тихонова для задачи восстановления многомерных нелинейных функций в MLP-сетях и RBF-сетях.....	94
3.3.1. Практический алгоритм регуляризации MLP-нейросетей в задаче восстановления.....	97
3.3.2. Регуляризация RBF-сетей.....	99
3.4. Байесовский подход к регуляризации нейросетевых моделей аппроксимации нелинейных функций многих переменных.....	112
3.4.1. Постановка задачи и её особенности.....	112
3.4.2. Обучение по Байесу. Основные понятия. Регуляризация обучения.....	116
3.4.3. Связь между ошибкой обобщения и минимальной длиной описания в байесовском подходе .....	123

3.4.4. Итерационный процесс байесовского обучения (EM – алгоритм).....	126
3.5. Прикладной приближенный метод байесовской регуляризации обучения нейросети при сильном зашумлении данных с неизвестной функцией распределения шума.....	134
3.6. Модифицированный метод вложенных математических моделей (МВММ), основанный на взаимосвязанном управлении информативностью данных и качеством их аппроксимации в байесовском ансамбле нейросетей.....	140
3.6.1. Логическая схема метода.....	140
3.6.2. Алгоритм I.1 априорного экспертного выбора типа нейросети, её архитектуры, активационных функций и алгоритма обучения.....	146
3.6.3. Алгоритм I.2 квазиоптимальной процедуры «дообучения экспертов» при спецификации переменных.....	150
3.6.4. Алгоритм нормировки и фильтрации переменных по их статистической значимости.....	156
3.6.5. Алгоритм I.4 удаления из данных противоречивых вектор-строк (примеров) наблюдений по критерию Липшица.....	168
3.6.6. Алгоритм I.5 оптимальной итерационной кластеризации данных с использованием последовательности НССМ.....	170
3.6.7. Алгоритм I.6 оптимальной итерационной очистки образованных кластеров данных от аномальных точек с использованием вспомогательных НССМ.....	175
3.6.8. Алгоритм I.7 «ремонта» вектор-столбцов данных в кластерах.....	178
3.6.9. Алгоритмы II.1 – II.6 для подмодели II в рамках МВММ-метода.....	180
3.7. Вычислительные эксперименты по апробации метода ВММ.....	181
3.7.1. Первоначальная версия алгоритма кластеризации, учитывающая итоговый показатель качества разбиения по нейросети.....	181
3.7.2. Вычислительные эксперименты по алгоритму I.5 оптимальной итерационной кластеризации данных (серия 3).....	202
3.7.3. Вычислительные эксперименты по алгоритму I.6 оптимальной итерационной очистки образованных кластеров данных от аномальных точек (серия 4).....	206
3.8. Метод синтеза оптимального плана выездных налоговых проверок.....	238
3.8.1. Алгоритм модификации I ранжирования налогоплательщиков на основе предположения о нормальном законе распределения отклонений.....	248
3.8.2. Алгоритм модификации II ранжирования налогоплательщиков на основе фрактального подхода.....	256
3.8.3. Оценка адекватности модели СПВМ и всей гибридной нейросетевой модели.....	262
3.9. Вычислительные и натурные эксперименты по верификации метода синтеза оптимального плана выездных проверок.....	265

