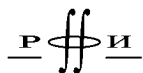


УДК 519.8
ББК 32.81
М 34



*Издание осуществлено при поддержке
Российского фонда фундаментальных
исследований по проекту 04-01-14027д*

Математические модели природы и общества / Н. Н. Калиткин, Н. В. Карпенко, А. П. Михайлов, В. Ф. Тишкин, М. В. Черненко. — М.: ФИЗМАТЛИТ, 2005. — 360 с. — ISBN 5-9221-0558-2.

Настоящая монография содержит основные результаты исследований в области построения математических моделей природы и общества, полученные Институтом математического моделирования РАН в содружестве с ГУИС ФАПСИ и социологическим факультетом МГУ им. М. В. Ломоносова. Эти исследования сконцентрированы на проблемах математического моделирования развития сложных ситуаций, возникающих в природе и обществе. С помощью мониторинговых и моделирующих систем поддержки принятия решений удаётся найти оптимальные способы и средства выхода из чрезвычайных ситуаций.

Результаты, полученные в работах, вошедших в данную монографию, активно используются в настоящее время в научных исследованиях и при принятии решений управленческими структурами общества. Монография будет полезна для научных сотрудников и управленческих кадров, а также для всех обучающихся и овладевающих знаниями в данной области математического моделирования.

© ФИЗМАТЛИТ, 2005

© Н. Н. Калиткин, Н. В. Карпенко,
А. П. Михайлов, В. Ф. Тишкин,
М. В. Черненко, 2005

ISBN 5-9221-0558-2

ОГЛАВЛЕНИЕ

Предисловие	5
Глава I. Создание систем мониторинга качества воздуха	8
§ 1. Механизмы распространения загрязняющих веществ после аварийных выбросов на промышленных объектах	8
§ 2. Создание программных комплексов, основанных на статистических представлениях	13
§ 3. Тестирование разработанного комплекса программ	47
§ 4. Архитектура пакета	49
§ 5. Исполнительные модули для интегрированной информационной системы	50
Список литературы	65
Глава II. Модели выбора и их применение к анализу социальных конфликтов	67
§ 1. Некоторые особенности моделирования социальных процессов	67
§ 2. Постановка проблемы, основные понятия и предположения.	74
§ 3. Математическая формулировка базовой модели и описываемые ею режимы эволюции	83
§ 4. Циклы политической самоидентификации дагестанского этноса — модельные прогнозы и сопоставление с реальностью	90
§ 5. Некоторые обобщения моделей выбора	97
Список литературы	108
Глава III. Компьютерные методы клиринговых расчётов	110
§ 1. Денежные расчёты: от курьера до электронной сети	110
§ 2. Взаимозачёт долгов методом фракционирования	119
§ 3. Переадресация	172
§ 4. Банковские платежи.	179
§ 5. Оценка долгов.	194
§ 6. Заключение	212
Список литературы	215

Глава IV. Стабильная пенсионная система	216
Введение	216
§ 1. Принципы пенсионной системы	218
§ 2. Балансные расчёты средних величин.	235
§ 3. Управление Пенсионным фондом	261
§ 4. Статистические данные.	276
Заключение	284
Список литературы	285
 Глава V. Математическое моделирование сложных трудноформируемых проблем социально-экономического развития региона	286
Введение	286
§ 1. Исследование динамики развития социального явления	287
§ 2. Прогнозирование значений показателей обучающимся генетическим алгоритмом прогнозирования LGAP.	295
§ 3. Исследование динамики развития социального явления при большом числе признаков	306
§ 4. Сравнительный межрегиональный анализ социального явления.	311
Заключение	321
Список литературы	322
 Глава VI. Моделирование взаимосвязей проблем при обработке текстов	323
Введение	323
§ 1. Проблемы и их представление в моделях поиска; выделение актуальных проблем для анализа взаимосвязей.	325
§ 2. Установление связей между значимыми рубриками	331
§ 3. Вероятностная модель определения области действия причинно-следственных связей.	338
§ 4. Модель поиска исходной информации для анализа взаимосвязей проблем	354
Список литературы	358

Предисловие

В последние десятилетия значительно расширилась сфера приложения математического моделирования. Одной из бурно развивающихся областей является прогнозирование средствами математического моделирования развития сложных ситуаций, возникающих в природе и обществе. С помощью мониторинговых и моделирующих систем поддержки принятия решений удаётся найти оптимальные способы и средства выхода из чрезвычайных ситуаций. В настоящее время рамки математического моделирования раздвинулись от уже традиционных проблем мониторинга окружающей среды до прогнозирования экологической, социальной, экономической, политической и государственной безопасности в масштабах произвольного региона и страны в целом. Математическое моделирование во многих случаях даёт возможность оценить степень опасности того или иного процесса в природе и обществе, провести оперативные расчёты рекомендуемых средств по предотвращению надвигающихся чрезвычайных обстоятельств или снижению величины ущерба от неизбежных катаклизмов и дать соответствующие прогнозы.

Несмотря на то, что математическое моделирование возникло почти одновременно с появлением вычислительной техники, до сих пор в отечественной и зарубежной литературе описание результатов решения подобного рода задач практически отсутствует. Общество испытывает насущную потребность в моделирующих системах и прогнозах, помогающих принимать решения по оценке складывающихся ситуаций, выработке стратегических направлений в развитии экономического и социального потенциала страны.

Авторы поставили перед собой задачу подробно изложить некоторые оригинальные результаты, полученные при тесном сотрудничестве Института математического моделирования РАН, ГУИС ФАПСи и социологического факультета МГУ им. М. В. Ломоносова, и провести их широкое обобщение.

В главе I нашли отражение многолетние исследования и разработки по развитию численных методов для описания распространения в атмосфере токсичных газов и капельно-воздушных инфекций. Изложены результаты вычислительных экспериментов, поставленных для задач, описывающих эти процессы, с начальными и краевыми условиями и когнитивными коэффициентами. Эксперименты проведены на специально разработанных вычислительных комплексах. Визуализация результатов даёт картину прогноза чрезвычайных ситуаций, что является необходимым условием выхода из такого положения. Результаты получены

авторским коллективом в составе к.ф.-м.н. Л. В. Ключковой, к.ф.-м.н. Д. В. Сузана, д.ф.-м.н. В. Ф. Тишкина.

В гл. II представлены оригинальные математические модели для мониторинга, анализа и прогноза типичных для российского социума этнополитических конфликтов, реализованных в соответствующих информационно-моделирующих системах поддержки принятия решений. Данный цикл работ выполнен коллективом специалистов в составе д.ф.-м.н. А. П. Михайлова, к.ф.-м.н. Л. В. Шведовского, к.ф.-м.н. А. И. Маслова, к.ф.-м.н. Л. Ф. Юхно, д.ф.-м.н. В. Ф. Ковалева.

Представленные в гл. III разработанные методы позволяют выполнять клиринговый взаимозачёт долгов предприятий, многократно уменьшающих суммарную задолженность в стране; проводить мониторинг экономики страны, отслеживая монополизацию и выявляя предприятия, скатывающиеся к банкротству. С их помощью можно определить, каким предприятиям целесообразно выдать кредиты для расшивания взаимных неплатежей, чтобы инициировать наиболее длинные цепочки выплат. Они позволяют определить разумную стартовую цену при аукционной продаже долгов.

В гл. IV сформулированы важнейшие требования к пенсионной системе: 1) она должна единообразно охватывать всех граждан независимо от профессиональной и ведомственной принадлежности; 2) размер пенсии должен зависеть от трудового вклада за всю жизнь; 3) система должна быть долгоживущей, не требующей законодательных изменений при сильной инфляции, обвалах экономики и банкротствах всевозможных компаний или фондов. Показано, что такая система должна быть 1) государственной, а не коммерческой; 2) распределительного типа, а не накопительного; 3) рассчитанной на граждан со средними заработками, а не сверхвысокими. Рассматриваемые сейчас Правительством проекты во многом не удовлетворяют этим требованиям, и в них заложены недостатки, которые через несколько лет станут источниками социальной напряжённости. Разработаны основы пенсионной системы требуемого типа. Даны простые формулы расчёта различных составляющих частей пенсии (социальной, трудовой, специальных надбавок) для разных категорий пенсионеров. Формулы содержат небольшое число коэффициентов, зависящих от демографической структуры населения и состояния экономики страны и определяемых на базе данных Пенсионного фонда, Госкомстата и налоговой службы. По статистическим данным найдена функция распределения населения России по доходам. Для основной массы населения она оказалась бистепенной, а учёт сверхвысоких доходов лишь незначительно изменил её. Это распределение исследовано, найден его децильный коэффициент. Расчёт баланса Пенсионного фонда с этим распределением оказался очень близким к расчёту по средним величинам. Это упрощает управление Пенсионным фондом. Собраны и проанализированы статистические данные, потребовавшиеся для данных расчётов. Даны рекомендации по распространению данных подходов на медицинское страхование

населения. Главы III и IV написаны чл.-корр. РАН Н. Н. Калиткиным, а изложенные в них результаты получены им в соавторстве с к.ф.-м.н. Л. В. Кузьминой. В этих работах принимали участие также д.ф.-м.н. А. П. Михайлов, к.ф.-м.н. С. В. Поляков (гл. III, § 2), Д. С. Гужев и к.ф.-м.н. И. О. Голосной (гл. III, § 1).

Представленные в гл. V математические модели разработаны на основе оригинальных модифицированных методов многомерного статистического анализа.

В гл. VI даётся описание процедуры формирования системы знаний по проблеме. Это даёт возможность на базе анализа отношений между документами, затронутых в них проблем и систем рубрик построить модель оригинальной информационной системы.

Цикл работ, изложенный в главах V и VI, выполнен коллективом специалистов в составе к.э.н. М. В. Черненкова, к.т.н. Н. В. Карпенко, к.ф.-м.н. А. Б. Карунина, С. В. Сычева.

В каждой главе своя нумерация формул, рисунков и таблиц, а также свой список литературы. Ни в одной главе нет ссылок на другие главы.

Быстродействие всех предложенных методов таково, что они пригодны для создания общегосударственной системы, работающей в режиме «боевого дежурства». Такая система может предотвращать большинство кризисных и аварийных ситуаций.

Глава I

СОЗДАНИЕ СИСТЕМ МОНИТОРИНГА КАЧЕСТВА ВОЗДУХА

§ 1. Механизмы распространения загрязняющих веществ после аварийных выбросов на промышленных объектах

Существует множество работ, в каждой из которых исследован какой-либо фактор или отражены несколько разных подходов для описания процессов распространения загрязняющих веществ над местностью, имеющей сложный ландшафт или в условиях промышленной застройки. Целью математического моделирования и проведения теоретических исследований механизмов распространения является оценка распределения воздушных потоков и примесей в них. Моделирование распространения примеси в атмосфере городов развивалось в нескольких направлениях. Например, использовались статистические модели распространения, основанные на функции распределения Гаусса [1]. Данное направление заключается в использовании моделей распространения, предназначенных для ровной подстилающей поверхности, модифицированных путём введения эмпирических коэффициентов, учитывающих возможное повышение концентрации в застойных зонах вблизи зданий и сооружений. Такой подход использован, например, в документе ОНД-86. Этот метод рекомендуется для установления нормативов ПДК в Российской Федерации. В этом документе вводится поправочный коэффициент, зависящий от взаимного расположения источника загрязнения атмосферы и близлежащих зданий. Этот подход практически эквивалентен введению понятия эффективной геометрии источника, поскольку застройка, расположенная на удалении от источника, не учитывается. Метод корректировки значений горизонтальной дисперсии при использовании гауссовых моделей, так же как и в ОНД-86, даёт возможность оценить вероятные повышения концентраций вблизи зданий.

При моделировании течений в уличных «каньонах» в [1] учитываются только здания, расположенные вблизи источника. Такие же предпосылки вводятся при решении уравнений термической гидродинамики и так называемых транспортно-диффузионных уравнений [2–8]. Моделирование течений в каньонах на основе решения уравнений

термической гидродинамики сопряжено с известными математическими трудностями, а также с принципиальными трудностями для всех моделей — заданием входных параметров:

- условий на границах (нижняя граница с потоком транспорта, здания со своим воздухообменом с уличным воздухом, верхняя граница, параметры которой зависят от многих метеорологических факторов);

- начальных значений, которые, как правило, должны зависеть от времени и, в частности, от тех же метеоусловий.

Кроме того, метеорологические модели в условиях больших городов могут иметь свои специфические особенности, например, они могут описывать образование острова тепла над промышленными и жилыми районами. Проблема при решении уравнений заключается ещё и в том, что необходимо задавать коэффициент переноса, зависящий от энергии турбулентных движений — функции многих величин. Наиболее простой способ определения этой функции следует из уравнения баланса турбулентной энергии. Адекватность приводимых моделей реальным условиям во многом определяется выбором значений эмпирических констант. Для описания формирования полей концентраций примеси часто используется полуэмпирическое уравнение переноса и диффузии. Так, в работе [4] сделана попытка на основе полуэмпирического уравнения переноса и диффузии примеси получить распределение примеси в отдельных уличных каньонах. В работе [5] распределение примеси над застройкой моделируется транспортно-диффузионным уравнением.

Физическое моделирование в аэродинамических трубах заключается в проведении физических экспериментов в аэродинамических трубах [9]. Эти эксперименты дают возможность оценить некоторые особенности распределения примеси в условиях застройки для таких метеорологических условий, которые можно с той или иной точностью воспроизвести в аэродинамической трубе. Следует отметить, что в трубах невозможно соблюсти подобие течения по достаточному набору критериев, например, задать число Рейнольдса одновременно с числом Росби. В то же время метод физического моделирования в аэродинамических трубах является практически единственным для определения некоторых необходимых для моделирования параметров и даёт возможность сравнения модели с измерениями, например, распределения потоков воздуха по улицам при различных направлениях ветра.

Моделирование потоков в аэродинамических трубах использовалось в работах Института гигиены и патологии с участием Института глобального климата и экологии РАН. Построение эмпирических моделей на основе комплексного подхода — анализа результатов натурных экспериментов, результатов численного моделирования и физического моделирования позволяет анализировать результаты натурных экспериментов. Результаты численного моделирования и физического моделирования сопряжены с построением параметрических моделей. Распределения примеси в уличных каньонах строятся в зависимости от метеоусловий: скорости и направления ветра, температурной стра-

тификации атмосферы, влажности и т. п. В параметрических моделях концентрацию загрязняющего вещества в уличном каньоне представляют как сумму концентраций:

- концентраций поступающих непосредственно от источников самого каньона (в основном автотранспорт);
- концентраций от сторонних источников (например, примесь, переносимая над данной местностью от промышленных предприятий);
- концентраций, обусловленных явлением рециркуляции внутри каньона.

