

0-793327

На правах рукописи

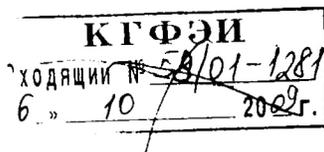
Гречин Сергей Борисович

**РАЗРАБОТКА МЕТОДОВ И ИНСТРУМЕНТАЛЬНЫХ
СРЕДСТВ ПРОГНОЗИРОВАНИЯ
ЭКОНОМИЧЕСКОГО СОСТОЯНИЯ ПРЕДПРИЯТИЙ
НА ОСНОВЕ ГЕНЕТИЧЕСКОГО
ПРОГРАММИРОВАНИЯ**

Специальность 08.00.13 –
Математические и инструментальные методы экономики

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата экономических наук

Нижний Новгород – 2009



Работа выполнена на кафедре «Экономической информатики» в ГОУ ВПО Нижегородский государственный университет им. Н.И. Лобачевского.

Научный руководитель: д.э.н., профессор,
Трифонов Ю.В.

Официальные оппоненты: доктор экономических наук,
профессор Ромашова Ирина
Борисовна;
кандидат экономических наук,
доцент Павленков Михаил
Николаевич

Ведущая организация: ГОУ ВПО «Нижегородский
государственный технический
университет им Р.Е. Алексеева»

Защита состоится 29 октября 2009 г. в 12 часов на заседании диссертационного совета Д212.166.03 при ГОУ ВПО «Нижегородский государственный университет им. Н.И. Лобачевского» по адресу: 603000, г. Нижний Новгород, ул. Б. Покровская, д. 60, экономический факультет, ауд. 512

С диссертацией можно ознакомиться в библиотеке Нижегородского государственного университета им. Н.И. Лобачевского.



Автореферат разослан “ ___ ” _____

Ученый секретарь
диссертационного совета

Ю.А. Лебедев

1. ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

Мировой финансовый кризис 2008-2009 годов ознаменовал наступление периода экономической нестабильности. Коснулся он и российских промышленных предприятий. Усложнившаяся внешняя среда организаций и непредсказуемость макроэкономических факторов ставят перед российскими фирмами все более трудные управленческие задачи. Предприятия, пытаясь выжить в эту сложную эпоху, изменяют свою структуру и поведение. Все это делает экономическое пространство более сложным для анализа и прогнозирования.

В то же время различные институты, такие как банки и налоговые органы, по-прежнему заинтересованы в получении прогнозов экономического состояния предприятий, поскольку эта информация является ключевой при принятии решений и затрагивает ряд аспектов управления рисками, важность которых в условиях кризиса чрезвычайно обострилась. Важны такие прогнозы и для руководства самих предприятий.

Перечисленные факторы обуславливают актуальность разработки методов прогнозирования экономического состояния предприятий, работающих на сверхкоротких (не более 5-6 наблюдений) исторических рядах. Методика, позволяющая на основании малого объема исторических данных сформировать практически применимый прогноз состояния предприятия, стала бы важным в условиях кризиса инструментом экономиста-аналитика.

При использовании аппарата генетического программирования (ГП), помимо высокой практической значимости, задача разработки этих методов затрагивает ряд фундаментальных вопросов из области символьной регрессии, не разработанных на данный момент ведущими учеными этого направления. Среди них: способы избегания оверфиттинга для моделей с соотношением количества входных параметров к числу исторических точек порядка $5/20$, способы априорной оценки качества сделанного прогноза и их достоверность и др.

Кроме того, при анализе литературы и интернет-источников была выявлена недостаточная проработанность вопроса о применении ГП для прогнозирования значений отдельных экономических показателей промышленного предприятия на основе российской бухгалтерской отчетности. Вопрос о применимости ГП для решения подобных задач является открытым и требует ответа.

Обозначенные положения определяют практическую и теоретическую значимость выбранной темы диссертационного исследования в контексте сложившейся экономической ситуации и применимость его результатов.

Диссертационное исследование проведено в соответствии с требованиями паспорта специальности 08.00.13 “Математические и инструментальные методы экономики” ВАК РФ в следующих пунктах: 1.4. Разработка и исследование моделей и математических методов анализа микроэкономических процессов и систем: отраслей народного хозяйства, фирм и предприятий, домашних хозяйств, рынков, механизмов формирования спроса и потребления, способов количественной оценки предпринимательских рисков и обоснования инвестиционных решений, 2.8. Развитие методов и средств аккумуляции знаний о развитии экономической системы и использование искусственного интеллекта при выработке управленческих решений.

Степень научной разработанности проблемы.

Проблеме экономического прогнозирования посвящено множество работ и публикаций. К отечественным авторам, имеющим публикации по данной проблеме, относятся, в частности: И.В. Бестужев-Лада, В.В. Леонтьев, Д.С. Львов, С.А. Айвазян, В.С. Мхитарян, Ю.П. Лукашин, Ю.В. Трифонов, Ф.Ф. Юрлов, С.Л. Чернышев, А.Ф. Плеханова, Н.Ф. Поляков и другие.

Из зарубежных авторов, внесших значительный вклад в разработку данного вопроса, необходимо отметить: Дж. Бокса и Г. Дженкинса, К. Гренджера и М. Хатанака, М. Кендалла, А. Стьюарта.

Также непосредственное отношение к теме диссертационного исследования является задача определения экономического состояния предприятия. Этой проблеме посвящено множество работ ученых-экономистов. К отечественным авторам, которые рассматривают данную проблему, можно отнести: А.Д. Шеремета, М. И. Баканова, О.В. Ефимову, В.В. Ковалева, Е.С. Стоянову, Р. С. Сайфулина, Г.В. Савицкую и др.

В зарубежной литературе этой проблеме посвящены работы: Е. Альтмана, Ю. Бриггема, Л. Гапенски, Дж. К. Ван Хорна, Р.Н. Холта, Дж. Г. Сигела, Дж. К. Шима и др.

Одним из ключевых аспектов данного исследования является применение аппарата искусственного интеллекта к решению прикладных экономических задач. Важнейший вклад в разработку этой проблемы внесли следующие отечественные и зарубежные ученые: Дж.Р. Коза, Д.Б. Фогель, М.Л. Кричевский, О.А. Цуранов, С.И.Родзин и другие. Ведущим практиком использования генетического программирования и основателем научного направления применения этого метода в экономике и финансах считается тайваньский ученый Shu-Heng Chen, имеющий ряд важнейших публикаций по этой теме. В связи с вопросами валидации результатов



применения генетического программирования необходимо отметить вклад исследователей Николаса Навета (Франция) и Махмуда Кабудана (США).