3.9.1. Вычислительные эксперименты.....	265
3.9.2. Верификация нейросетевых моделей на основе натуральных экспериментов.....	269
3.9.3. Оценка адекватности ГНСМ путем ее сравнения с альтернативной непараметрической моделью Estimation Tax.....	271
<b>ГЛАВА IV. ПРИКЛАДНЫЕ НЕЙРОСЕТЕВЫЕ МОДЕЛИ И КОМПЬЮТЕРНЫЕ МЕТОДИКИ В СИСТЕМЕ НАЛОГОВОГО АДМИНИСТРИРОВАНИЯ.....</b>	<b>272</b>
4.1. Разработка структурно-функциональной схемы интеллектуальной информационно-аналитической системы поддержки принятия решений по налоговому администрированию (СНА).....	272
4.1.1. Методические основы разработки СНА и места в ней ГНСМ.....	272
4.1.2. Описание работы структурно-функциональной схемы СНА.....	277
4.2. Разработка прикладной гибридной нейросетевой модели для информационной системы налогового контроля.....	284
4.3. Разработка прикладной гибридной нейросетевой модели для подсистемы налогового планирования.....	292
4.3.1. Спецификация и база данных для гибридной многофакторной прогнозной модели доходной части бюджета МО.....	292
4.3.2. Вычислительные эксперименты по формированию «русла» путем оптимизации архитектуры нейросети.....	297
4.3.3. Расчет коэффициента риска выхода ошибки прогноза за назначенные границы прогноза.....	309
4.3.4. Обобщающие выводы по вычислительным экспериментам формирования «русла».....	312
4.3.5. Обоснование адекватности НСМ путем байесовской апостериорной фильтрации нейросетей – гипотез в ансамбле.....	318
4.3.6. Компьютерная методика прогноза наполнения муниципального бюджета с учетом риска выхода ошибки прогноза за назначенный уровень.....	321
4.3.7. Факторный анализ доходной части бюджета по нейросетевой модели.....	335
4.4. Метод кластеризации на основе нейросетей Кохонена с селекцией признаков и байесовской регуляризацией. Прикладная нейросетевая модель кластеризации налогоплательщиков для поддержки принятия решений по налоговому регулированию.....	342
4.4.1. Предварительные замечания.....	342
4.4.2. Постановка задачи нейросетевого моделирования.....	344
4.4.3. Метод кластеризации на основе нейросетей Кохонена с селекцией признаков и байесовской регуляризацией.....	347
4.4.4. Нейросетевой метод самоорганизующихся карт Кохонена.....	360
4.4.5. Результаты кластеризации.....	361
4.5. Гибридный метод оценки кредитоспособности (ГМОК) для разработки модели поддержки принятия решений по налоговому регулированию с использованием нечетких правил вывода.....	375

4.5.1. Идея метода.....	375
4.5.2. Модификация а) ГМОК с оперированием всех объясняющими переменными.....	376
4.5.3. Модификация б) ГМОК с оперированием одним обобщенным критерием.....	383
4.6. Бинарные модели прогнозирования вероятности банкротства.....	383
4.6.1. Направления исследований оценки вероятности банкротства компаний.....	383
4.6.2. Методологическая основа исследования прогнозирования банкротства на развивающихся рынках.....	388
4.6.3. Методика моделирования вероятности банкротства на развивающихся рынках.....	393
4.6.4. Выводы и рекомендации для дальнейшего исследования прогнозирования вероятности банкротства.....	404
ЗАКЛЮЧЕНИЕ.....	406
ЛИТЕРАТУРА.....	410
СПИСОК ЧАСТО ВСТРЕЧАЮЩИХСЯ УСЛОВНЫХ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ.....	422
ПРИЛОЖЕНИЯ.....	426
Приложение 1. Два подхода к построению нейросетевой модели – аппроксимация и интерполяция функций.....	426
Приложение 2. Исходная выборка для торговых предприятий г. Уфы по Орджоникидзевскому району (номера предприятий закодированы числами).....	438
Приложение 3. Исходные данные для построения гибридной нейросетевой модели ранжирования налогоплательщиков на выборках $Z^I$ и $Z^{II}$ .....	444
Приложение 4. Выборка данных образовательных учреждений г. Стерлитамака.....	458
Приложение 5. База данных к задаче прогнозирования наполнения бюджета за счет налоговых и неналоговых поступлений одного из МО г. Стерлитамака.....	480
Приложение 6. Корреляционные матрицы финансовых коэффициентов, используемых при прогнозировании банкротства.....	485