Распределение примеси по этим моделям зависит от скорости ветра в каньоне и от коэффициента дисперсии, который, в свою очередь, зависит от координаты, скорости ветра, начальной дисперсии, связанной с масштабами начальных выбросов в приземный подслои, а также дисперсией значений турбулентной скорости. Эта последняя величина определяется характером вертикальных потоков над поверхностью земли.

Кроме того, в этих работах представлены подмодели химических трансформаций оксидов азота в атмосферном воздухе. Там же имеется сравнение с экспериментальными данными, полученными в Дании, Норвегии и Голландии.

В работах обращается внимание на факторы, влияющие на возникновение опасных концентраций в местах скопления пешеходов. Отмечается, что наибольшие колебания значений концентрации отмечаются на перекрёстках. При этом наибольшие значения концентрации наблюдаются при направлениях ветра, параллельных улицам. Одним из возможных путей развития такого направления является моделирование течений в уличных каньонах. Это осуществляется путём решения уравнений сохранения. Используются вспомогательные приёмы оценки характера течения вблизи зданий на основе сравнений параметров подобия. Например, при моделировании течения над местностью, имеющей сложный рельеф с перепадами высот. На основе оценки числа Фруда заключается, будет ли поток двигаться вверх по склону горы или обтекать её по горизонтальному уровню.

В настоящее время такие исследования проводятся в ряде научных центров мира (Калифорнийский университет, Международный институт Системного анализа в Австрии, Германский Национальный Исследовательский центр Информационных технологий). Однако проблемы, отвечающие в полной мере поставленной задаче, не нашли отражения в научных публикациях, что заставило приступить к разработке новых математических моделей.

В Институте математического моделирования в течение ряда лет разрабатывались вопросы, связанные с построением моделей распространения примесей в турбулентной атмосфере над местностью, имеющей сложный рельеф и в условиях городской застройки [10–17]. В основе этих моделей лежат фундаментальные разработки специальных вычислительных алгоритмов для решения определённого класса

газодинамических задач. Особенно важным вопросом в поставленной задаче является построение ветрового поля. Это сопряжено с необходимостью решения ряда задач, таких как изучение поведения ветра с высотой, влияния рельефа и температурной стратификации на скорость ветра и т. д. С другой стороны, поставленная задача требует чёткого понимания картины процессов, происходящих в пограничном слое атмосферы (знание термодинамики атмосферы, изучение явления турбулентности и составление эмпирических формул для нахождения коэффициентов турбулентности, владение методикой расчётов выбросов загрязнителей от различного вида источников и т. д.).

Решение поставленной задачи требует детального знания физических и химических свойств загрязнителей (интенсивности сухого поглощения и влажного осаждения, скорости химической трансформации и т. п.).

Важное место в создании программного комплекса занимает принятая физическая модель атмосферы, которая существенным образом влияет на построение поля ветра и на описание адвективных процессов. Атмосфера представляет собой подвижную среду, в которой происходят разнообразные по масштабам, направлению и скорости движения. Обычно эти движения имеют турбулентный характер и характеризуются непостоянством поля скоростей. При таких движениях образуются беспорядочные, изменяющиеся по направлению и силе потоки воздуха и вихри. В них можно выделить элементарные массы воздуха, которые отрываются от общего потока и движутся самостоятельно, а затем разрушаются. Всё это приводит к сильному перемешиванию и взаимодействию между различными частями среды. Перемешивание приводит к переносу любых физических субстанций, таких как количество движения, теплосодержание, концентрация примеси и т. п. При математическом описании процесса турбулентного перемешивания важную роль играет коэффициент турбулентности, так как он используется в выражениях для турбулентных потоков различных физических субстанций. Вертикальный турбулентный поток P субстанции q записывается в виде:

$$P = -k_z \rho \frac{\partial q}{\partial z}, \quad (1)$$

где ρ — плотность, а k_z — коэффициент турбулентной диффузии в вертикальном направлении.

Аналогичны выражения для турбулентного потока и по любым иным направлениям. Коэффициенты турбулентности не отражают свойств переносимой субстанции. Поэтому турбулентные потоки различных физических величин выражаются через одни и те же коэффициенты. Теории турбулентности достаточно сложны, и существует множество эмпирических методов для вычисления коэффициентов турбулентности. Коэффициент k_z зависит от множества параметров: от шероховатости подстилающей поверхности, распределения температуры

по высоте (так называемой температурной стратификации), скорости ветра и т. д.

Аналитическое решение для системы уравнений транспортно-диффузионной задачи для стационарного случая при некоторых упрощающих допущениях:

$$c = \frac{Q_e}{4\pi K r} e^{-\frac{u}{2K}(r-x)}. \quad (2)$$

На основе этой системы уравнений разработана оперативная трёхмерная модель прогнозирования концентраций примеси во времени и пространстве для нестационарного случая аварийного выброса облака токсичных веществ.

Построение транспортно-диффузионной модели и создание на её основе программного комплекса «TIMES» обеспечивает решение системы уравнений, описывающей процесс распространения загрязняющих веществ в ветровом поле и его графическое отображение в условиях городской застройки и над местностью, имеющей сложный рельеф, без подобного рода ограничивающих точность предположений. Эффективность разработанного пакета заключается в удовлетворительной работе модели при изменении в разумных пределах исходных данных и универсальности подхода. Другим достоинством пакета является его модульная архитектура, что делает возможным изменение некоторых его блоков (использование различных методик решения уравнений, изменение ряда факторов, учитываемых физической моделью процесса распространения примесей и т. п.), без ущерба надёжности работы комплекса.

В процессе разработки программного комплекса предпринимались попытки вставить различные схемы расчёта переноса. Однако эти схемы не удовлетворяли требованиям, предъявляемым при проведении эксперимента. Так, гауссова модель не предназначена для вычисления переноса загрязнений при следующих условиях: в переменном ветровом поле; при наличии нескольких источников переменной мощности; в области со сложным рельефом; для сложных метеоусловий; на близких к источникам расстояниях. Были также проведены вычислительные эксперименты по различным компенсационным схемам. Однако заложенные принципы компенсации не только не дают существенного улучшения, но и обладают дополнительными серьёзными недостатками: используемые в них явные разностные схемы не всегда удовлетворяют требованию консервативности.

Описанная методика вычисления с использованием сеточно-характеристического метода для моделирования адвекции хотя и удовлетворяет вышеперечисленным требованиям и является консервативной, но не позволяет исключить вычислительную диффузию, возникающую из-за погрешностей численных методов. Особенно ошибка заметна при расчётах для мгновенно действующего источника, а также для постоянного источника при различных направлениях ветра.

Для численного моделирования процессов переноса в транспортно-диффузионной модели малых примесей в ветровом поле атмосферного воздуха был разработан специальный метод. Отправной моделью для создания метода послужил классический метод частиц. Однако по принятым допущениям он оригинален и не имеет аналогов в литературе.

Новая схема расчёта полей концентраций при переносе в поле ветра позволяет быстро и эффективно рассчитывать концентрации в любой точке пространства и времени, но, главное, позволяет избежать принципиальных счётных ошибок при построении алгоритмов переноса.

§ 2. Создание программных комплексов, основанных на статистических представлениях

Программные комплексы создаются с целью получения средств мониторинга и получения данных для принятия решений по сохранению экологической безопасности окружающей среды в масштабах произвольного региона. Основной проблемой при этом является совместимость с геоинформационными системами (ГИС).

2.1. Применение статистических моделей, основанных на распределении Гаусса. С целью проведения сравнительного анализа построена модифицированная статистическая модель Гаусса, позволяющая рассчитать концентрацию загрязнителей в уличных каньонах и между зданиями.

Чтобы обеспечить расчёт концентрации загрязнителей в случае зданий, расположенных произвольно относительно направления ветра, каждое здание «накрывается» поверхностью, граница которой аппроксимирует профиль здания. Для каждого здания такая поверхность (рис. 1) вводится следующим уравнением:

$$z(x, y) = \frac{b}{\sigma_1 \sigma_2} \exp \left(-\frac{(x - a_1)^8}{2\sigma_1^8} - \frac{(y - a_2)^8}{2\sigma_2^8} \right), \quad (3)$$

где параметры a_1 , a_2 описывают положение здания относительно начала координат; σ_1 , σ_2 — параметры, характеризующие длину и ширину здания; b — параметр, характеризующий высоту здания, которая может быть вычислена как $h_b = \frac{b}{\sigma_1 \sigma_2}$.

Модель обтекания строится на основе статистического распределения Гаусса. Концентрации загрязняющих веществ над гладкой подстилающей поверхностью в нестационарном случае

$$\chi(x, y, z, t) = \frac{Q}{(2\pi)^{3/2} \sigma_x \sigma_y \sigma_z} \exp \left[-\frac{(x - \bar{u}t)^2}{2\sigma_x^2} \right] \exp \left[-\frac{y^2}{2\sigma_y^2} \right] \times \\ \times \left\{ \exp \left[-\frac{(z - H)^2}{2\sigma_z^2} \right] + \exp \left[-\frac{(z + H)^2}{2\sigma_z^2} \right] \right\}, \quad (4)$$

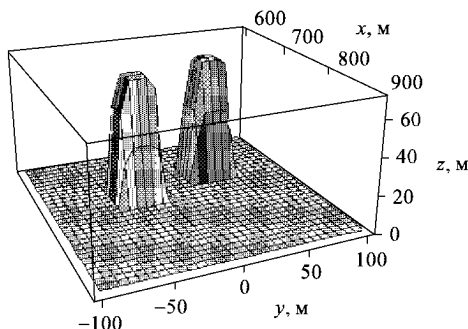


Рис. 1. Аппроксимация зданий поверхностями с помощью формулы (3)

а для стационарного случая концентрации загрязнений определяются выражением

$$n(x, y, z) = \frac{Q}{2\pi u \sigma_y \sigma_z} \exp\left(-\frac{(y - y_0)^2}{2\sigma_y^2}\right) \times \\ \times \left\{ \exp\left(-\frac{(z - H)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z + H)^2}{2\sigma_z^2}\right) \right\}, \quad (5)$$

где (x_0, y_0) — координаты основания источника; Q — мощность непрерывного точечного источника, [г/с]; u — скорость ветра на высоте H вдоль оси Оу [м/с]; $\sigma_y(x, x_0)$ — горизонтальная дисперсия, [м]; $\sigma_z(x, x_0)$ — вертикальная дисперсия, [м]; $|x - x_0|$ — расстояние от источника, [м]; $|y - y_0|$ — поперечное расстояние от оси шлейфа, [м]; z — высота над поверхностью земли, [м]; H — конечный подъём шлейфа над землёй (эффективная высота подъёма шлейфа), [м], в который для поля концентрации вводятся смещения в областях, примыкающих к зданиям. При этом выполняется закон сохранения массы вещества.

Это достигается следующим образом. В случае здания, аппроксимированного поверхностью, заданной уравнением $z = f(x, y)$, вводится дополнительная поверхность, задаваемая уравнением $z = 2f(x, y)$ такая, что объём между этими двумя поверхностями равен объёму здания. Тогда смещение поля концентрации в каждой точке вычисляется по формуле $S_z = -f(x, y)$. Это смещение вводится в формулу гауссовой модели для области, заключённой между этими двумя поверхностями. Концентрация по гауссовой модели с учётом зданий записывается в следующем виде:

$$n(x, y, z) = \frac{Q}{2\pi u \sigma_y \sigma_z} \exp\left(-\frac{(y - y_0)^2}{2\sigma_y^2}\right) \times \\ \times \left\{ \exp\left(-\frac{(z + S_z - H)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z + S_z + H)^2}{2\sigma_z^2}\right) \right\}. \quad (6)$$

На основе этой математической модели построен программный комплекс. Проведены вычислительные эксперименты, позволившие сделать сравнительный анализ количественных характеристик процессов распространения загрязнителей по различным моделям и, в частности, с помощью вновь создаваемой транспортно-диффузионной модели.

Расчёты, согласно построенной модели, проводились для трёх случаев. Значения параметров и физических величин, используемые при расчёте, приводятся ниже:

- 1) подстилающая поверхность за пределами урбанизированной местности — гладкая;
- 2) класс устойчивости атмосферы — D;
- 3) координаты основания источника $(x_0, y_0) = (15, 500)$;
- 4) эффективная высота источника — 12 м;
- 5) расчёт концентрации выполнен на высоте 2 м;
- 6) скорость ветра на эффективной высоте — 4,1 м/с;
- 7) мощность точечного источника — 100 г/с;
- 8) вертикальная дисперсия

$$\sigma_z = 0,06(x - x_0)(1 + 0,0015(x - x_0))^{-1/2};$$

- 9) горизонтальная дисперсия

$$\sigma_y = 0,08(x - x_0)(1 + 0,0001(x - x_0))^{-1/2}.$$

Урбанизированная местность в обоих случаях включала два здания:

1. Высота зданий (площадью $60 \times 40 \text{ м}^2$) составляет 70 м; координаты центров зданий (690, 540) и (720, 470). Результаты расчётов приведены на рис. 2 а, б, в.

2. Высота зданий (площадью $200 \times 60 \text{ м}^2$) составляет 70 м; координаты центров зданий (750, 560) и (750, 450). Результаты расчётов приведены на рис. 3 а, б.

Сечению $x = 675$ в 1-м случае соответствует область непосредственно перед зданием (по направлению ветра), координаты центра которого (720, 470). Сечению $x = 735$ в 1-м случае соответствует область непосредственно после здания (по направлению ветра), координаты центра которого (690, 540). Сечению $x = 750$ во 2-м случае соответствует область, содержащая оба здания. Продольные сечения во всех случаях берутся точно посередине между двумя зданиями.

Однако вычислительные эксперименты показали, что применение построенной модели для оценки концентрации загрязнений в районе застройки приводит к завышенным результатам.

Перераспределение концентрации лишь на основе закона сохранения массы даёт завышенные значения концентрации в уличных каньонах. При этом наблюдаются скачки значений концентрации в районе застройки. Это объясняется тем, что, в отличие от транспортно-диффузионной, данная модель не учитывает перераспределения потоков ветра в зависимости от конфигурации застройки.