Признавая вклад упомянутых ученых, следует отметить, что:

1) при анализе литературы и интернет-источников была выявлена недостаточная проработанность вопроса о применении ГП для прогнозирования значений отдельных экономических показателей промышленного предприятия на основе российской бухгалтерской отчетности;

2) при прогнозировании курса валют или ценных бумаг с применением генетического программирования, проводившегося многими исследователями, базой исторических наблюдений являлись сотни и тысячи отчетов, что по объему на несколько порядков превышает доступную историческую базу по состоянию промышленного предприятия, особенно в условиях кризиса и нестабильности, когда исторические данные теряют свою актуальность с точки зрения их прогностической ценности.

В свете обозначенных положений были выбраны цель и задачи диссертационного исследования.

Цель диссертационного исследования заключается в:

1) разработке **методики** прогнозирования экономического состояния российских предприятий на основании их бухгалтерской отчетности с применением генетического программирования, включая:

1.1) создание метода получения прогнозов экономического показателя предприятия;

1.2) разработку метода статистического анализа результатов прогнозирования;

1.3) разработку метода предварительной оценки эффективности метода на модельных данных.

2) разработке **программных комплексов**, реализующих предлагаемые методики;

3) проведение **верификации методик** на реальных данных, что включает в себя:

3.1) определение источников данных о предприятии и обоснование возможности их использования;

3.2) сбор необходимых данных;

3.3) применение метода к накопленным данным и анализ полученных результатов.

Поставленная цель обусловила необходимость решения **следующих задач**:

1) проанализировать текущую разработанность темы методов оценки и прогнозирования состояния предприятий;

2) разработать алгоритм получения единичных прогнозов показателей промышленного предприятия на основе его бухгалтерской отчетности с применением генетического программирования;

3) разработать алгоритм анализа качества и применимости полученного прогноза на основании статистической обработки множества результатов запуска алгоритма получения единичного прогноза;

4) реализовать разработанные алгоритмы в виде программных комплексов;

5) провести верификацию метода на реальных экономических данных российских предприятий.

6) провести прогнозирование состояния выбранных предприятий.

Объектом исследования выступают российские предприятия.

Предметом исследования являются способы определения и прогнозирования экономического состояния промышленных предприятий с использованием программного обеспечения, реализующего интеллектуальные математические методы анализа данных.

Теоретической и методологической основой диссертационного исследования являются современная теория экономического анализа, математического моделирования и интеллектуальных методов.

В работе использовались открытые статистические материалы по отдельным предприятиям Нижегородской области, электронные информационные источники, применялось специализированное программное обеспечение.

Научная новизна исследования состоит в следующем:

1) **Разработан метод получения единичных прогнозов** коэффициентов соотношения групп балансовых статей промышленных предприятий на сверхкоротком временном ряде, базирующийся на генетическом программировании и осуществляющий поиск символьной зависимости прогнозного значения целевого показателя от значений других показателей и балансовых статей этого предприятия за предыдущий год. Метод использует эволюционную технику машинного обучения, оптимизирующую популяции компьютерных программ с точки зрения их способности решать сформулированную проблему прогнозирования целевого коэффициента;

2) **Разработана методика априорной оценки качества полученного прогноза** на основании вида распределения исходов множества запусков алгоритма прогнозирования и его сравнения с результатом прогнозирования на модельных

хаотических данных, позволяющая идентифицировать проявления оверфиттинга и учитывать эту информацию при принятии решений на основании прогноза;

3) **Сформулирован критерий критического качества прогноза**, дающий численное значение для оценки применимости полученного прогноза;

4) **Разработана методика апостериорной оценки эффективности прогноза**, основывающаяся на прогнозных данных и реализовавшейся динамике прогнозируемого показателя;

5) **Скорректирована традиционная методика генетического программирования** с целью повышения эффективности ее применения к проблеме прогнозирования экономических коэффициентов предприятий. Коррекция позволила значительно ускорить сходимость метода и в результате получить большее количество экспериментальных данных, позволивших исследовать статистические характеристики результатов прогнозирования. Изменения коснулись адаптированной процедуры кроссовера и вычисления функции приспособляемости. В результате коррекции в ряде случаев удалось более, чем на 80% снизить использование вычислительных ресурсов.

Кроме того, в рамках проведения диссертационного исследования, автором были разработаны два **программных комплекса**:

1) Программный комплекс, реализующий разработанный метод прогнозирования с учетом коррекций традиционной постановки задачи генетического программирования, который может использоваться как самостоятельно, так и предоставлять свое математическое ядро для встраивания в пользовательские приложения через интерфейс прикладного программирования. Программный комплекс написан на языке C++ и работает в операционных системах Windows, Solaris. Полная платформенезависимость позволяет с легкостью портировать программный комплекс на другие операционные системы.

2) Программный комплекс обработки результатов прогнозирования, реализующий методику априорной оценки качества получаемого прогноза. Программный комплекс реализован на языке Visual Basic for Applications, интегрирован в документ Microsoft Excel для удобства обработки и отображения данных и может применяться на любых платформах, поддерживаемых Microsoft Office.

Практическая значимость диссертационного исследования состоит в возможности применения его результатов для прогнозирования состояния экономических объектов, а также в том, что разработанный в рамках ее

программный комплекс был успешно верифицирован и применялся для прогнозирования состояния российских предприятий в работе финансовых отделов ООО “Мера НН” и ООО “НПП Бинар”.

Апробация результатов исследования. Основные положения работы докладывались на международной научно-практической конференции “Государственное регулирование экономики, региональный аспект” (Н. Новгород, 2007), публиковались в виде тезисов в трудах XI международной научно-практической конференции “Системный анализ в проектировании и управлении” (Санкт-Петербург, 2007), докладывались на межвузовской научной конференции аспирантов и студентов “Инвестиционная составляющая рыночной экономики России” (Н.Новгород, 2008), I всероссийской научно-практической интернет-конференции “Современность и экономические науки” (Новосибирск, 2009), VI Международной научно-практической конференции “Основные направления повышения эффективности экономики, управления и качества подготовки специалистов” (Пенза, 2009).

Основные положения диссертации отражены в 9 научных публикациях.

Структура и объем работы. Диссертация объемом 151 машинописных листов состоит из введения, трех глав, заключения, списка литературы, состоящего из источников 110 источников, и трех приложений.

Во **введении** обосновывается актуальность темы исследования, выявляется степень научной разработанности различных аспектов проблемы в настоящее время, ставятся цели диссертационной работы и определяются задачи для их достижения, формулируется объект и предмет исследования, обосновывается научная новизна и практическая значимость работы.

Первая глава “Аналитическое исследование проблем оценки и прогнозирования экономического состояния промышленных предприятий” посвящена обзору и анализу вопросов сбора информации о предприятии, существующих подходов к определению состояния предприятия и проблемам его прогнозирования. Также в этой главе раскрывается используемая далее в работе семикомпонентная методика оценки состояния предприятий на основе безразмерных экономических коэффициентов.