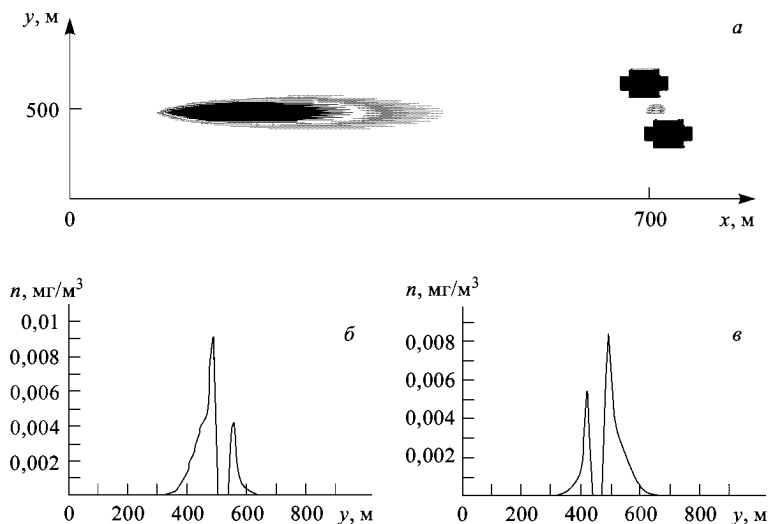


Рис. 2. Распространение загрязнений от непрерывно действующего точечного источника на высоте 2 метра. Обтекание зданий площадью $200 \times 60 \text{ м}^2$ (а). Тонами чёрного цвета передаётся величина концентрации. Концентрации примесей в поперечном сечении шлейфа при: $x = 675 \text{ м}$, $z = 2 \text{ м}$ (б); $x = 735 \text{ м}$, $z = 2 \text{ м}$ (в)

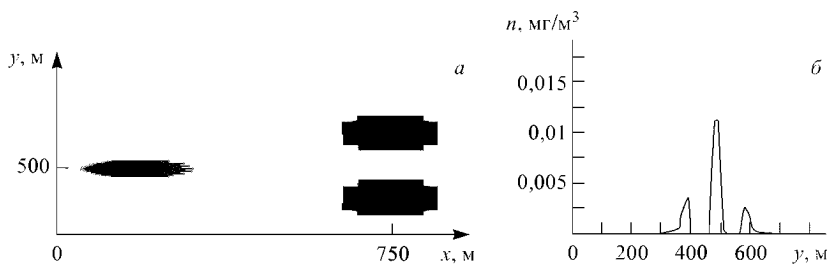


Рис. 3. Модель Гаусса для двух домов площадью $200 \times 60 \text{ м}^2$. Распространение загрязнений от непрерывно действующего точечного источника на высоте 2 метра (а). Концентрации примесей в поперечном сечении шлейфа при $x = 750 \text{ м}$, $z = 2 \text{ м}$ (б)

2.2. Транспортно-диффузионная модель.

2.2.1. Общие описание. Построенная математическая транспортно-диффузионная модель распространения примесей на основе системы уравнений параболического типа также выражает закон сохранения вещества. Перенос, изменение градиента концентраций, источники и стоки за счёт физико-химических процессов в такой модели описываются как

$$\frac{\partial c}{\partial t} + \text{div } \mathbf{c} \mathbf{w} - \text{div } (K \text{ grad } c) + rc = Q, \quad (7)$$

где c — концентрация субстанции, K — некоторый когнитивный коэффициент турбулентной диффузии, $\mathbf{w}$ — поле скоростей адвекции, Q — поле эмиссии, r — коэффициент, характеризующий интенсивность источников и стоков.

Модель процессов распространения загрязнителей для многослойной расчётной области сложной конфигурации может быть выражена системой уравнений, включающей модель ветрового поля в условиях городской застройки и над местностью, имеющей слабохолмистый ландшафт:

$$\begin{cases} \frac{\partial c}{\partial t} + \operatorname{div} c\mathbf{w} - \operatorname{div} (K \operatorname{grad} c) - \frac{\partial}{\partial z} K_z \frac{\partial c}{\partial z} + rc = Q, & z > h_t \quad (x, y) \\ c = 0, & z \leq h_l \quad (x, y), \end{cases} \quad (8)$$

где c — концентрация примеси, K — коэффициент горизонтальной турбулентной диффузии атмосферного воздуха, K_z — коэффициент вертикальной турбулентной диффузии, $\mathbf{w}$ — скорость ветра, Q — поле эмиссии, r — коэффициент, характеризующий интенсивность распада вещества, $h_l(x, y)$ — кусочно-непрерывная функция, которая описывает рельеф местности. Причём, $h_l(x, y) = h_b$, если точка $(x, y, 0)$ принадлежит основанию возвышения высотой h_b , и $h_l(x, y) = 0$, если в точке возвышения нет.

На основе этой математической модели разработан программный комплекс «TIMES» (Транспортная информационная модель для экологических систем), позволяющий строить решение системы уравнений и его графическое отображение.

В качестве исходных данных задаётся поле скоростей ветра над местностью сложного рельефа и коэффициенты вертикальной и горизонтальной турбулентности.

Расчётная область имеет форму прямоугольника с характерными размерами в несколько десятков километров. Характерная высота слоя перемешивания, в пределах которого происходит интенсивный перенос примесей в атмосфере, составляет 250–2000 м.

Учитывается влияние следующих факторов на процессы распространения: поле ветра, турбулентное движение масс воздуха, сухое поглощение и влажное осаждение и химическая трансформация веществ, составляющих выбросы (например, окисление SO_2 до SO_4^-). Модель рассчитывает распространение загрязнителей от нескольких источников, что её выгодно отличает от других ранее развитых моделей. По вертикали вся трёхмерная область разбивается на слои переменной высоты. Скорость ветра, высота верхней границы слоя перемешивания, интенсивность осаждения, класс атмосферной стабильности, температурный градиент и другие физические величины определяются по метеорологическим приборам.

Для приведённого примера, где в системе происходит окисление SO_2 , запись в расширенном виде с номерами слоёв вместо перемен-

ной z :

$$\begin{aligned} \frac{\partial h_i c_{1,i}}{\partial t} + \operatorname{div} h_i c_{1,i} \mathbf{w}_i - \operatorname{div} (K_i h_i \operatorname{grad} c_{1,i}) - \\ - D_{2,i} + (v_{d_2} + k_{\omega_2} h_i) c_{2,i} - k_t h_i c_{1,i} = \beta h_i \bar{Q}_i, \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{\partial h_i c_{1,i}}{\partial t} + \operatorname{div} h_i c_{1,i} \mathbf{w}_i - \operatorname{div} (K_i h_i \operatorname{grad} c_{1,i}) - \\ - D_{2,i} + (v_{d_2} + k_{\omega_2} h_i) c_{2,i} - k_t h_i c_{1,i} = \beta h_i \bar{Q}_i, \end{aligned} \quad (10)$$

в области $\Omega \times (0, T)$, $\Omega = L_x \times L_y$,

$$D_{1,i} = \begin{cases} K_{V_i} \frac{c_{1,i+1} - c_{1,i}}{0,5(h_i + h_{i+1})}, & i = 1, \\ K_{V_i} \frac{c_{1,i+1} - c_{1,i}}{0,5(h_i + h_{i+1})} - K_{V_{i-1}} \frac{c_{1,i} - c_{1,i-1}}{0,5(h_i + h_{i-1})}, & i = 1, \dots, nz - 1, \\ -K_{V_{i-1}} \frac{c_{1,i} - c_{1,i-1}}{0,5(h_i + h_{i-1})}, & i = nz, \end{cases} \quad (11)$$

$$D_{2,i} = \begin{cases} K_{V_i} \frac{c_{2,i+1} - c_{2,i}}{0,5(h_i + h_{i+1})}, & i = 1, \\ K_{V_i} \frac{c_{2,i+1} - c_{2,i}}{0,5(h_i + h_{i+1})} - K_{V_{i-1}} \frac{c_{2,i} - c_{2,i-1}}{0,5(h_i + h_{i-1})}, & i = 1, \dots, nz - 1, \\ -K_{V_{i-1}} \frac{c_{2,i} - c_{2,i-1}}{0,5(h_i + h_{i-1})}, & i = nz, \end{cases} \quad (12)$$

где $i = 1, \dots, nz$ — номер слоя, $c_{1,i}$, $c_{2,i}$ — концентрация, соответственно, SO_2 и SO_4^- , K_i — коэффициент горизонтальной турбулентной диффузии, K_{V_i} — коэффициент турбулентной диффузии, $\mathbf{w}_i$ — скорость ветра, $\bar{Q}_i$ — поле эмиссии, v_{d_1}, v_{d_2} — скорость сухого осаждения, соответственно, SO_2 и SO_4^- , k_{ω_1} , k_{ω_2} — скорость влажного осаждения, соответственно, SO_2 и SO_4^- , k_t — скорость химической трансформации SO_2 в SO_4^- , β — доля SO_4^- в примеси, h_i — толщина слоя, L_x , L_y — линейные размеры области Ω , $D_{1,i}$, $D_{2,i}$ — члены, описывающие вертикальную турбулентную диффузию.

Значения коэффициента горизонтальной турбулентной диффузии вычисляются по формуле

$$K = \sigma_\Theta^2 \cdot \max(0, 5, |\mathbf{w}|) \cdot h, \quad (13)$$

где σ_Θ — угол горизонтальной флуктуации направления ветра в радианах, определяемый из таблицы, $\mathbf{w}$ — скорость ветра в м/с, H — высота слоя перемешивания в м.

Класс атмосферной стабильности s задаётся на станциях.

Таблица 1. Угол горизонтальной флуктуации направления ветра в зависимости от класса атмосферной стабильности

Класс атмосферной стабильности, s	A	B	C	D	E	F	G
$\sigma_\Theta, ^\circ$	25	20	15	10	5	2,5	1,7

Значения коэффициента вертикальной турбулентности на некоторой высоте z вычисляются по следующим эмпирическим формулам:

$$K_z(z) = \begin{cases} \frac{z}{h} K_z(h), & z < h \\ K_z(h), & z \geq h, \end{cases} \quad (14)$$

где h — высота приземного слоя;

Таблица 2. Коэффициент вертикальной турбулентной диффузии в зависимости от класса атмосферной стабильности

Класс атмосферной стабильности	A	B	C	D	E	F	G
$K_z, \text{м}^2/\text{с}$	160	100	70	15	5	1,5	0,13

Таблица 3. Высота приземного слоя в зависимости от класса атмосферной стабильности

Класс атмосферной стабильности	A-C	D	E-G
$h, \text{м}$	100	150	250

Высота верхней границы слоя перемешивания H изменяется в пространстве и времени.

$$H_n(x, y, t) = a_n(x, y)t^2 + b_n(x, y)t + c_n(x, y), t_n \leq t \leq t_{n+1}, \quad (15)$$

где

$$a_1(x, y) = 0, t_1 = 0,$$

$$b_1(x, y) = \frac{F_2(x, y) - F_1(x, y)}{t_2},$$

$$c_1(x, y) = F_1(x, y),$$

$$a_n(x, y) = \frac{F_{n+1}(x, y) - F_n(x, y)}{(t_{n+1} - t_n)^2} - \frac{2a_{n-1}(x, y)t_n + b_{n-1}(x, y)}{t_{n+1} - t_n},$$

$$b_n(x, y) = 2a_{n-1}(x, y)t_n + b_{n-1}(x, y) - 2a_n(x, y)t_n,$$

$$c_n(x, y) = F_n(x, y) - a_n(x, y)t_n^2 - b_n(x, y)t_n.$$

где

$$F_n(x, y) = \bar{H}_n + \sum_{i=1}^{N_{zt}} \frac{k_{ni}}{1 + \varepsilon r_i(x, y)},$$

$$\bar{H}_n = \frac{1}{L_x L_y} \iint_{\Omega} H_a(x, y, t_n) dx dy,$$

$$H(x, y, t_n) = \bar{H} + \sum_{i=1}^{N_{st}} \frac{k_{ni}}{1 + \varepsilon r_i(x, y)},$$

k_{ni} — неизвестные коэффициенты, N_{st} — число станций, $r_i(x, y)$ — расстояние от точки (x, y) до i -й метеостанции, $H_a(x, y, t)$ — функция, значение которой в точке (x, y) в момент времени t равно значению H в этот момент на ближайшей станции, ε — малый постоянный параметр.

Чтобы вычислить k_{ni} , предыдущее условие записывается в точках расположения метеостанций. Для каждой станции с координатами (x_j, y_j)

$$H(x_j, y_j, t_n) = \bar{H} + \sum_{i=1}^{N_{st}} \frac{k_{ni}}{1 + \varepsilon r_i(x_j, y_j)}, \quad j = 1, \dots, N_{st}, \quad (16)$$

причём $H(x_j, y_j, t_n)$ — экспериментально измеренное на j -й станции значение высоты. Получается система N_{st} линейных уравнений относительно k_{ni} , которая решается методом Гаусса.

Разработана вторая схема расчёта верхнего слоя перемешивания. Высота слоя перемешивания аппроксимируется в виде суперпозиции колоколообразных функций, что позволяет получить ограниченный профиль значений.

Шероховатость z_0 , которая входит в коэффициент вертикальной турбулентности, можно оценивать по нижеприведённой таблице или из соотношения, по которому она составляет 1/7–1/10 средней высоты неровностей на подстилающей поверхности.

Таблица 4. Шероховатость для различных видов поверхности

Вид поверхности	z_0 , см	$\bar{z}_0$, см
Снег	0,005–0,1	0,05
Оголённая почва	0,1–1	0,5
Трава высотой 2–30 см	0,3–3	1
Сельскохозяйственные культуры	1–10	5
Лес	50–200	100

2.2.2. Источники. Данная модель предполагает наличие как площадных (расположенных на некотором участке земной поверхности), так и точечных (заводских труб) источников. Площадные источники характеризуются интенсивностью эмиссии E_s , измеряемой в $г/(м^2/с)$ и численно равной массе загрязнителя, выделяемой с единицы площади в единицу времени. Точечный источник характеризуется интенсивно-

стью эмиссии E_e , измеряемой в $г/(м^3/с)$ и численно равной массе загрязнителя, выделяемой в единицу объёма в единицу времени трубой высотой h_e и диаметром d_e . Высота и диаметр трубы, а также скорость вылета v_e и температура t_e газов, влияют на эффективную высоту подъёма шлейфа.