Вторая глава “Разработка методов и инструментальных средств прогнозирования экономического состояния промышленного предприятия на основе генетического программирования” представляет описание авторских методик, использовавшихся в процессе получения прогноза показателей и статистического

анализа результатов прогнозирования. Раскрывается разработанный автором обобщенный подход к анализу методик прогнозирования на основании понятия характеристики прогноза. Также приводятся описания программных комплексов, реализующих разработанные методики.

Третья глава “Результаты верификации и применения разработанных методов с использованием бухгалтерской отчетности российских предприятий” содержит результаты верификации методов на данных пяти российских предприятий, включая три предприятия нефтехимического комплекса нижегородской области и два открытых акционерных общества, за период 2000-2009 гг.

Основные положения, выносимые на защиту.

1. Разработан оригинальный комбинированный метод, совмещающий получение прогнозов экономических показателей промышленных предприятий на сверхкоротком историческом ряде на основе модифицированного метода генетического программирования и статистическую обработку результатов запуска прогнозирующего алгоритма с целью априорной оценки качества прогноза и определения прогнозного значения показателя с помощью автоматизированного исследования статистических характеристик множества вариантов прогноза.

Под состоянием предприятия понимается значения его экономических параметров (коэффициентов), допускающих численное представление. Задача прогнозирования состояния предприятия сводится, таким образом, к составлению обоснованных суждений о поведении каждого из рассматриваемых параметров в будущем.

Для некоторого коэффициента допустим, что:

$$S(t) = D(t) + N(t),$$

где $S(t)$ – значение прогнозируемого показателя как функция времени, $D(t)$ – детерминированная компонента, $N(t)$ – недетерминированная, шумовая, стохастическая составляющая. Исходной посылкой задачи прогнозирования является то, что компонента $D(t)$ является определяющей, в то время как влияние помех и неучтенных факторов $N(t)$ не является существенным. Это утверждение нуждается в проверке, более того, соотношение детерминированной и шумовой составляющих зависит от метода, которым мы пользуемся для выявления зависимости.

Получив компоненту $D(t)$, возможно экстраполировать ее в будущее, получая тем самым прогнозные значения показателя с точностью, которую допускает присутствие шумовой компоненты $N(t)$. Вопрос о том, является ли детерминированная компонента константной во времени или может изменяться приводит к необходимости еще одного допущения, а именно признания преемственности (неизменности) детерминированной составляющей поведения системы в будущем относительно ее поведения в прошлом, что также нуждается в проверке.

В случае задачи прогнозирования состояния промышленного предприятия под $S(t)$ мы будем понимать значение одного из производных коэффициентов соотношения групп балансовых статей этого предприятия. Детерминированная компонента $D(t)$ будет определяться как символьная зависимость:

$$I_{jk} = I_{jk}(t) = I_{jk}(I_{jk}(t - \Delta t), I_{jk}(t - 2\Delta t), \dots, I_{jk}(t - \Delta t * b)), \quad j = 1..n,$$

где Δt - шаг временных отсчетов исторических данных, b -параметр ретроспекции (число предшествовавших временных отсчетов, значение параметров предприятия в которые будут учитываться при получении прогноза), k -уникальный код предприятия, j -номер балансовой статьи или коэффициента, участвующей в расчете прогноза, k -номер исследуемого коэффициента.

Таким образом, предположительно существующая зависимость связывает значения i -го показателя предприятия в момент времени t от значений его показателей в моменты времени $t - \Delta t, t - 2\Delta t, t - 3\Delta t$ и т.д.

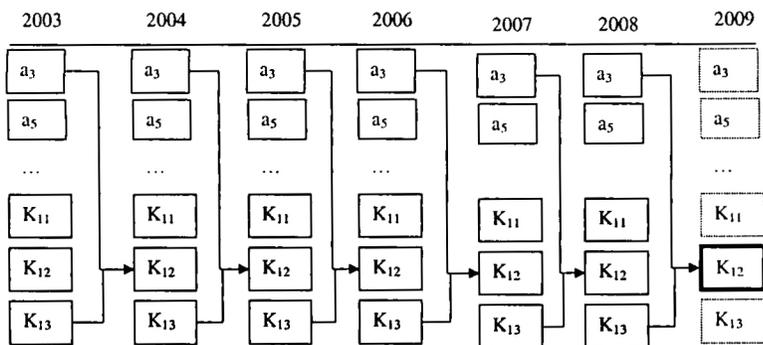


Рисунок 1. Схема вычисления прогноза. a_1 - a_n – балансовые статьи исследуемого предприятия, K_{ij} – производные коэффициенты, стрелками обозначена некоторая функция, которая на основе значений a_3, a_5, K_{21}, K_{13} для предприятия за предыдущий год вычисляет значение K_{12} в следующем году. Вычисление значения этой функции на последних известных данных (2008 год) дает прогнозное значение K_{12} на 2009 год.

Применение найденной зависимости к последним известным данным из обучающей выборки дает прогноз значения коэффициента на следующий отчетный период.

Предлагается искать упомянутую зависимость с помощью метода модифицированного генетического программирования, в рамках которой задача отыскания оптимальной структуры регрессионной модели формулируется следующим образом: задана выборка — множество $\{x_1, \dots, x_N | x \in \mathfrak{X}^M\}$ значений свободных переменных и множество $\{y_1, \dots, y_N | y \in \mathfrak{Y}\}$ соответствующих им значений зависимой переменной. Обозначим оба эти множества как множество исходных данных D .

Также задано множество $G = \{g | g : \mathfrak{X} \times \dots \times \mathfrak{X} \rightarrow \mathfrak{Y}\}$ гладких функций. Рассмотрим произвольную суперпозицию, состоящую из не более, чем r функций g . Эта суперпозиция задает регрессионную модель $f = f(x)$. Обозначим $\Phi = \{f_i\}$ — множество всех суперпозиций, порожденное элементами множества G .

Требуется выбрать такую модель f_i , которая доставляет минимум заданного функционала $P(\Phi)$, под которым может пониматься величина, зависящая, в том числе, от степени соответствия значений входных и выходных переменных. Таким образом, задача поиска данной символьной закономерности сводится к минимизации этого функционала на множестве возможных формул.

В данном исследовании функциональное множество генетического поиска составляют арифметические операции сложения, вычитания, умножения, деления, натурального логарифма, операции взятия корня второй и третьей степени. Также была введена функция синуса для моделирования потенциально существующей периодичности. Терминальное множество составляют значения балансовых статей и производных показателей, а также константы различных порядков: 10^{-3} , 10^{-2} , 10^{-1} , 1, 10, 100, 1000.