2.2.3. Учёт точечных источников. Учёт каждого точечного источника происходит следующим образом:

- 1) вычисляется эффективная скорость вылета: $v_{gr} = 0,5h_e^{0,6}$ м/с;
- 2) вычисляется тепловая эмиссия;
- 3) вычисляются индикаторы:
 $k_{hol} = 1,5v_e d_e + 0,00974 \cdot Q$ — формула Холланда,
 $k_{con} = 1,126 \cdot Q^{0,58}$ — формула Конкейва;
- 4) вычисляется эффективная высота шлейфа:

$$H_p = \begin{cases} h_e, & h_e \leq v_{gr}, \\ h_e + k_{hol} / \|\mathbf{w}\|, & h_e > v_{gr} \text{ и } Q \leq 2 \cdot 10^4, \\ h_e + k_{con} / \|\mathbf{w}\|^{0,7}, & h_e > v_{gr} \text{ и } Q > 2 \cdot 10^4, \end{cases} \quad (17)$$

где $\|\mathbf{w}\| = \sqrt{u^2 + v^2}$;

- 5) вычисляется промежуточный параметр b ;
- 6) вычисляется расстояние до точки максимальной концентрации:

$$x_m = C \left(\frac{H_p - h_i}{B} \right)^{1/b} \text{ м}, \quad (18)$$

где m, b, c — коэффициенты, зависящие от стабильности атмосферы, которые указаны в Таблице 5;

Таблица 5. Зависимость некоторых параметров от класса атмосферной стабильности

Стабильность	A	B	C	D	E	F
Параметры						
M	0,08	0,143	0,196	0,27	0,363	0,44
B	1,284	1,108	0,978	0,822	0,66	0,551
C	0,815	0,771	0,727	0,657	0,553	0,457

- 7) вычисляются эффективные координаты точечного источника:

$$\begin{aligned} x_0 &= x_e + x_m \frac{u}{\|\mathbf{w}\|}, \\ y_0 &= y_e + x_m \frac{v}{\|\mathbf{w}\|}, \end{aligned} \quad (19)$$

где x_0, y_0, y_e — координаты соответствующей заводской трубы.

Суммарный источник

$$\bar{Q}_i = \sum_{j=1}^{n_i^Q} Q_{e_i}^j \delta(x_{0_i}^j, y_{0_i}^j), \quad (20)$$

где суммирование проводится по всем источникам, эффективные высоты шлейфов которых находятся в i -м слое, j — индекс суммирования, n_i^Q — число таких источников, $Q_{e_i}^j$ — значение Q_e для источника с номером j для i -го слоя, $x_{0_i}^j, y_{0_i}^j$ — эффективные координаты этого источника.

Проведены исследования источников и стоков загрязняющих веществ за счёт физико-химических превращений. Проведён обзор литературных данных [18].

Управление качеством воздуха предполагает наличие стандартов качества воздуха. Вредными веществами являются пыль, сернистый ангидрид, окись углерода, двуокись азота, сажа (копоть), сероводород и т. д. В предлагаемой системе возможно осуществление мониторинга за концентрациями каждого вещества в отдельности. С помощью блока физико-химических превращений исходная концентрация может быть уточнена, если возможна аппроксимация сложной системы физико-химических превращений в атмосфере. Концентрация контролируемой вредной примеси аппроксимируется реакцией первого порядка с обобщённой константой скорости для частиц вещества в газовой фазе, многофазном состоянии или в капельно-жидкостной фазе. Тогда можно учесть процессы химических превращений, сухого поглощения, седиментации и влажного осаждения.

Для программного комплекса «TIMES-URBAN» стандартными встроенными блоками для учёта химических реакций, сухого поглощения и влажного осаждения разработаны системы для описания превращений соединений серы и азота.

Атмосфера в силу наличия в ней свободного кислорода представляет собой систему, обладающую окислительными свойствами. Практически все реакции соединений серы и азота идут в сторону образования сульфатов и нитратов, т. е. высших форм окисления. Наиболее интенсивно реакции протекают со свободными радикалами. Важную роль в образовании радикалов имеют фотохимические процессы. Поэтому скорости реакций существенным образом зависят от времени суток, сезона, широты местности и наличия облачного покрова. При наличии молекул воды и в жидкой фазе окислы более высокого порядка быстро окисляются до серной и азотной кислоты соответственно.

Существенными параметрами являются константы скорости реакций и время жизни окислов. Многие реакции являются радикальными и прекращаются с наступлением темноты.

Встроенный блок для расчёта двуокиси серы за счёт химических превращений, сухого поглощения и влажного осаждения учитывает ряд характерных процессов.

Газо-фазное окисление двуокиси серы идёт в основном по реакциям, в которых особую важность во всей цепочке превращений играют озон и интенсивность солнечного света. С наступлением темноты практически мгновенно прекращаются радикальные реакции. В ночное время газо-фазное окисление обусловлено только молекулярными реакциями, которые существенно медленнее, чем радикальные. Летом в северных широтах (45°) суммарный коэффициент среднесуточной скорости газо-фазных реакций окисления до серной кислоты принят $0,015 \text{ ч}^{-1}$, а соответствующее время жизни — 70 ч. Данные оценки характерны для условий загрязнителей над промышленными регионами. В дневное время эта скорость на несколько порядков выше, чем ночью. Поэтому ночью скорость окисления можно положить равной нулю. Образующиеся по реакциям окисления окислы серы практически мгновенно превращаются в серную кислоту.

Многофазные реакции окисления на поверхности частиц летучей золы и сажи являются другим механизмом окисления двуокиси серы. Общий эффект многофазных реакций в окислении двуокиси серы в существенной степени зависит от концентрации частиц, площади поверхности на единицу объёма, физико-химических свойств поверхности и т.д. При 25°C и относительной влажности 60–70 % частицы летучей золы покрыты слоем воды толщиной меньше молекулярного, а при 90 % — несколькими молекулярными слоями. В условиях высокой влажности скорость окисления может возрасти на 1–2 порядка. На первом этапе существования дымового факела скорость окисления высокая, но по мере очищения атмосферы от частиц, снижения их объёмной плотности и насыщения поверхности вероятность протекания газо-фазных реакций будет уменьшаться. В изменяющихся условиях факела обобщённый коэффициент скорости газо-фазных реакций будет уменьшаться в пределах от 0,3 до $0,003 \text{ ч}^{-1}$. Для условий реальной атмосферы над крупным промышленным регионом среднесуточная скорость летом $0,01 \text{ ч}^{-1}$, а соответствующее время жизни двуокиси серы относительно гетерогенных реакций — 100 ч.

Третьим механизмом окисления двуокиси серы служит окисление *в жидко-капельной фазе облаков и туманов*. Реакции могут носить каталитический характер, при этом скорость возрастает на 1–2 порядка. В качестве катализаторов могут выступать ионы марганца и железа, концентрация которых над промышленными районами $3 \cdot 10^{-7} \div 4 \cdot 10^{-6}$ моль/л. Обобщённая константа скорости превышает 1 ч^{-1} . Однако скорость выведения двуокиси серы из атмосферы определяется не только химическими реакциями, но и вероятностью попадания молекулы в жидкофазную среду. Эта вероятность, в свою очередь, зависит от множества факторов: климатических условий региона, времени года, высоты выброса двуокиси серы над поверхностью земли, значения pH облачных и дождевых капель. У европейского региона, для которого характерно статистическое распределение «сухих»

и «мокрых» периодов, коэффициент равен $0,007 \text{ ч}^{-1}$, соответствующее время жизни — 140 ч.

Кроме того, константа обобщённой скорости окисления в жидкофазных реакциях, как показали экспериментальные исследования, зависит от времени переноса загрязнённых воздушных масс.

Из таких компонентов складывается обобщённая скорость окисления двуокиси серы, установленная экспериментальным путём. Существенен суточный ход обобщённой скорости окисления двуокиси серы в различные сезоны года. Вне зависимости от сезона ночью скорость падает до нуля. Дневная скорость окисления на порядок выше летом, чем зимой.

Обзор литературных данных показывает, что оценки скорости окисления колеблются от $0,005$ до $0,15 \text{ ч}^{-1}$. Принято считать среднесуточную суммарную скорость окисления равной $0,027 \text{ ч}^{-1}$, а соответствующее время жизни относительно окислительных процессов — 37 часов. Данные оценки являются среднесуточными и характерны для условий загрязнения над промышленными регионами летом.

В программном комплексе суточные колебания обобщённой скорости окисления аппроксимированы синусоидальной кривой с амплитудой, равной соответствующей сезонной максимальной обобщённой скорости окисления. При этом в исходных данных необходимо указывать фазу $6,28/24t_{\text{сут}} = 0,262t_{\text{сут}}$.

На величину концентрации двуокиси серы, кроме процесса окисления, влияет процесс сухого поглощения. В отличие от процесса седиментации, протекающего при выбросе крупных тяжёлых частиц (диаметром порядка 10 мкм), процесс сухого поглощения протекает при контакте молекул газа с подстилающей поверхностью: молекулы газа оказываются захваченными поверхностным слоем и таким образом выводятся из атмосферы. Этот процесс характерен для газо-фазных и гетерофазных реакций (размер частиц порядка $0,1 < d < 1,0 \text{ мкм}$). Поскольку при захвате молекул на поверхности раздела фаз происходит адсорбция и хемосорбция, скорость во многом будет определяться как физико-химическими свойствами поглощающей поверхности, так и свойствами поглощаемых газов и частиц. Химически активная природа окислов серы делает сухое поглощение важным механизмом выведения соединений серы. При моделировании процессов делается ряд существенных предположений. Модели сложны, наиболее надёжный способ ориентироваться на среднестатистические экспериментальные данные. Поскольку в свободной атмосфере перемешивание газов и частиц происходит с совершенно одинаковой скоростью, которая определяется, в основном, скоростью ветра, макромасштабной неровностью подстилающей поверхности, температурным градиентом, неравномерностью нагрева различных участков подстилающей поверхности. Измерения показывают, что влияние поверхности в условиях лесного массива приводит к изменению слоя турбулентного перемешивания до высоты 250 м при скорости ветра примерно 10 м/с . Скорость турбулентного

перемешивания может изменяться в широких пределах. В условиях ночной инверсии турбулентный перенос полностью прекращается, скорость падает до нуля.

Механизмы переноса через приповерхностный слой для молекул и частиц резко отличаются друг от друга. Существует три типа сопротивления захвату. Скорость поглощения единицей поверхности в слое поглощения пропорциональна концентрации и обратно пропорциональна сопротивлению. Константа скорости изменения концентраций газов или аэрозольных частиц в воздушной массе в процессе переноса над поверхностью за счёт сухого поглощения, τ -поглощения, зависит от слоя и скорости поглощения. Слой турбулентного перемешивания для Центральной Европы для двуокиси серы — 1200 м. Эмпирическим путём установлено, что можно описать изменения концентраций газов или аэрозольных частиц в воздушной массе в процессе переноса над поверхностью за счёт сухого поглощения так же, как сток, в виде λc .

Вымывание двуокиси серы из атмосферы осадками также состоит из множества сложнейших физико-химических реакций. Сложность и разнообразие механизмов вымывания существенно затрудняют создание достаточно точных моделей. Поэтому на основе экспериментальных данных о составе и количестве вещества, поступающего на подстилающую поверхность с осадками, вывели приближённую зависимость обобщённой скорости вымывания двуокиси серы. Для условий Европы можно принять, что для двуокиси серы $\tau_{wet} = 0,01 \text{ ч}^{-1}$. Соответствующее время жизни — 100 ч. Эмпирическим путём установлено, что можно описать изменения концентраций газов или аэрозольных частиц в воздушной массе в процессе вымывания так же, как сток, в виде λc .

Таким образом, расчётная формула концентрации двуокиси серы при учёте химического превращения, сухого поглощения, влажного осаждения, времени суток и сезона разработана, учитывая вышеизложенные соображения,

$$\frac{\partial c}{\partial t} = \sum_{i=1}^3 \frac{\partial}{\partial x_i} K_i \frac{\partial c}{\partial x_i} + \mathbf{w} \operatorname{div} c - [(0,027 + 0,015) \sin(0,216 t_{\text{дн}}) + 0,01] c + Q \quad (21)$$

где t — момент времени наблюдений, $t_{\text{дн}}$ — время суток момента наблюдений, сезон — июль.

В программном комплексе имеется блок для расчёта изменения концентраций окислов азота. Характерной особенностью этих реакций является то, что некоторые из них носят циклический характер и по существу не приводят к выведению окислов азота из атмосферы. В результате этих циклических превращений устанавливается равновесное состояние между озоном, окисью и двуокисью азота. Выведение окислов азота из атмосферы с образованием азотной кислоты осуществляется

тремя газо-фазными реакциями, некоторые из которых могут протекать лишь в светлое время суток.

При сжигании топлива образуется окись азота NO. Затем NO доокисляется до двуокиси азота NO₂. При выходе из дымовой трубы 90 % окислов азота в выбрасываемых газах составляют газы окиси азота, 10 % — двуокись азота. В зоне максимума концентрации выброса окись азота составляет уже приблизительно 35 % всех выброшенных окислов азота, а двуокись — 65 %. Поэтому шкала пересчёта [NO₂] = 1,53 [NO].

Измерения показывают, что ночью отношение объёмных концентраций окиси и двуокиси азота меньше и составляет 0,05. При солнечной погоде оно составляет 0,7.

Особую важность для всех реакций окислов азота играет вся цепочка превращений озона и интенсивность солнечного света. С наступлением темноты практически мгновенно прекращаются радикальные реакции. В ночное время газо-фазное окисление обусловлено только молекулярными реакциями, которые существенно медленнее, чем радикальные. Летом в северных широтах (45°) суммарный коэффициент среднесуточной скорости газо-фазных реакций окисления до азотной кислоты принят 0,015 ч⁻¹, а соответствующее время жизни 70 ч. Данные оценки характерны для условий загрязнителей над промышленными регионами. В дневное время эта скорость на несколько порядков выше, чем ночью, и поэтому ночью скорость окисления можно положить равной нулю.

С учётом сухого и влажного осаждения и времени суток для июля разработана расчётная формула изменения концентрации окислов азота.

$$\frac{\partial c}{\partial t} = \sum_{i=3} \frac{\partial}{\partial x_i} K_i \frac{\partial c(x_i, t)}{\partial x_i} + \mathbf{w} \operatorname{div} c + \\ + Q - [(0,12 + 0,025 + 0,01) \sin(0,216t_{\text{дн}}) + 0,01]c \quad (22)$$

где t — момент времени наблюдений, $t_{\text{дн}}$ — время суток момента наблюдений.

В настоящее время модель предполагает химическую трансформацию любых веществ, но отдельным блоком в качестве примера разработана схема окисления SO₂ → SO₄⁻, а также сухое поглощение и влажное осаждение. Однако модульная архитектура пакета позволяет описать любой сложный процесс путём введения дополнительных членов и уравнений.

2.3. Модели ветра. Разработка математических моделей ветровых полей как части общей транспортно-диффузионной математической модели необходима для прогноза распространения загрязнений в масштабах произвольного региона над слабо холмистой местностью и в условиях промышленной застройки и при других крутых препят-

ствиях (например, состоящих из горных массивов с длинными грядками хребтов).