В дополнение к случайному поиску (операция мутации в классической постановке и “область хаоса” в модифицированном методе ГП, рассматриваемом в данной работе) ГП позволяет вести планомерную работу по планомерному приближению к минимуму от итерации к итерации (операция скрещивания), сохраняя при этом лучшие найденные на текущий момент решения (операция репродукции).

$$r = \sin \sqrt{\frac{A_{214}^{-2} - \frac{A_{190}^{-1}}{A_{159}^{-2}}}{\sin \left(\sqrt{\frac{A_{510}^{-2}}{A_{210}^0} A_{214}^{-2}} \right) - 0,1(A_{K_{21}}^0 - \sin(A_{K_{21}}^{-1})) \sqrt{\frac{A_{510}^{-2}}{A_{210}^0} A_{214}^{-2}}}}$$

Рисунок 2. Пример функции, полученной в ходе эволюционного поиска

Критерием меры близости решения к искомому является значение функционала невязки, определяющемуся как суммарное отклонение прогнозных значений в точках, где возможно его сопоставление с реальными значениями исследуемого показателя (на основе исторических данных), которое определяется как:

$$M^*(f) = \sum_{i=1,n} |f(t_i) - R(t_i)|,$$

где f – функция – кандидат, $t_{1..n}$ – моменты времени в прошлом, на которых возможно сравнение прогноза, основывающегося на исторических данных с реальными историческими значениями прогнозируемого показателя, $R(t_i)$ – реальные значения прогнозируемой величины в моменты времени t_i . Дополнительно, функции кандидаты меньшей сложности расцениваются как более приспособленные в соответствии с формулой:

$$M(f) = M^*(f) \frac{C}{C^*},$$

где $M(f)$ – результирующая функция невязки, $M^*(f)$ – функция невязки, вычисленная как суммарное отклонение на опорных точках, C – количество узлов в функции-решении, C^* – коэффициент штрафа за сложность особи-решения. Второй множитель в формуле вычисления функции невязки при равных степенях совпадения на исторических данных оценивает более простые решения как более приспособленные и производит таким образом отсев чрезмерно сложных кандидатов, уменьшая тем самым вероятность подгонки под кривую (curve fitting), что особенно важно с учетом малого количества исторических данных.

Отметим, что метод поиска прогноза коэффициента соотношения групп балансовых статей, описанная в предыдущем разделе определяет способ получения единичного прогноза и не учитывает стохастическую природу самого метода.

Также следует заметить, что монотонный спуск к минимуму не означает сам по себе выявления скрытых закономерностей в данных. Для выявления оверфиттинга в прогнозировании конкретного показателя предприятия мы будем использовать понятие **качества прогноза**.

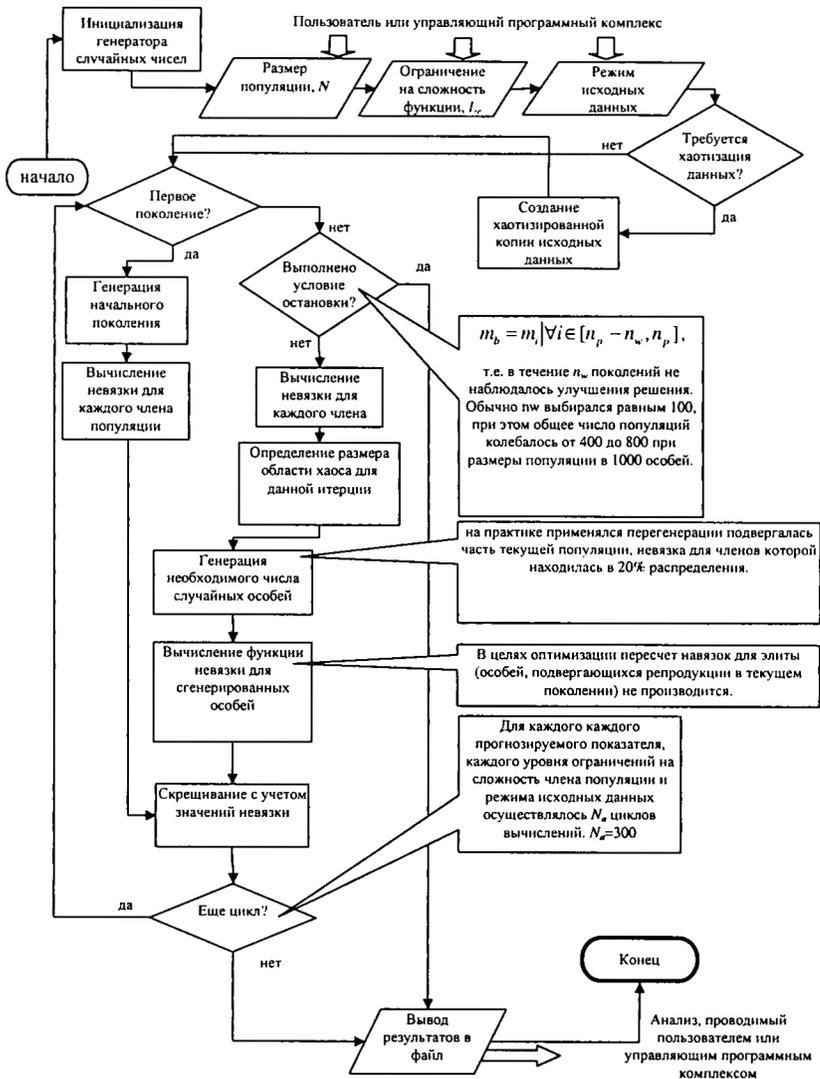


Рисунок 3. Схема реализации метода получения единичного прогноза экономического показателя предприятия на основе ГП.

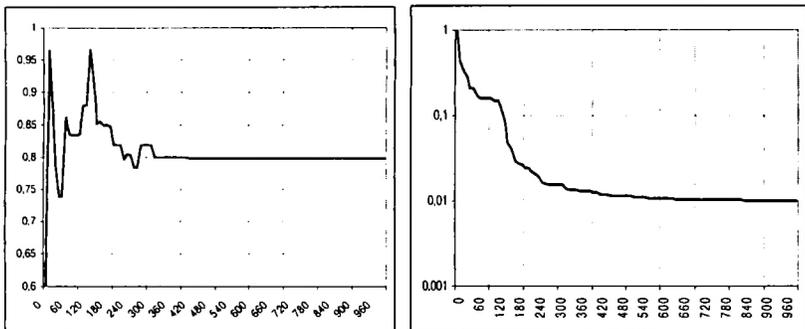


Рисунок 4. График сходимости прогноза промежуточного коэффициента покрытия ЗАО “Авиационные и технические масла” покрытия (слева) и уменьшения функции невязки (справа). По горизонтальной оси откладывается номер итерации, а по вертикальной: для **графика слева** – значение прогноза с использованием наилучшего решения в данной итерации, для **графика справа** – значение функции невязки для наилучшего решения в данном поколении.