Исходными предпосылками для моделирования является физическая модель атмосферы, которая рассматривает атмосферу как подвижную среду, в которой происходят разнообразные по масштабам, направлению и скорости движения. Обычно эти движения имеют турбулентный характер и характеризуются непостоянством поля скоростей. Всё это приводит к сильному перемешиванию и взаимодействию между различными частями среды. Перемешивание приводит к переносу любых физических субстанций, таких как количество движения, теплосодержание, примеси и т. п. Математическое описание процесса турбулентного перемешивания и переноса примеси в данной работе строится на подборе соответствующих коэффициентов турбулентности и потоков, огибающих и обтекающих препятствия. Выражения для турбулентного потока по любым направлениям аналогичны, но различаются коэффициентами. Коэффициенты турбулентности не отражают свойств переносимой субстанции, поэтому турбулентные потоки различных физических величин могут выражаться через одни и те же коэффициенты.

В программном комплексе изменение ветра с высотой может быть описано двумя моделями. Во-первых, как степенная зависимость вектора скорости ветра от высоты, и, во-вторых, системой уравнений Экмана, описывающей вертикальный профиль ветра.

Геострофический ветер, или градиентный ветер, определяется как ветер при прямолинейных изобарах, направленный перпендикулярно градиенту давления. Пакет предоставляет возможность учёта как геострофического ветра, так и эффекта изменения вертикального профиля ветра с высотой, описываемого эмпирической степенной функцией.

Важное место в создании моделей ветра занимает проблема математического описания безотрывного течения воздушных потоков над местностью, имеющей сложный рельеф. Описание течения над местностью, имеющей слабо холмистый рельеф, существенно отличается от описания обтекания крутых препятствий. Были построены несколько моделей. Все они базируются на расчёте угла поворота вектора скорости ветра и изменении его модуля при обтекании ландшафта (пологих холмов или домов в условиях городской застройки).

Для модели, описывающей поле ветра между крутыми препятствиями и, в частности, между домами городской застройки, наиболее сложным является задание больших градиентов высоты рельефа. Здесь возникает необходимость задания граничных условий: внутри домов поле равно нулю, на фронтальной стене скорость равна нулю, на стене, перпендикулярной фронту, вектор скорости направлен параллельно стене. С помощью граничных условий производится разрыв потоков.

При этом учитывается угол поворота вектора скорости ветра и изменение модуля вектора скорости при бездивергентном обтекании воздушными массами домов или крутых препятствий, производится

корректировка начального приближения для ветрового поля в каждой точке пространства между домами путём минимизации дивергенции и сглаживания.

Поле ветра рассчитывается по следующему алгоритму:

- 1) задаётся начальное приближение на высоте анемометра h_a ;
- 2) рассчитывается вертикальный профиль;
- 3) обращается в нуль величина $\operatorname{div} \mathbf{A} = \operatorname{div} \mathbf{w}h$, $\mathbf{w} = (u, v)$, где h — толщина горизонтального слоя в данной точке;
- 4) производится итерационное сглаживание поля ветра.

Таблица 6. Показатель экспоненты в зависимости от класса атмосферной стабильности

Класс атмосферной стабильности	A	B	C	D	E	F	G
m	0.08	0.165	0.215	0.31	0.405	0.43	0.44

Затем повторяется минимизация дивергенции и выполняется подстановка величины скорости ветра в транспортную модель.

Разработана схема минимизации дивергенции поля ветра. Описание базируется на корректировке начального приближения для ветрового поля в каждой точке пространства между домами путём минимизации дивергенции и сглаживания. Дивергенция обращается в нуль путём проецирования заданного ветрового поля на пространство сеточных векторов. На первом этапе моделирования задаётся начальное приближение распределения векторов скоростей и минимизируется дивергенция горизонтальных компонент векторов скоростей (условие неразрывности $\operatorname{div} \mathbf{w} = 0$).

Пусть задано начальное приближение в прямоугольной области $\Omega = L_x \times L_y$ для двумерного векторного поля:

$$\mathbf{w}^o(x, y) = (u^o(x, y), v^o(x, y)). \quad (23)$$

На разностной сетке начальное приближение имеет вид:

$$\mathbf{w}_{ij}^o = (u_{i,j+1/2}^o, v_{i+1/2,j}^o), \quad (24)$$

где $i = 1, \dots, N_1$, $j = 1, \dots, N_2$.

Построим на разностной сетке двумерное векторное поле (рис. 4):

$$\mathbf{w}_{ij} = (u_{i,j+1/2}, v_{i+1/2,j}), \quad (25)$$

где $i = 1, \dots, N_1$, $j = 1, \dots, N_2$, обладающее следующими свойствами: разностная дивергенция

$$\operatorname{div} \mathbf{w}_{ij} = \frac{u_{i+1,j+1/2} - u_{i,j+1/2}}{\Delta x} + \frac{v_{i+1/2,j+1} - v_{i+1/2,j}}{\Delta y}$$

равна нулю, и отклонение $\mathbf{w}_{ij}$ от начального приближения $\mathbf{w}_{ij}^o$ минимально для всех $i = 1, \dots, N_1$, $j = 1, \dots, N_2$. С этой целью введём

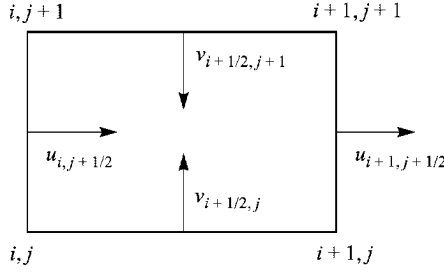


Рис. 4. Задание компонент скорости ветра на сторонах ячеек разностной сетки

следующее семейство сеточных векторных функций, значения которых будут определяться в узлах сетки:

$$\mathbf{A}_{pk} = \begin{cases} (0, 0, 0), & i \neq p \cup j \neq k, \\ (0, 0, A), & i = p \cap j = k; \quad i, p = 1, \dots, N_1, \quad j, k = 1, \dots, N_2, \end{cases} \quad (26)$$

где A — некоторая произвольная константа. Запишем разностные роты для данных функций:

$$\mathbf{F}_{pk} \equiv \text{rot } \mathbf{A}_{pk} = \begin{pmatrix} \tilde{u}_{p,k-1/2} - \tilde{u}_{p,k+1/2} \\ \tilde{v}_{p+1/2,k} - \tilde{v}_{p-1/2,k} \end{pmatrix}, \quad (27)$$

где $\tilde{u}_{p,k-1/2}$, $\tilde{u}_{p,k+1/2}$, $\tilde{v}_{p+1/2,k}$, $\tilde{v}_{p-1/2,k}$ — скалярные сеточные функции, значения которых задаются на сторонах ячеек сетки, такие что:

$$\begin{aligned} \tilde{u}_{p,k+1/2} &= \begin{cases} 0, & i \neq p \cup j \neq k, \\ A/\Delta y, & i = p \cap j = k; \quad i, p = 1, \dots, N_1, \quad j, k = 1, \dots, N_2, \end{cases} \\ \tilde{u}_{p,k-1/2} &= \begin{cases} 0, & i \neq p \cup j \neq k, \\ A/\Delta y, & i = p \cap j = k; \quad i, p = 1, \dots, N_1, \quad j, k = 1, \dots, N_2, \end{cases} \\ \tilde{v}_{p+1/2,k} &= \begin{cases} 0, & i \neq p \cup j \neq k, \\ A/\Delta x, & i = p \cap j = k; \quad i, p = 1, \dots, N_1, \quad j, k = 1, \dots, N_2, \end{cases} \\ \tilde{v}_{p-1/2,k} &= \begin{cases} 0, & i \neq p \cup j \neq k, \\ A/\Delta x, & i = p \cap j = k; \quad i, p = 1, \dots, N_1, \quad j, k = 1, \dots, N_2. \end{cases} \end{aligned} \quad (28)$$

Получившийся элементарный вихрь изображён на рис. 5.

Линейная оболочка всех элементарных роторов на разностной сетке образует некоторое линейное пространство, в котором данные роты играют роль базиса. Очевидно, любая их линейная комбинация также является ротором некоторой сеточной функции. Если функция $\text{rot } \mathbf{A}(x, y)$ принадлежит указанному линейному пространству, то её разложение по базису $\mathbf{F}_{pk}$ будет иметь вид:

$$(\text{rot } \mathbf{A})_{ij} = \left(\sum C_{pk} \mathbf{F}_{pk} \right)_{ij}, \quad (29)$$

где $i = 1, \dots, N_1$, $j = 1, \dots, N_2$.

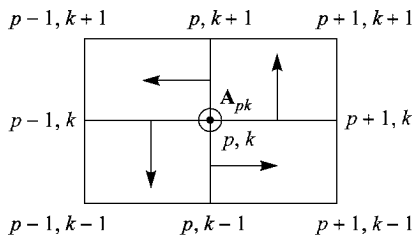


Рис. 5. Компоненты элементарного ротора на разностной сетке

Запишем начальное приближение для двумерного векторного поля в виде:

$$\mathbf{w}^\circ(x, y) = \text{rot } \mathbf{A}(x, y) + \boldsymbol{\alpha}(x, y), \quad (30)$$

где $\mathbf{A}(x, y)$ и $\boldsymbol{\alpha}(x, y)$ — некоторые векторные поля. Если спроецировать начальное приближение на описанное выше линейное пространство, то искомое векторное поле будет иметь вид:

$$\mathbf{w}(x, y) = \text{rot } \mathbf{A}(x, y). \quad (31)$$

Начальное приближение для двумерного векторного поля на разностной сетке можно записать как

$$\mathbf{w}_{ij}^\circ = (\sum C_{pk} \mathbf{F}_{pk})_{ij} + \boldsymbol{\alpha}_{ij}, \quad (32)$$

тогда искомое векторное поле запишется в виде

$$\mathbf{w}_{ij} = (\sum C_{pk} \mathbf{F}_{pk})_{ij}. \quad (33)$$

Скалярно помножим обе части уравнения (32) на величину $\mathbf{F}_{ij}$:

$$\mathbf{F}_{ij} \mathbf{w}_{ij}^\circ = \mathbf{F}_{ij} (\sum C_{pk} \mathbf{F}_{pk})_{ij} + \mathbf{F}_{ij} \boldsymbol{\alpha}_{ij}. \quad (34)$$

Для того чтобы норма функции $\boldsymbol{\alpha}$, являющейся отклонением $\mathbf{w}_{ij}$ от начального приближения $\mathbf{w}_{ij}^\circ$, была минимальной, проецирование начального приближения на линейное пространство элементарных роторов должно быть ортогональным, то есть, должно выполняться условие ортогональности $\mathbf{F}_{ij}$ второму слагаемому в правой части формулы (32) для всех i и j :

$$\mathbf{F}_{ij} \boldsymbol{\alpha}_{ij} = 0, \quad i = 1, \dots, N_1, \quad j = 1, \dots, N_2. \quad (35)$$

Из (34) и (35) имеем

$$\mathbf{F}_{ij} \mathbf{w}_{ij}^\circ = \mathbf{F}_{ij} (\sum C_{pk} \mathbf{F}_{pk})_{ij}, \quad i = 1, \dots, N_1, \quad j = 1, \dots, N_2. \quad (36)$$

Решая систему (36) $N_1 \times N_2$ уравнений, находим коэффициенты C_{pk} . Подставляя их в (33) и проведя элементарные упрощения, находим $\mathbf{w}_{ij}$:

$$\mathbf{w}_{ij} = (u_{i,j+1/2}, v_{i+1/2,j}) = \left((C_{i,j+1} - C_{ij}) \frac{A}{\Delta y}, (C_{ij} - C_{i+1,j}) \frac{A}{\Delta x} \right). \quad (37)$$

Может возникнуть ситуация, когда искомые весовые коэффициенты для элементарных вихрей, относящихся к соседним узлам сетки, не являются независимыми относительно друг друга. Рассмотрим для этого ветровое поле в качестве примера для векторного поля.

На рис. 6 изображена ячейка, в центре которой находится метеостанция. Так как наличие метеостанции предполагает значения соответствующих компонент искомого поля ветра на нижней и левой границах ячейки известными, это требует некоторой модернизации системы (36), чтобы её решение не входило в противоречие с измерениями на метеостанции. Принимая во внимание (37), можно видеть, что для этого в системе (36) уравнения для узлов $(m+1, n)$ и $(m, n+1)$ надо заменить условиями

$$\begin{aligned} (c_{mn} - c_{m+1,n}) \frac{A}{\Delta x} &= V_{m+1/2,n}, \\ (c_{m,n+1} - c_{mn}) \frac{A}{\Delta y} &= U_{m,n+1/2}. \end{aligned} \quad (38)$$

Перепишем систему (36) в виде

$$\mathbf{F}_{ij} \mathbf{w}_{ij}^{\circ} = \mathbf{F}_{ij} \left(\sum_{\substack{(p,k) \neq (m,n) \cap \\ \cap (p,k) \neq (m+1,n) \cap \\ \cap (p,k) \neq (m,n+1)}} C_{pk} \mathbf{F}_{pk} + C_{mn} \mathbf{F}_{mn} + C_{m+1,n} \mathbf{F}_{m+1,n} + C_{m,n+1} \mathbf{F}_{m,n+1} \right)_{ij} \quad (39)$$

и учтём условия (38):

$$\begin{aligned} \mathbf{F}_{ij} \mathbf{w}_{ij}^{\circ} = \mathbf{F}_{ij} \left(\sum_{\substack{(p,k) \neq (m,n) \cap \\ \cap (p,k) \neq (m+1,n) \cap \\ \cap (p,k) \neq (m,n+1)}} C_{pk} \mathbf{F}_{pk} + C_{mn} \mathbf{F}_{mn} + \right. \\ \left. + \left(C_{mn} - V_{m+1/2,n} \frac{\Delta x}{A} \right) \mathbf{F}_{m+1,n} + \left(C_{mn} + U_{m,n+1/2} \frac{\Delta y}{A} \right) \mathbf{F}_{m,n+1} \right)_{ij}. \end{aligned} \quad (40)$$

Окончательно преобразуем (40):

$$\begin{aligned} \mathbf{F}_{ij} \mathbf{w}_{ij}^{\circ} + V_{m+1/2,n} \frac{\Delta x}{A} \mathbf{F}_{m+1,n} - U_{m,n+1/2} \frac{\Delta y}{A} \mathbf{F}_{m,n+1} = \\ = \mathbf{F}_{ij} \left(\sum_{\substack{(p,k) \neq (m,n) \cap \\ \cap (p,k) \neq (m+1,n) \cap \\ \cap (p,k) \neq (m,n+1)}} C_{pk} \mathbf{F}_{pk} + C_{mn} (\mathbf{F}_{mn} + \mathbf{F}_{m+1,n} + \mathbf{F}_{m,n+1}) \right)_{ij}. \end{aligned} \quad (41)$$

Из (41) следует, что в формуле (32) нужно требовать ортогональность несоленоидальному слагаемому в правой части не от элементарного

вихря $\mathbf{F}_{mn}$, а от выражения $\mathbf{F}_{mn} + \mathbf{F}_{m+1,n} + \mathbf{F}_{m,n+1}$, так как именно оно стоит под знаком суммы. Новый базисный вихрь $\mathbf{F}_{mn} + \mathbf{F}_{m+1,n} + \mathbf{F}_{m,n+1}$ показан на рис. 6 белыми стрелками.

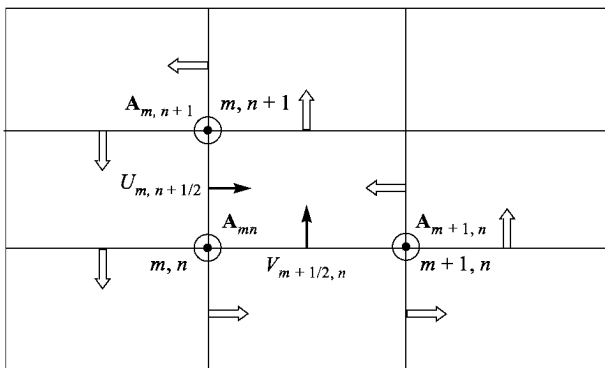


Рис. 6. Базисный вихрь при введении в ячейке дополнительного ограничения на ветровое поле, связанного с метеозмерениями

Другой случай, требующий введения изменений в систему (36), обусловлен наличием возвышений рельефа и зданий. Рассмотрим пример, проиллюстрированный на рис. 7. Здание обозначено заштрихованной областью. Так как ветер внутри него должен отсутствовать, а компоненты ветра, перпендикулярные стенам, также должны обращаться в нуль, то все коэффициенты C_{ij} в системе (36), относящиеся к элементарным вихрям вокруг узлов расчётной сетки, находящихся внутри и на границе здания, должны быть равны между собой:

$$C_{mn} = C_{m+1,n} = C_{m+2,n} = C_{m,n+1} = C_{m+1,n+1} = C_{m+2,n+1} = C_{m,n+2} = \\ = C_{m+1,n+2} = C_{m+2,n+2} = C_{m,n+3} = C_{m+1,n+3} = C_{m+2,n+3}. \quad (42)$$

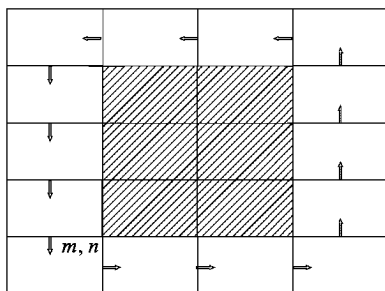


Рис. 7. Базисный вихрь при введении в нескольких ячейках дополнительного ограничения на ветровое поле, связанного с наличием здания

Проведя преобразования, аналогичные (39)–(41), получим для ситуации, изображённой на рис. 7, следующий новый базисный вихрь (на рисунке он обозначен стрелками):

$$\mathbf{F}_{mn} + \mathbf{F}_{m+1,n} + \mathbf{F}_{m+2,n} + \mathbf{F}_{m,n+1} + \mathbf{F}_{m+1,n+1} + \mathbf{F}_{m+2,n+1} + \mathbf{F}_{m,n+2} + \\ + \mathbf{F}_{m+1,n+2} + \mathbf{F}_{m+2,n+2} + \mathbf{F}_{m,n+3} + \mathbf{F}_{m+1,n+3} + \mathbf{F}_{m+2,n+3},$$

а окончательная система уравнений после применения формулы (42) и выражения всех зависимых коэффициентов через C_{mn} примет вид:

$$\mathbf{F}_{ij} \mathbf{w}_{ij}^{\circ} = \mathbf{F}_{ij} \left(\sum_{\substack{(p,k) \neq (m,n) \cap \dots \\ \dots \cap (p,k) \neq (m+2,n+2)}} C_{pk} \mathbf{F}_{pk} + C_{mn} (\mathbf{F}_{mn} + \mathbf{F}_{m+1,n} + \right. \\ \left. + \mathbf{F}_{m+2,n} + \mathbf{F}_{m,n+1} + \mathbf{F}_{m+1,n+1} + \mathbf{F}_{m+2,n+1} + \right. \\ \left. + \mathbf{F}_{m,n+2} + \mathbf{F}_{m+1,n+2} + \mathbf{F}_{m+2,n+2} + \mathbf{F}_{m,n+3} + \mathbf{F}_{m+1,n+3} + \mathbf{F}_{m+2,n+3}) \right)_{ij}. \quad (43)$$

Системы (36), (41) и (43) переопределены, так как искомые коэффициенты определяются с точностью до аддитивной константы. Поэтому для определённости заменим, например, уравнение для узла (0, 0) условием $C_{00} = 0$ и исключим его из систем (36), (41) и (43). Теперь эти системы могут быть решены с помощью метода сопряжённых градиентов, так как их матрицы симметричны и положительно определены. Затем, используя (38) и (42), находим значения остальных коэффициентов.

При наличии нескольких метеостанций, возвышений рельефа или их комбинаций, обуславливающих наличие дополнительных связей между весовыми коэффициентами для соседних элементарных вихрей, все вычисления аналогичны вышеприведённым. Причём, для каждой серии таких вихрей возникает новый базисный вихрь, являющийся их векторной суммой.

На рис. 8–11 представлены результаты построения ветрового поля при наличии на пути воздушного потока зданий для различных азимутов направления ветра и конфигурации застройки. Существенным для URBAN — модели является выбор сетки. Её размеры должны обеспечивать правильную картину разделения потоков вокруг домов, т. е. обеспечивать достаточную точность решений.

Диагностическая модель поля ветра над слабо холмистым ландшафтом (модель DWM) существенно отличается от описания обтекания крутых препятствий. В этой модели применены некоторые энергетические характеристики распределения потоков вокруг и сверх холмов.

Поле ветра рассчитывается по следующему алгоритму:

- 1) задаётся начальное приближение: $u(x, y, z) = u_0$, $v(x, y, z) = v_0$;
- 2) рассчитывается вертикальный профиль;

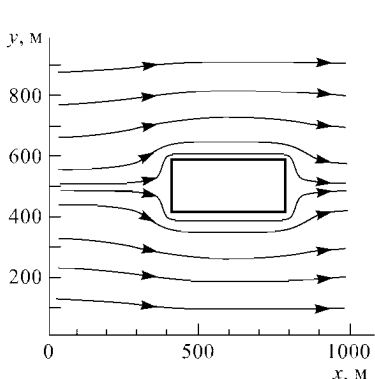


Рис. 8. Ветровое поле при наличии одного дома. Преобладающее направление ветра перпендикулярно короткой стене дома

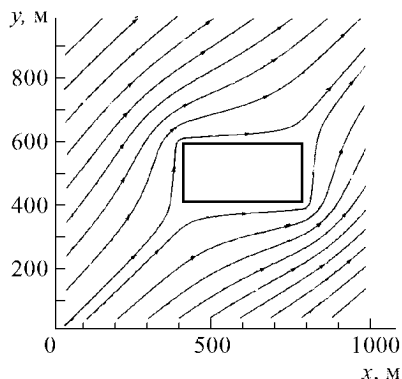


Рис. 9. Ветровое поле при наличии одного дома. Преобладающее направление ветра составляет угол 45 градусов со стенами дома

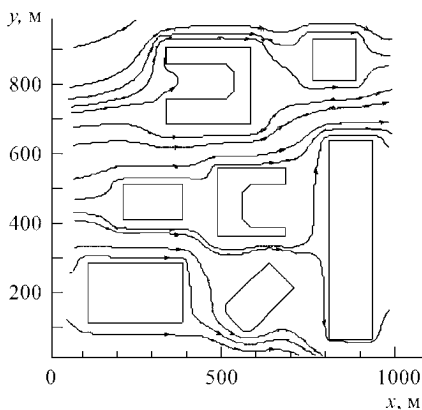


Рис. 10. Ветровое поле в области с семью домами различной конфигурации и ориентации

- 3) итерационно минимизируется величина $\text{div } \mathbf{A} = \text{div } \mathbf{w}h$, где $\mathbf{w} = (u, v)$, h — толщина горизонтального слоя в данной точке;
- 4) рассчитывается влияние холмов;
- 5) учитывается поворот ветра у препятствия в зависимости от числа Фруда;
- 6) производится итерационное сглаживание поля ветра;
- 7) осуществляется повторение п. 3.

2.4. Решение системы уравнений модели. Система уравнений транспортно-диффузионной модели решается методом физического расщепления на процессы. Разработаны две схемы расщепления.

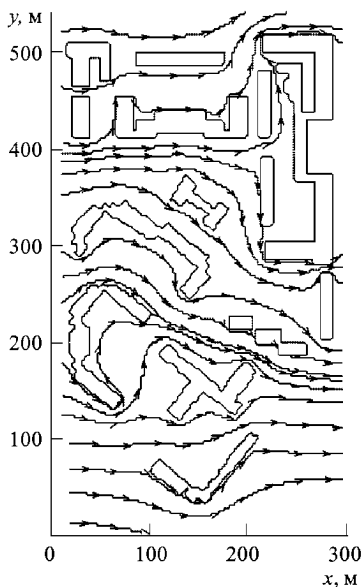


Рис. 11. Ветровое поле в одном из кварталов Москвы

2.4.1. Схема расщепления 1 на основе сеточно-характеристического метода. Для схемы расщепления 1 был разработан сеточно-характеристический метод, адаптированный к конкретной проблеме построения мониторинговой системы.

В данной схеме выделяются следующие процессы: перенос ветром вдоль оси X и Y , горизонтальная турбулентная диффузия и вертикальная турбулентная диффузия. Запишем закон сохранения вещества в интегральной форме для каждого из них.

1. Для переноса ветром вдоль оси X :

$$h_i^{n+1} c_{2,i}^{(1)} - h_i^n c_{2,i}^{(0)} = - \int_t^{t+\Delta t} \frac{\partial h c_{2,i} u_i}{\partial x} dt,$$

$$h_i^{n+1} c_{2,i}^{(1)} - h_i^n c_{2,i}^{(0)} = - \int_t^{t+\Delta t} \frac{\partial h c_{2,i} u_i}{\partial x} dt;$$

2. Для переноса ветром вдоль оси Y :

$$h_i^{n+1} c_{1,i}^{(2)} - h_i^n c_{1,i}^{(1)} = - \int_t^{t+\Delta t} \frac{\partial h c_{1,i} v_i}{\partial y} dt,$$

$$h_i^{n+1} c_{2,i}^{(2)} - h_i^n c_{2,i}^{(1)} = - \int_t^{t+\Delta t} \frac{\partial h_i c_{2,i} v_i}{\partial y} dt;$$

3. Для горизонтальной турбулентной диффузии:

$$\begin{aligned} h_i^{n+1} c_{1,i}^{(3)} - h_i^n c_{1,i}^{(2)} &= h_i^{n+1} c_{2,i}^{(3)} - h_i^n c_{2,i}^{(2)} = \int_t^{t+\Delta t} (\operatorname{div} (h_i^{n+1} K_i^{n+1} \operatorname{grad} c_{1,i}^{(3)}) - \\ &- (v_{d1} + k_{\omega_1} h_i^{n+1} + k_l h_i^{n+1}) c_{1,i}^{(3)} + (1 - \beta) h_i^{n+1} \bar{Q}_i^{n+1}) dt = \\ &= \int_t^{t+\Delta t} (\operatorname{div} (h_i K_i \operatorname{grad} c_{1,i}) - (v_{d1} + k_{\omega_1} h_i + k_l h_i) c_{1,i} + (1 - \beta) h_i \bar{Q}_i) dt, \end{aligned}$$

$$\begin{aligned} h_i^{n+1} c_{2,i}^{(3)} - h_i^n c_{2,i}^{(2)} &= \\ &= \int_t^{t+\Delta t} \operatorname{div} (h_i K_i \operatorname{grad} c_{2,i}) - (v_{d2} + k_{\omega_2} h_i) c_{2,i} + k_l h_i c_{1,i} + \beta h_i \bar{Q}_i dt; \end{aligned}$$

4. Для вертикальной турбулентной диффузии:

$$\begin{aligned} h_i^{n+1} c_{1,i}^{(4)} - h_i^n c_{1,i}^{(3)} &= \int_t^{t+\Delta t} D_1 dt, \\ h_i^{n+1} c_{2,i}^{(4)} - h_i^n c_{2,i}^{(3)} &= \int_t^{t+\Delta t} D_2 dt, \end{aligned}$$

где n — номер временного слоя, Δt — величина шага по времени, причём $t = t^n$ и $t + \Delta t = t^{n+1}$, индекс (0) означает, что значение концентрации берётся с предыдущего (n -го) временного слоя. Из выражения для переноса ветром вдоль оси X находится значение концентрации $c^{(1)}$ и подставляется в выражение для переноса ветром вдоль оси Y и т. д., пока не найдено значение концентрации $c^{(4)}$, которое соответствует значению концентрации на текущем ($n + 1$)-м временном слое.

2.5. Построение разностной сетки. При построении разностной сетки трёхмерная область разбивается на ячейки переменной высоты (рис. 12).

Во всех разностных схемах к граням и к узлам ячеек относятся разные переменные, как показано на рис. 13, на котором изображена ячейка разностной сетки с пронумерованными узлами ($i = 0, \dots, nx$, $j = 0, \dots, ny$, $m = 1, \dots, nz + 1$).

Перенос ветром вдоль оси X описывается уравнением вида:

$$\frac{\partial f}{\partial t} + \frac{\partial f u}{\partial x} = 0.$$

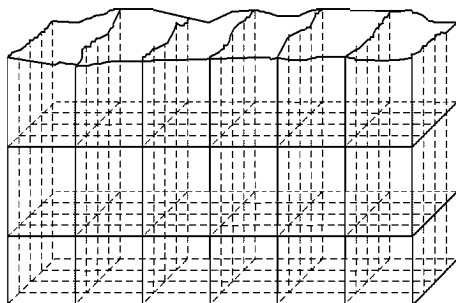


Рис. 12. Построение разностной сетки при верхнем слое переменной толщины

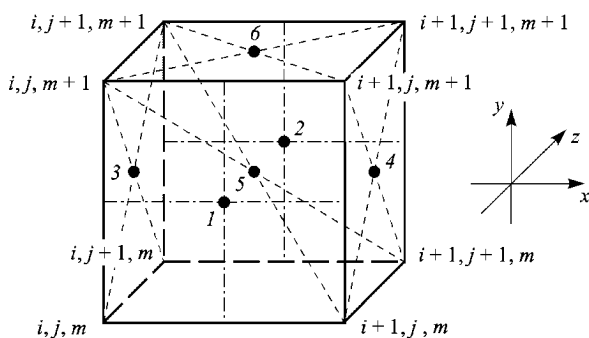


Рис. 13. Задание физических величин в различных точках ячейки разностной сетки

Таблица 7. Соотнесения переменных разностной схемы с различными точками ячейки

Номер точки на рис. 13	Переменные разностной схемы
1	v_{ijm}
2	$v_{i,j+1,m}$
3	u_{ijm}
4	$u_{i+1,j,m}$
5	$q_{ijm}, c_{ijm}, h_{ijm}, K_{ijm}, u_{gijm}, v_{gijm}$
6	K_{vijm}

Значения величины f на разностной сетке относятся к центрам ячеек, а значения компонент ветра — к граням, как показано на рис. 15. Метод проиллюстрирован на рис. 14 а–в.