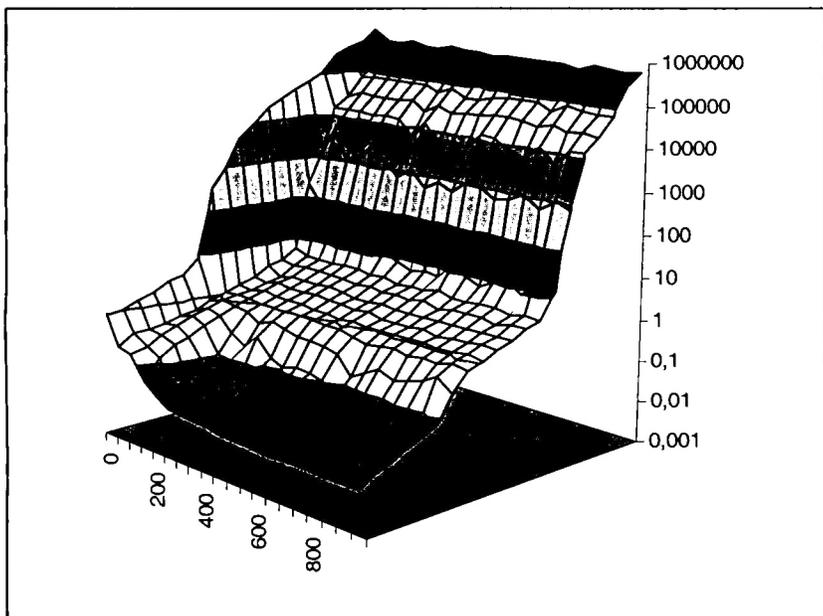


Рисунок 5. Изменение структуры популяций в процессе поиска решения. **Вертикальная ось** – значение невязки для члена популяции (логарифмическая шкала), **горизонтальная ось с делениями** – номер итерации, **горизонтальная ось без делений** – номер члена популяции.

Таблица 1. Модификации традиционной реализации метода ГП и их эффект

| Модификация традиционного метода ГП | Традиционная постановка | Достигнутый эффект |
|--|---|--|
| Множественные попытки кроссовера с выбором наилучшего потомка | Разовая попытка кроссовера | Ускорение сходимости на 10-12% (в реальном времени) и 20-25% процентов (измеряемое как количество поколений). |
| Область хаоса переменного размера. | Хаотическая генерация только первого поколения. | Ускорение до 80% в ходе начальных популяций. Ускорение порядка 10% в ходе последующих популяций. В измерениях использовалось реальное время. |
| Переход от максимизации функции приспособляемости к минимизации функции невязки. | Максимизация нормализованной (приведенной к интервалу 0..1) функции приспособляемости | Более выраженный эффект от ограничения функции невязки на рассогласование на реальных данных. |

Понятие качества прогноза является ключевым в в разработанной методике статистической обработки результатов прогнозирования с использованием вышеописанного метода, реализующей следующую функциональность:

1) На **первом шаге** производится многократный поиск прогнозирующей формулы на основании неизменных исходных данных.

2) Для **второго шага** создается копия исходных данных, к которой применяется операция хаотизации, т.е. случайного перемешивания исходных рядов и повторяется процедура первого шага.

Мы можем ожидать, что в случае, когда немодифицированные исходные данные содержат некоторый сигнал, множество запусков алгоритма поиска прогноза на исходных и хаотизированных данных выдадут статистически отличающиеся множества прогнозов. В случае же когда исходные данные представляют собой чистый шум, мы можем ожидать похожих результатов.

В качестве количественной меры сходства результатов, полученных в с помощью двух шагов первого функционального блока алгоритма, была предложена мера качества прогноза Q , определяющаяся на основании следующих выражений:

$$S = \int_{x_{\min}}^{x_{\max}} d(x) dx \frac{1}{\max(d(x)) | x \in [x_{\min}, x_{\max}] },$$

где $d(x)$ – значение плотности распределения вероятности прогноза показателя, построенного на неизменных начальных данных. x_{\min}, x_{\max} – значения максимального и минимального пределов значений в выборке.

$$S' = \int_{x'_{\min}}^{x'_{\max}} d'(x) dx \frac{1}{\max(d'(x)) | x \in [x'_{\min}, x'_{\max}] },$$

где $d'(x)$ – значение плотности распределения вероятности прогноза показателя, построенного на начальных данных, подвергнутых хаотизации, а x'_{\min}, x'_{\max} определяются аналогично предыдущему случаю:

$$x'_{\min} = \min(x'_i) | i = 1..N_a, \quad x_{\min} = \min(x_i) | i = 1..N_a$$

$$Q = \begin{cases} \left| \frac{S'}{S} - 1 \right| \frac{S'}{S} \geq 1 \\ \left| \frac{S'}{S} \right| < 1 \end{cases},$$

где Q – искомая характеристика прогноза, его **качество**.

Фактически, качество прогноза характеризует отношение площадей нормализованных функций распределения прогнозов для неизменных и хаотизированных данных. Другими словами, этот показатель характеризует степень относительной выраженности уплотнения распределения в области предполагаемого результирующего прогнозного значения и различия формы распределения для немодифицированных и модифицированных исходных данных.

Гипотезой H_0 в этом случае будет являться утверждение, что большие значения характеристики Q соответствуют большей прогностической ценности прогноза, альтернативная гипотеза H_1 утверждает, что корреляция между ними отсутствует. Статистическая проверка гипотез проводится в разделе анализа результатов.

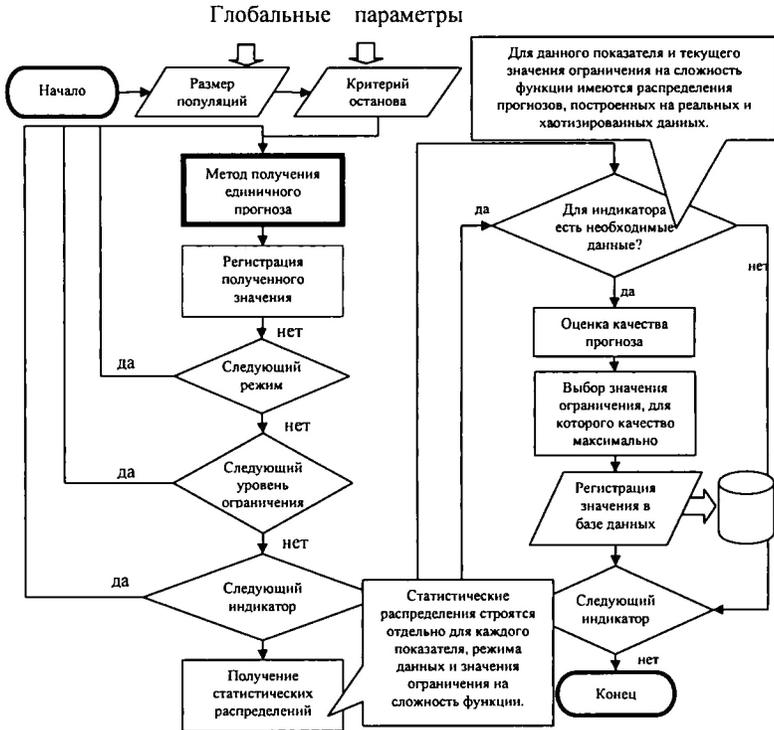


Рисунок 6. Схема второго блока алгоритма. Блок, выделенный жирными линиями, представляет собой точку вызова первого блока алгоритма, представленного на рисунке 3.