На рис. 14 а–в по горизонтали отложены номера узлов ячеек вдоль оси X , а по вертикали — номера временных слоёв. При заданном

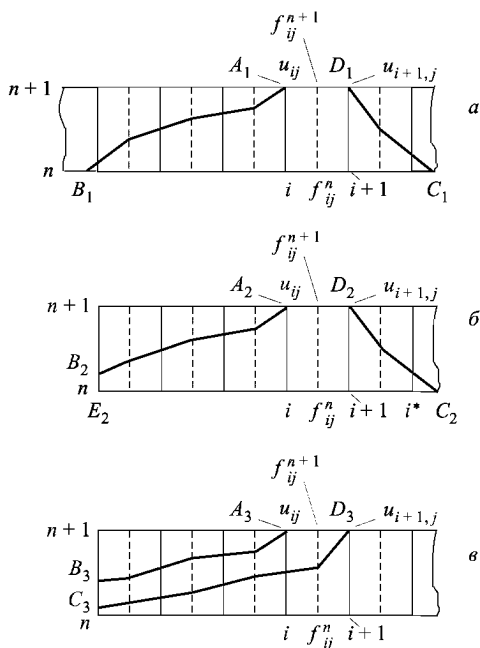


Рис. 14. Иллюстрация сеточного характеристического метода. Из двух характеристик линию $t = t^n$ пересекают: а) обе; б) одна; в) ни одна

значении f_{ij}^n на n -м слое находится значение f_{ij}^{n+1} на $(n+1)$ -м слое. Вертикальные сплошные линии соответствуют узлам, в которых заданы скорости, а пунктирные линии — центрам ячеек, в которых задана величина f . Из точек А и D опускаются характеристики — кривые, определяемые уравнением $\frac{dx}{dt} = u(x, t)$. При построении характеристик получают вспомогательные точки: $A_1B_1C_1D_1$ — рис. 14 а, $A_2B_2C_2E_2D_2$ — рис. 14 б, $A_3B_3C_3D_3$ — рис. 14 в

$$\int_{D_1}^{A_1} f dx + \int_{B_1}^{C_1} f dx = 0,$$

$$\int_{D_2}^{A_2} f dx + \int_{B_2}^{E_2} u f dt + \int_{B_2}^{C_2} f dx = 0,$$

$$\int_{D_3}^{A_3} f dx + \int_{B_3}^{C_3} u f dt = 0.$$

Интегралы записываются в виде разностных схем, откуда находятся искомые значения для концентрации загрязнения. Например, разност-

ная схема следующая:

$$\sum_{i=0}^{i^*-1} f_{ij}^n \Delta x + f_{ij}^n (x(C_2) - i^* \Delta x) - f_{ij}^{n+1} \Delta x - f_{0j}^{n+1} (t(B_2) - 0) \cdot u_{0j} = 0,$$

откуда

$$f_{ij}^{n+1} = \sum_{i=0}^{i^*-1} f_{ij}^n + \frac{f_{ij}^n (x(C_2) - i^* \Delta x) - f_{0j}^{n+1} t(B_2) u_{0j}}{\Delta x}.$$

Аналогично записываются выражения и для других случаев. Далее вводятся $f^n = h^n c^{(0)}$ и $f^{n+1} = h^{n+1} c^{(1)}$. По известному значению $c^{(0)}$ находится $c^{(1)}$.

Перенос ветром вдоль оси Y описывается аналогично переносу вдоль оси X . Уравнение переноса ветром вдоль оси Y :

$$\frac{\partial f}{\partial t} + \frac{\partial f v}{\partial y} = 0.$$

Вводятся $f^n = h^n c^{(1)}$ и $f^{n+1} = h^{n+1} c^{(2)}$. По известному значению $c^{(1)}$ находится $c^{(2)}$.

Пояснения для обеих методик приведены на рис. 15.

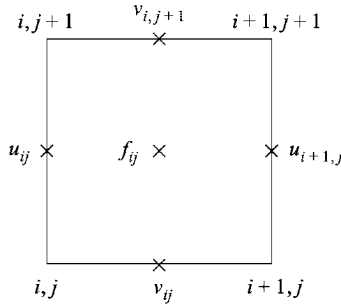


Рис. 15. Задание проекций векторного поля ветра и искомой скалярной величины в ячейке разностной сетки при решении уравнений адвекции

Разностная схема для горизонтальной турбулентной диффузии:

$$PM_{ij} c_{ij}^{(3)} + AM_{ij} c_{i-1,j}^{(3)} + AM1_{ij} c_{i+1,j}^{(3)} + BM_{ij} c_{i,j-1}^{(3)} + BM1_{ij} c_{i,j+1}^{(3)} = FM_{ij}.$$

Во внутренних точках области ($i = 1, \dots, nx - 2$, $j = 1, \dots, ny - 2$):

$$AM_{ij} = 0,5 \frac{K_{ij}^{n+1} h_{ij}^{n+1} + K_{i-1,j}^{n+1} h_{i-1,j}^{n+1}}{(\Delta x)^2} \cdot \Delta t,$$

$$AM1_{ij} = 0,5 \frac{K_{ij}^{n+1} h_{ij}^{n+1} + K_{i+1,j}^{n+1} h_{i+1,j}^{n+1}}{(\Delta x)^2} \cdot \Delta t,$$

$$BM_{ij} = 0,5 \frac{K_{ij}^{n+1} h_{ij}^{n+1} + K_{i,j-1}^{n+1} h_{i,j-1}^{n+1}}{(\Delta y)^2} \cdot \Delta t,$$

$$BM1_{ij} = 0,5 \frac{K_{ij}^{n+1} h_{ij}^{n+1} + K_{i,j+1}^{n+1} h_{i,j+1}^{n+1}}{(\Delta y)^2} \cdot \Delta t,$$

$$PM_{ij} = h_{ij}^{n+1} + AM_{ij} + AM1_{ij} + BM_{ij} + BM1_{ij} + (v_{d1} + k_{\omega 1} h_{ij}^{n+1} + k_t h_{ij}^{n+1}) \Delta t,$$

$$FM_{ij} = (1 - \beta) h_{ij}^{n+1} q_{ij}^{n+1} \cdot \Delta t + c_{1ij}^{(2)} h_{ij}^n \quad \text{для } \text{SO}_2,$$

$$PM_{ij} = h_{ij}^{n+1} + AM_{ij} + AM1_{ij} + BM_{ij} + BM1_{ij} + (v_{d2} + k_{\omega 2} h_{ij}^{n+1}) \Delta t,$$

$$FM_{ij} = \beta h_{ij}^{n+1} q_{ij}^{n+1} \cdot \Delta t + k_i h_{ij}^{n+1} c_{1ij}^{(3)} \cdot \Delta t + c_{2ij}^{(2)} h_{ij}^n \quad \text{для } \text{SO}_4^-.$$

В граничных точках области:

для $i = 1, \dots, nx - 2$:

$$AM_{i0} = 0, \quad BM_{i0} = 0, \quad AM1_{i0} = 0,$$

$$BM1_{i0} = \frac{\beta_{Bi}}{\Delta y}, \quad FM_{i0} = \gamma_{Bi}, \quad PM_{i0} = \alpha_{Bi} + BM1_{i0},$$

$$AM_{i,ny-1} = 0, \quad BM_{i,ny-1} = \frac{\beta_{Ti}}{\Delta y}, \quad AM1_{i,ny-1} = 0,$$

$$BM1_{i,ny-1} = 0, \quad FM_{i,ny-1} = \gamma_{Ti}, \quad PM_{i,ny-1} = \alpha_{Ti} + BM_{i,ny-1};$$

для $j = 0, \dots, ny - 1$:

$$AM_{0j} = 0, \quad BM_{0j} = 0, \quad AM1_{0j} = \frac{\beta_{Lj}}{\Delta x},$$

$$BM1_{0j} = 0, \quad FM_{0j} = \gamma_{Lj}, \quad PM_{0j} = \alpha_{Lj} + AM1_{0j},$$

$$AM_{nx-1,j} = \frac{\beta_{Rj}}{\Delta x}, \quad BM_{nx-1,j} = 0, \quad AM1_{nx-1,j} = 0,$$

$$BM1_{0j} = 0, \quad FM_{nx-1,j} = \gamma_{Rj}, \quad PM_{nx-1,j} = \alpha_{Rj} + AM_{nx-1,j},$$

индексы T, L, R, B означают, что соответствующие значения $\alpha_\Omega, \beta_\Omega$ и γ_Ω относятся к различным сторонам прямоугольной границы.

Решение данной пятидиагональной системы полностью неявным методом даёт значение $c^{(3)}$.

Вертикальная турбулентная диффузия описывается одномерным параболическим уравнением с условиями второго рода на верхней и нижней границах. Для выбранной разностной сетки схема, описывающая процесс вертикальной турбулентности:

$$\frac{c_m^{(4)} - c_m^{(3)}}{\Delta t} = \frac{B_m - A_m}{h_m^{n+1}}, \quad m = 1, \dots, 3,$$

где

$$A_m = K_{V_{m-1}}^{n+1} \frac{c_m^{(4)} - c_{m-1}^{(4)}}{0,5(h_m^{n+1} + h_{m-1}^{n+1})}, \quad m = 2, \dots, 3,$$

$$B_m = K_{V_m}^{n+1} \frac{c_{m+1}^{(4)} - c_m^{(4)}}{0,5(h_{m+1}^{n+1} + h_m^{n+1})}, \quad m = 1, \dots, 2,$$

$$A_1 = 0, \quad B_3 = 0.$$

Решение данной системы методом прогонки даёт значение $c^{(4)}$.

Однако такой метод расщепления приводит к большим вычислительным погрешностям, которые особенно очевидны при расчёте на прямоугольной сетке распространения потока загрязняющих веществ под действием ветра, направленного по оси X , и под углом к оси X (например, по диагонали).

Для иллюстрации рассмотрим одномерный случай для наблюдаемого на практике перемещения загрязнений в ветровом поле (рис. 16). Пусть значение ветра постоянно вдоль оси X и не изменяется во времени. Выберем шаг по времени $\Delta t = \Delta x / (2u)$, где Δx — шаг по пространству, u — значение ветра. На рис. 16 по вертикали отмечены номера временных слоёв. Пусть на начальном, нулевом временном слое, концентрация вещества во всех ячейках, кроме крайней левой, равна 0, а в крайней левой — 1 условная единица. Тогда на следующем, 1-м временном слое половина вещества из крайней левой ячейки перейдёт в соседнюю с ней ячейку. На 2-м временном слое уже всё вещество покинет крайнюю левую ячейку и т. д. Процесс приблизительно адекватно отражается на рис. 16. Сеточно-характеристический метод предполагает, что на каждом временном шаге после того, как вещество заполняет часть ячейки разностной сетки, происходит его равномерное распределение по этой ячейке. Так, вначале на 1-м временном шаге в крайней левой и соседней с ней ячейках вещество занимает половину площади каждой ячейки (2-я строка снизу на рис. 17), а затем равномерно перераспределяется по этим ячейкам. Причём, концентрация в каждой из них составляет 0,5 условных единиц (3-я строка снизу на рис. 17). За счёт равномерного распределения вещества по ячейкам передняя граница фронта вещество-пустота движется в 2 раза быстрее, чем в реальности. Задняя граница при этом покоится, так как на каждом шаге оставшееся в крайней левой ячейке вещество распределяется по всей её площади, отодвигаясь к её левой границе. Очевидно, на n -м временном шаге вещество будет занимать $(n + 1)$ ячеек. Причём, его концентрация в крайней левой ячейке и крайней правой, в которой присутствует вещество, составит 2^{-n} условных единиц, а между ними значения будут изменяться, достигая максимума посередине.

В случае наличия второй компоненты ветра вычислительная диффузия будет ещё интенсивнее за счёт одновременного распределения вещества по двум направлениям. Таким образом, происходит размазывание вещества по пространству. Уже на 3-м шаге вместо максимальной концентрации 1 условная единица концентрации будет составлять $3/8$ условных единицы.

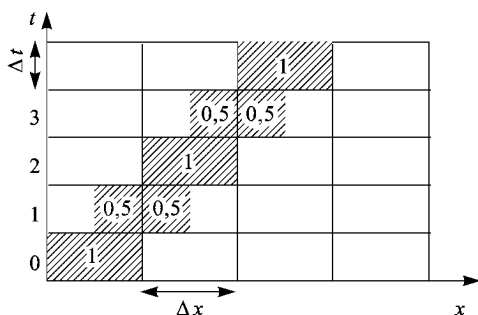


Рис. 16. Реальное распределение концентраций примеси с течением времени в одномерном однородном ветровом поле

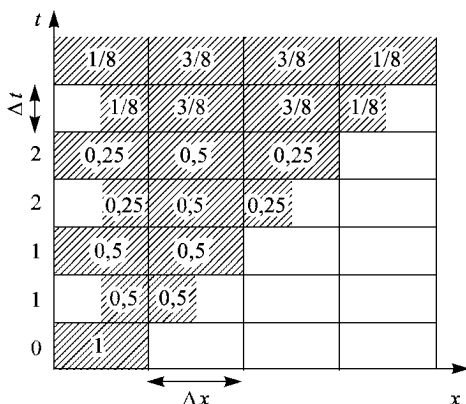


Рис. 17. Распределение концентраций примеси с течением времени в одномерном однородном ветровом поле по результатам расчёта с помощью сеточно-характеристического метода

Недостатки вышеупомянутых вычислительных схем побудили к разработке специально адаптированного для данной задачи метода моделирования переноса.

2.6. Схема расщепления по процессам на основе метода частиц в ячейках. Новая схема расчёта полей концентраций при переносе в поле ветра должна была позволить быстро и эффективно рассчитывать концентрации в любой точке пространства и времени, но, главное, избежать принципиальных счётных ошибок при построении алгоритмов переноса, свойственных характеристическому методу и явным схемам с разнообразными компенсационными схемами.

Отправной моделью для создания схемы расщепления 2 по физическим процессам послужил классический метод частиц в ячейке [21].

Однако по принятым допущениям он оригинален и не имеет аналогов в литературе.