Рисунок 7. Пример функций распределения вероятности для неизменных данных (слева) и хаотизированных данных (справа). На этих рисунках горизонтальная ось соответствует множеству прогнозов, на которой определена функция плотности (откладывается по вертикали), определяемая как количество прогнозов попадающих в окрестность текущей точки.

При разработке и отладке метода возникла необходимость оценки эффективности прогноза на тестовой выборке. Абсолютная и относительная погрешность не являются удовлетворительными для этой задачи, поскольку не учитывают динамику реализовавшегося движения показателя, поэтому в рамках данной работы был предложен **показатель эффективности прогноза E** , определяющийся как:

$$E = \frac{E_0 - |\arctan(\Delta_r) - \arctan(\Delta_p)|}{E_0},$$

где E_0 – базисная величина угла расхождения графиков, в нашем случае принималась равным $\frac{\pi}{6}$, а значения Δ_r и Δ_p проиллюстрированы на приводимом ниже рисунке 8.

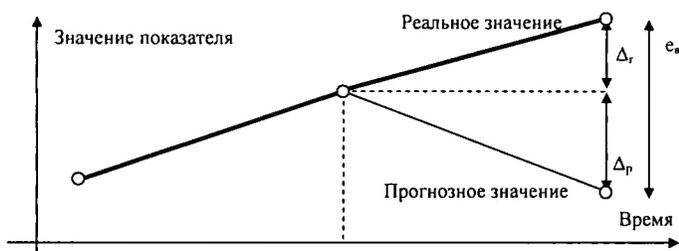


Рисунок 8. Схема определения реального и прогнозного изменений значения показателей.

Геометрическая аналогия, используемая при вычислении значения эффективности прогноза, опирается на вычисление угла между графиками действительного и прогнозного значений показателей.

Теоретически можно предположить существование двух возможных выводов, сделанных на этой стадии предварительного анализа:

- 1) эффективность прогноза не зависит от качества;
- 2) наблюдается прямая корреляция между значениями качества и эффективности.

Понятие характеристики прогноза позволяет описать метод прогнозирования как множество исходов его запуска на тестовой выборке и получить наглядное представление результатов на двумерной плоскости.

Отметим, что разработанная методика может применять не только для оценки применимости методов, основанных на генетическом программировании, но и как

общий подход, применимый для любых стохастических методик прогнозирования независимо от математического аппарата, ими используемого.

2. Предложена методика оценки состояния предприятия на основании семи коэффициентов соотношения групп балансовых статей.

Для получения данных о предприятиях выбрана балансовая отчетность, которая является основным, а в ряде случаев и единственным источником информации о предприятии.

Ниже приводится использованная в данной работе схема образования укрупненных статей баланса, частично заимствованная из существующих методик, применяемых в банковской практике для оценки кредитоспособности хозяйствующих субъектов.

Согласно применяемой методике балансового анализа, **активы и пассивы баланса** разделяются на 8 основных групп.

Таблица 2. Структура обобщенных групп актива баланса

| Обозначение | Название | Состав группы |
|----------------|----------------------------|--|
| A ₁ | Быстрореализуемые активы | Денежные средства (стр. 260) |
| A ₂ | Среднереализуемые активы | Отгруженные товары, краткосрочная дебиторская задолженность, краткосрочные финансовые вложения (стр. 215,240,250). |
| A ₃ | Медленнореализуемые активы | Запасы за исключением отгруженных товаров, прочие оборотные активы (строки 211, 212, 213, 214, 216, 220, 270). |
| A ₄ | Труднореализуемые активы | Нематериальные активы, основные средства, незавершенное строительство, доходные вложения в материальные ценности, долгосрочные финансовые вложения, прочие внеоборотные активы (стр.110, 120, 130, 135, 140, 150). |

Таблица 3. Структура обобщенных групп пассива баланса

| Обозначение | Название | Состав группы |
|----------------|------------------------------|---|
| P ₁ | Краткосрочные пассивы | Краткосрочные займы и кредиты (стр. 610) |
| P ₂ | Среднесрочные пассивы | Кредиторская задолженность и прочие краткосрочные обязательства (стр. 620, 660) |
| P ₃ | Долгосрочные пассивы | Долгосрочные займы и кредиты (стр. 510, 520) |
| P ₄ | Постоянные пассивы (капитал) | Уставной капитал, добавочный капитал, резервный капитал (стр. 410, 420, 430). |

Обозначим выделенные группы пассивов и активов как A₁, A₂, A₃, A₄, P₁, P₂, P₃, P₄ соответственно. Валюту баланса обозначим символом V. На основании проведенной группировки могут быть выделены следующие безразмерные коэффициенты:

$$K_{11} = \frac{A_1}{P_1 + P_2}; K_{12} = \frac{A_1 + A_2}{P_1 + P_2}; K_{13} = \frac{A_1 + A_2 + A_3}{P_1 + P_2 + P_3}; K_{14} = \frac{A_1 + A_2 + A_3}{A_4};$$

$$K_{21} = \frac{P_4}{P_1 + P_2 + P_3}; K_{22} = \frac{P_3 + P_4}{V}; K_{23} = \frac{P_4}{A_4}.$$

Используемые коэффициенты в существующей практике принято называть: K₁₁ – коэффициентом оперативной (срочной) ликвидности, K₁₂ – промежуточным коэффициентом покрытия, K₁₃ – общим коэффициентом покрытия, K₁₄ – коэффициентом соотношения ликвидных и неликвидных средств, K₂₁ – коэффициентом соотношения собственных и заемных средств, K₂₂ – коэффициентом финансовой устойчивости, K₂₃ – коэффициентом инвестирования.

Используемые коэффициенты отвечают следующим основным требованиям:

- 1) безразмерность получаемых коэффициентов;
- 2) простота использования;
- 3) относительная независимость получаемых значений коэффициентов относительно финансовой структуры предприятия и его размеров, что позволяет без дополнительной обработки производить сравнительный анализ предприятий.

3. Получены результаты тестирования метода на модельных и реальных данных. Сформулирован критерий критического качества прогноза. Проведена верификация методов на данных пяти российских предприятий.

Верификация и апробация разработанного метода и программного комплекса, построенного на его базе проходила в несколько этапов с применением разработанной методики оценки методов прогнозирования на основе понятий качества и эффективности прогноза.

На **первом этапе** проводился поиск прогноза существующей (случайным образом сгенерированной) зависимости некоторой величины от значений балансовых статей и производных коэффициентов каждого из пяти предприятий отдельно. **Второй этап** верификации аналогичен первому с тем отличием, что вместо зависимостей прогнозируемый ряд состоял из чисто случайных величин. При этом также проводилось множество попыток прогнозирования, причем для каждой попытки ряд генерировался заново. Критическое значение качества прогноза выбиралось как точка максимального расхождения между графиками.



Рисунок 9. Кривые распределения плотностей качеств прогнозов для модельных данных, основанных на информации по ЗАО “Авиационные и технические масла”. По горизонтальной оси отложено значение, для которого вычисляется плотность распределения, линия с маркерами – распределение для реальных данных, линия без маркеров – для хаотизированных данных.

Таблица 4. Критические значения качества прогноза.

| Название предприятия | Диапазон исходных данных | Значение $Q_{кр}$ |
|---------------------------------------|--------------------------|------------------------|
| ЗАО “Авиационные и технические масла” | 2000-2006 гг. | 0,6 |
| ОАО “Лукойл-Нижегороднефтеоргсинтез” | 2002-2006 гг. | 0,63 |
| ООО “Нефтемаслозавод Варя” | 2003-2006 гг. | 0,83 (значение спорно) |
| ОАО “Газпром” | 2002-2009 гг. | 0,63 |
| ОАО “РЖД” | 2004-2009 гг. | 0,66 |

На **третьем этапе** верификации разработанные методы применялись непосредственно к прогнозированию коэффициентов соотношений групп балансовых статей пяти российских предприятий и последующему сравнению полученных данных с известными реальными значениями. В исследовании использовалась балансовая отчетность следующих предприятий:

1. ЗАО “Авиационные и технические масла”
2. ОАО “Лукойл-Нижегороднефтеоргсинтез”
3. ООО “Нефтемаслозавод Варя”
4. ОАО “Газпром”
5. ОАО “Российские железные дороги”

Распределение характеристик прогнозов 35 коэффициентов рассматриваемых предприятий приведено на рисунке 10. Прямая линейной регрессии показывает статистическую связь между значениями качества и эффективности прогнозов.

На **четвертом этапе** проверки методов для двух акционерных обществ проводилось прогнозирование значений семи коэффициентов соотношения балансовых статей двух открытых акционерных обществ на начало 2010 года.

Для коэффициента срочной ликвидности ОАО “Газпром” необходимо отметить, что прогнозирующая модель плохо валидировалась на данных 2009 года, что должно учитываться в интерпретациях прогноза на 2010 год. Также должно учитываться сравнительно низкое (0,49) значение качества прогноза.

Промежуточный коэффициент покрытия этого предприятия в период 2007-2009 гг. оставался практически неизменным. Характеристики прогноза на 2010 год, тем не менее, не демонстрирует убедительного качества $Q=0,2$. Таким образом, прогноз на 2010 год для этого коэффициента не является убедительным и, по всей видимости, просто экстраполирует сложившуюся динамику изменения.

Аналогичная ситуация складывается в отношении общего коэффициента покрытия и коэффициента соотношения собственных и заемных средств **ОАО “Газпром”**. Для трех оставшихся коэффициентов (коэффициента соотношения ликвидных и неликвидных средств, коэффициента финансовой устойчивости и коэффициента инвестирования) значения качества прогноза превысили критические.

Прогноз коэффициента соотношения ликвидных и неликвидных средств предсказывает дальнейший рост этого показателя с некоторым замедлением, что подтверждается сложившейся тенденцией. Следует отметить, что этот прогноз обладает весьма высоким качеством (0,89).

Прогноз значения коэффициента финансовой устойчивости, наоборот, свидетельствует о возможном изменении направления изменения этого показателя в 2010 году но не прогнозирует резких, значительных изменений.

Прогноз коэффициента инвестирования предсказывает незначительное снижение этого показателя в 2010 по отношению к предыдущей отчетной дате.

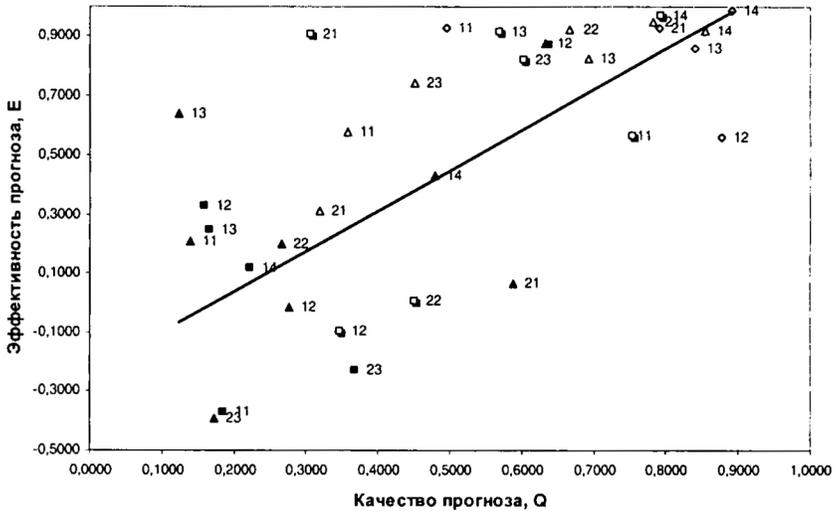


Рисунок 10. Распределение характеристик прогнозов коэффициентов соотношения групп балансовых статей пяти российских предприятий. Условные обозначения:

- ЗАО "Авиационные и технические масла", ▲ ОАО "Лукойл-Нижегороднефтеоргсинтез", ■ - ООО "Нефтемаслозавод Варя", △ - ОАО "Газпром", ◇ - ОАО "РЖД"



Рисунок 11. Историческая и прогнозная динамика изменения промежуточного коэффициента покрытия ОАО "Газпром"

Прогноз коэффициента срочной ликвидности ОАО "РЖД" на 2010 год предсказывает изменение динамики показателя: на смену росту 2008-2009 года должно прийти некоторое уменьшение. Прогноз этого коэффициента не обладает убедительным качеством (0,494) и поэтому не может считаться в достаточной мере достоверным.

Напротив, прогноз коэффициента промежуточного коэффициента покрытия обладает весьма высоким качеством (0,877) и предполагает изменение тенденции изменения данного показателя в 2010 году. По всей видимости, в данном случае возможна экстраполяция тенденции, выявленной на предыдущих отчетах. Учитывая динамику показателя в прошлом и достаточно высокое качество прогноза, данная оценка может считаться вполне достоверной.

В случае общего коэффициента покрытия ОАО «РЖД» налицо простейший случай экстраполяции практически линейной зависимости. Высокое качество прогноза и очевидная зависимость делает прогноз вполне применимым при принятии решений. Аналогичная ситуация складывается относительно прогноза коэффициента соотношения собственных и заемных средств, коэффициента финансовой устойчивости и коэффициента инвестирования этого предприятия. Коэффициент соотношения ликвидных и неликвидных средств ОАО «РЖД» демонстрирует высочайшее качество прогноза для данного предприятия (0,891) и предсказывает продолжение роста этого показателя с некоторым уменьшением его темпов.

Основные результаты работы. Ниже перечислены основные результаты диссертационного исследования.

1) Была обоснована актуальность разработки методов прогнозирования состояния промышленных предприятия по данным их бухгалтерской отчетности. Особое внимание было уделено специфике экономической информации о промышленном предприятии, характеризующейся, как правило, недостаточностью исторических данных.

2) Были проанализированы возможные источники получения данных о предприятии. Было показано, что бухгалтерская отчетность является наиболее подходящим способом получения информации для использования в формализованном и алгоритмизированном методе прогнозирования, который должен быть доступен для максимально широкого круга заинтересованных лиц, а также отвечать требованиям достоверности и сопоставимости исходных данных и результатов своей работы;

3) На основании существующей банковской практики был использован способ приведения бухгалтерского баланса к упрощенному виду на основе преобразования исходного вида баланса к набору из 7 коэффициентов соотношения групп статей, характеризующих состояние предприятия с точки зрения трех основных аспектов экономического статуса: платежеспособности, кредитоспособности и устойчивости.

4) Были разработана методика получения единичного прогноза целевого коэффициента на основе модифицированного генетического программирования. Использование ГП в отличие от традиционных методов позволило расширить область поиска возможных прогнозирующих зависимостей. Авторские модификации традиционного ГП позволили значительно ускорить сходимость метода.

5) Разработанная методика была реализована в виде алгоритма, включающего ряд авторских разработок в области оптимизации эволюционного поиска, позволивших значительно ускорить сходимость метода, что являлось ключевым фактором в процессе постановки численного эксперимента.

6) Разработана обобщенная методика оценки применимости стохастических методов программирования, основывающаяся на понятиях эффективности и качества прогноза. Методика позволяет наглядно оценить возможность применения выбранного метода прогнозирования на имеющихся исходных данных и основывается на последовательном сравнении результатов прогнозирования на реальных и специальным образом модифицированных исходных данных.

7) Методика была реализована в виде алгоритма, реализующего необходимые статистические и математические функции. В самой основе алгоритма была заложена возможность его использования как управляющей системы, контролирующей поток входных данных и результаты работы первого алгоритма получения единичного прогноза. Выбранная схема допускает возможность разделения вычислительных функций на два различных блока, которые могут выполняться различными программами и даже физически независимыми ЭВМ, соединенными сетью.

8) Разработанные алгоритмы легли в основу двух программных комплексов, модульная структура которых позволила эффективно распределить вычислительную нагрузку на кластере ЭВМ. Оптимальный выбор технологий реализации программных комплексов обусловил такие качества разработанной системы как открытость и масштабируемость. Модульная архитектура системы позволила эффективно распараллелить вычисления.

9) Проведена верификация метода на реальных данных по пяти российским предприятиям.

10) Проведено прогнозирование состояния двух открытых акционерных обществ на 2010 год.

На основании полученных результатов можно сформулировать следующие выводы:

1) Прогнозирование состояния предприятий является важной и практически значимой задачей.

2) Бухгалтерская отчетность является подходящим источником данных для получения прогнозов экономического состояния предприятия в будущем.

3) Генетическое программирование как метод символьной регрессии может применяться для прогнозирования экономического состояния промышленных предприятия по данным российской бухгалтерской отчетности и является оптимальным с точки зрения соотношения трудозатрат на различных этапах применения метода.

4) На основании верификации методов на известных исторических данных была показана его применимость для решения данного класса задач на примере пяти российских предприятий.

5) Полученные прогнозные данные на 2010 год могут применяться заинтересованными лицами с целью поддержки принятия решений, опирающихся на ожидаемое экономическое состояние исследуемых предприятий в будущем.

Из перечисленных выше результатов и выводов следует, что задачи диссертационного исследования были выполнены и его цели – достигнуты.

Список работ, опубликованных по теме диссертации:

1. Гречин, С.Б. Прогнозирование финансового состояния предприятий по данным бухгалтерской отчетности с применением генетического программирования // Вестник Белгородского Университета Потребительской Кооперации, №2 2009. – 0,5 п.л. (издание из списка ВАК)

2. Гречин, С.Б., Трифонов, Ю.В. Методика группировки балансовых коэффициентов как часть задачи экономического прогнозирования // Основные направления повышения эффективности экономики, управления и качества подготовки специалистов: сборник статей VI Международной научно-практической конференции. - Пенза, 2009. – 0,2 п.л., в т.ч. авторских 1,5 п.л.

3. Гречин, С.Б. Применение генетического программирования для оценки состояния предприятий // Межвузовский сборник научных трудов «Менеджмент качества и устойчивое развитие экономических систем».- Санкт-Петербург.- 2007. – 0,2 п.л., в т.д. авторских 1 п.л.

4. Гречин, С.Б., Трифонов, Ю.В. Генетический подход к прогнозированию балансовых коэффициентов // Вестник Нижегородского университета им. Н.И. Лобачевского, №2, 2008. – 0,2 п.л., в т.ч. авторских 1,5 п.л.

5. Гречин, С.Б., Трифонов Ю.В. Применение генетического программирования для прогнозирования состояния нефтехимических предприятий // Российский экономический интернет-журнал [Электронный ресурс]: Интернет-журнал АТиСО / Акад. труда и социал. отношений — Электрон. журн. — М.: АТиСО, 2002—№ гос. регистрации 0420600008. — Режим доступа: http://www.e-rej.ru/Articles/2008/Grechin_Trifonov.pdf, свободный — Загл. с экрана. - 0,2 п.л., в т.ч. авторских 1,5 п.л.

6. Гречин, С.Б., Трифонов Ю.В. Применение генетического программирования для прогнозирования состояния предприятий // Вестник Нижегородского университета им. Н.И. Лобачевского. №5. – Н. Новгород: Издательство ННГУ им Н.И. Лобачевского, 2007. - 0,4 п.л., в т.ч. авторских 0,3 п.л.

7. Гречин, С.Б. О способе оценки эффективности стохастических методов экономического прогнозирования при недостаточности исходных данных // Сборник материалов I Всероссийской научно-практической интернет-конференции “Современность и экономические науки”, Новосибирск, 2009. – 0,4 п.л.

8. Гречин, С.Б. Применение генетического программирования для оценки состояния предприятий // Труды XI Международной научно-практической конференции “Системный анализ в проектировании и управлении”, Ч.3. СПб.: Издательство Политехнического университета, 2007. - 0,3 п.л.

9. Гречин, С.Б. Применение генетического программирования для оценки состояния предприятий // материалы VI международной научно-практической конференции “Государственное регулирование экономики: региональный аспект” 17-19 апреля 2007 г. Нижний Новгород ННГУ. - 0,2 п.л.

.10 2