Пусть имеется некоторое распределение аддитивной неотрицательной скалярной величины f в пространстве (например, концентрации примеси в атмосфере). На разностной сетке оно будет отражено матрицей

$$(f_{ij})_{N_x N_y},$$

где f_{ij} — значение f в ячейке с номером i, j , где $i = 0, \dots, N_x - 1$, $j = 0, \dots, N_y - 1$. Назовём «частицей» структуру типа (x_k, y_k, f_k) , где x_k и y_k — её координаты, а f_k — некоторое весовое значение величины f , относящееся к частице K (k — номер частицы). Размеры частицы считаются нулевыми. Пусть в каждой ячейке сетки с номером ij содержатся N_{ij} частиц, тогда, по определению,

$$f_{ij} = \sum_{k_{ij}=0}^{N_{ij}-1} f_{k_{ij}}, \quad (44)$$

где индекс ij показывает, что суммирование производится только по тем частицам, которые находятся в ячейке с номером ij . Если $x_{k_{ij}}^n$, $y_{k_{ij}}^n$ — координаты частицы в момент времени t^n (везде далее верхний индекс означает номер шага по времени), то её координаты в момент времени t^{n+1} вычисляются с помощью итераций

$$x_{k_{ij}}^{n+1} \approx x_{k_{ij}}^n + u_{k_{ij}} \Delta t, \quad y_{k_{ij}}^{n+1} \approx y_{k_{ij}}^n + v_{k_{ij}} \Delta t, \quad (45)$$

где $u_{k_{ij}}$, $v_{k_{ij}}$ — компоненты скорости ветра в точке с координатами $x_{k_{ij}}^n$, $y_{k_{ij}}^n$, Δt — шаг по времени. На каждом временном шаге итерации проводятся для всех частиц. Далее проводится проверка, какие частицы в результате переноса перешли в другую ячейку за интервал времени Δt . Например, если частица с номером k_{ij} перешла в ячейку lm , то частица привязывается к новой ячейке, причём число N_{ij} уменьшается на единицу, число N_{lm} увеличивается на единицу. Если частица вышла за границы области, она уничтожается. После этого по формуле (44) вычисляется новое значение величины f на разностной сетке:

$$f_{ij}^{n(1)} = \sum_{k_{ij}=0}^{N_{ij}^{n(1)}-1} f_{k_{ij}}^{n(1)}, \quad (46)$$

где $i = 0, \dots, N_x - 1$, $j = 0, \dots, N_y - 1$, а индекс (1) означает, что координаты частиц вычислены для временного шага t^{n+1} , но не учтён вклад не адвективных процессов, влияющий на значения весовых функций и распределение величины f на разностной сетке, т.е. указанные значения не изменяются по сравнению с временным шагом t^n .

Различные не адвективные процессы (турбулентная диффузия, химические реакции и т.д.) моделируются для этого распределения с помощью численных методов на временном шаге t^{n+1} , откуда находится

распределение $(f_{ij}^{n+1})_{N_x N_y}$, после чего осуществляется переход к следующему шагу. Перед этим необходимо провести обратное преобразование: из нового распределения величины $(f_{ij}^{n+1})_{N_x N_y}$ получить распределение по частицам. Для тех ячеек, в которых после моделирования адвекции осталось по одной частице,

$$f_{0,ij}^{n+1} = f_{ij}^{n+1}, \quad (47)$$

и при этом координаты частиц остаются прежними. Для тех же ячеек, в которых после моделирования адвекции находится по несколько частиц (их координаты также остаются прежними), сумма значений их весовых функций будет иметь уже значение f_{ij}^{n+1} , а сами значения — те же пропорции по отношению к значению величины f в ячейке, т. е.

$$\frac{f_{k_{ij}^{n(1)}}}{f_{ij}^{n(1)}} = \frac{f_{k_{ij}}^{n+1}}{f_{ij}^{n+1}}$$

для любого $k_{ij} = 0, \dots, N_{ij}^{n+1} - 1$.

Положим, что в начальный момент времени

$$N_{ij}^0 = \begin{cases} 0, & f_{ij}^0 = 0, \\ 1, & f_{0,ij}^0 = f_{ij}^0, \quad f_{ij}^0 > 0, \end{cases} \quad i = 0, \dots, N_x - 1, \quad j = 0, \dots, N_y - 1, \quad (48)$$

то есть, в ячейках, где $f_{ij}^0 > 0$, содержится по одной частице, такой частице соответствует весовое значение f_{ij}^0 , и она находится в центре соответствующей ячейки. Метод предполагает возможность «рождения» в ячейке новой частицы в результате какого-либо не адвективного процесса (например, диффузии) в случае, когда $f_{ij}^{n(1)} = 0$, а $f_{ij}^{n+1} > 0$. Тогда создаётся новая частица с координатами в центре соответствующей ячейки и $f_{0,ij}^{n+1} = f_{ij}^{n+1}$.

Рассмотрим применение метода на конкретном примере. Пусть ветер задаётся в центрах граней ячеек, как показано на рис. 15. В любой точке координатной плоскости для каждой компоненты ветра можно выделить 4 точки. Рассмотрим фрагмент расчётной сетки (рис. 18). Возьмём точку с координатами (x, y) и обозначим соответствующие точки для компоненты $u(x_0, y_0)$, через (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , а для компоненты v — (x_4, y_4) , (x_5, y_5) , (x_6, y_6) , (x_7, y_7) (рис. 18). Значения величин u и v в них, соответственно, $u_0, u_1, u_2, u_3, v_4, v_5, v_6, v_7$. Значения u и v в точке (x, y) вычисляются по интерполяционным формулам: $u(x, y) = a + bx + cy + dxy$, $v(x, y) = e + fx + gy + hxy$. Коэффициенты a, b, c, d, e, f, g, h определяются из системы соотношений $u(x_0, y_0) = u_0$, $u(x_1, y_1) = u_1$, $u(x_2, y_2) = u_2$, $u(x_3, y_3) = u_3$, $v(x_4, y_4) = v_4$, $v(x_5, y_5) = v_5$, $v(x_6, y_6) = v_6$, $v(x_7, y_7) = v_7$. Для получения координат частиц интерполированные значения компонент ветра подставляются в формулу (45).

Полученные таким образом координаты частиц, которым соответствуют определённые распределения концентраций примеси, подстав-

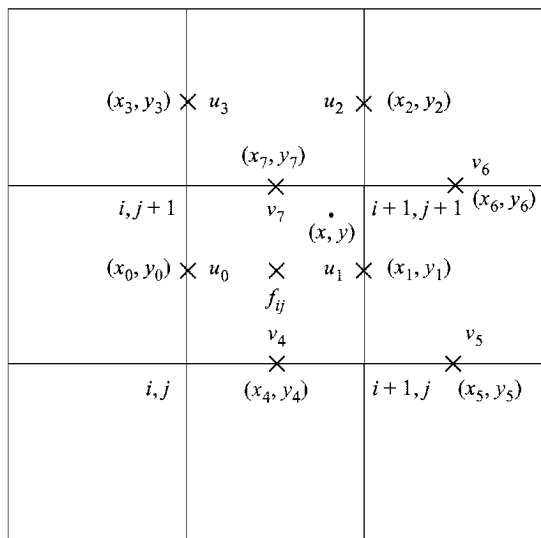


Рис. 18. Фрагмент расчётной сетки. Здесь интерполяция ветра в специальном методе частиц соответствует точкам

ляются в соответствующее место программы пакета «TIMES». С помощью численных методов для этого распределения на текущем временном шаге моделируются различные не адвективные процессы (турбулентная диффузия, физико-химические процессы и т. д.), находится распределение $(f_{ij}^{n+1})_{N_x, N_y}$ и осуществляется переход к новому временному шагу. При расчёте полагаем в пределах каждого атмосферного слоя $f_{ij} \equiv c_{ij} h_{ij}$, где h — толщина слоя, c — концентрация.

С целью получения более точного решения для поставленной задачи нахождения поля концентрации загрязняющего вещества в атмосфере описанный выше метод частиц в ячейке был усовершенствован путём введения в рассмотрение частиц с ненулевыми размерами. Здесь под «частицей» понимается структура типа (x_k, y_k, f_k) , где x_k и y_k — её координаты, а f_k — некоторое весовое значение величины f , относящееся к частице K (k — номер частицы). Причём, частица имеет размеры ячейки расчётной сетки $\Delta X \times \Delta Y$, где ΔX и ΔY — шаг сетки в направлении X и Y , соответственно.

Крупная частица может располагаться одновременно в нескольких ячейках — одной, двух или четырёх (рис. 19).

Пусть в каждой ячейке сетки с номером ij содержатся полностью или частично N_{ij} частиц, тогда

$$f_{ij} = \sum_{k_{ij}=0}^{N_{ij}-1} \alpha_{k_{ij}} f_{k_{ij}} = \sum_{k_{ij}=0}^{N_{ij}-1} f_{k_{ij}}. \quad (49)$$

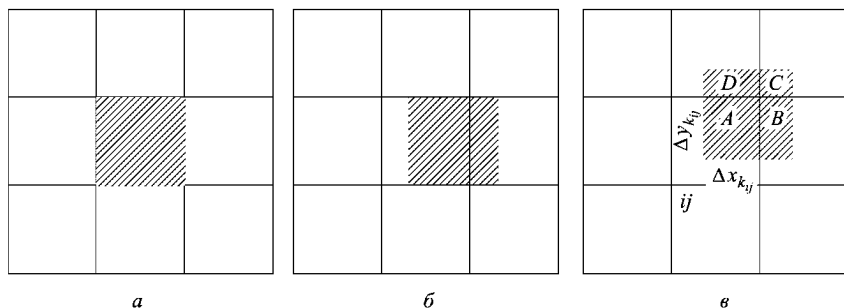


Рис. 19. Различные варианты расположения крупной частицы по отношению к разностной сетке

Символ $\circ$ над весовыми значениями частиц указывает на то, что в расчёте используется только сектор частицы, находящийся в ячейке ij : $\alpha_{k_{ij}} = \frac{(\Delta x)_{k_{ij}} (\Delta y)_{k_{ij}}}{\Delta x \Delta y}$ — относительная доля общего размера частицы, находящаяся в указанной ячейке (рис. 19 в).

Координаты центра крупной частицы в момент времени t^{n+1} вычисляются по формуле (41). После этого вычисляется новое значение величины f на разностной сетке по формуле

$$f_{ij}^{n(1)} = \sum_{k_{ij}=0}^{N_{ij}^{n+1}-1} \alpha_{k_{ij}}^{n+1} f_{k_{ij}}^{n(1)}, \quad (50)$$

где $i = 0, \dots, N_x - 1$, $j = 0, \dots, N_y - 1$, индекс (1) по-прежнему означает, что координаты частиц вычислены для временного шага t^{n+1} , но без вклада не адвективных процессов.

После моделирования не адвективных процессов обратное преобразование — получение распределения по частицам из нового распределения величины $(f_{ij}^{n+1})_{N_x N_y}$ — производится так же, как и для точечных частиц. Однако здесь необходимо учитывать, что крупная частица может включать в себя до 4-х участков, принадлежащих разным ячейкам сетки. При этом не адвективные процессы не влияют на координаты центров частиц. Для тех ячеек, в которых после моделирования адвекции находится по несколько частиц, сумма значений «вкладов» их весовых функций будет иметь значение f_{ij}^{n+1} , а сами значения — те же пропорции по отношению к значению величины f в ячейке, т.е.

$$\frac{f_{ij}^{n(1)}}{f_{ij}^{n(1)}} = \frac{f_{k_{ij}}^{n+1}}{f_{ij}^{n+1}}$$

для любого $k_{ij} = 0, \dots, N_{ij}^{n+1} - 1$.

Обозначим за T параметр, показывающий, какой из секторов частицы рассматривается. Например, на рис. 19 T может принимать значения: A (левый нижний сектор), B (правый нижний сектор), C (правый верхний сектор) и D (левый верхний сектор). Пусть $f_k(T)$ — весовое значение величины f в секторе T частицы K (k — номер частицы) после моделирования не адвективных процессов, тогда общее весовое значение величины f для этой частицы составит после $(n + 1)$ -го шага по времени

$$f_k^{n+1} = \sum_{T=\{A,B,C,D\}} f_k^{n+1}(T). \quad (51)$$

§ 3. Тестирование разработанного комплекса программ

Было проведено тестирование разработанного комплекса программ. На рис. 20 и 21 показаны шлейфы от загрязняющих веществ, образующиеся в процессе работы непрерывно действующего точечного источника в застройке из семи домов при различных направлениях ветра. Из рисунков видно, что шлейф имеет одинаковую ширину как при преобладающем направлении ветра вдоль оси X , так и при преобладающем направлении ветра под углом 45 градусов к ней. На рис. 22 показан шлейф от загрязняющих веществ, образующийся в процессе работы непрерывно действующего площадного источника в той же застройке.

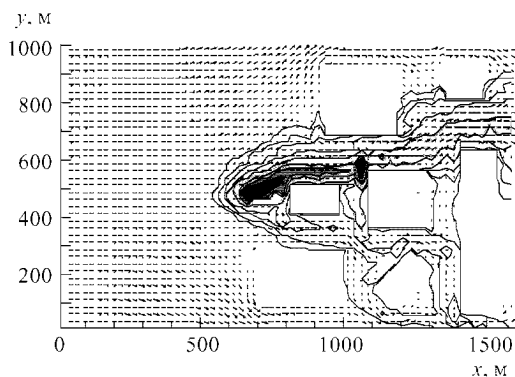


Рис. 20. Шлейф при работе непрерывно действующего точечного источника при преобладающем направлении ветра вдоль координатной оси X при наличии 7-ми домов различной формы и ориентации. Градации серого передают величину концентрации

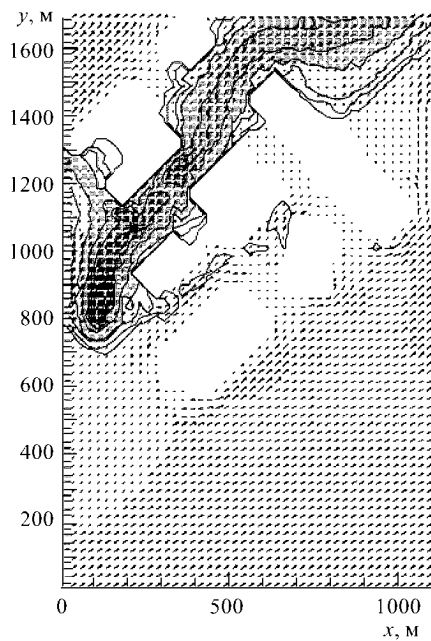


Рис. 21. Шлейф при работе непрерывно действующего точечного источника при наличии 7-ми домов различной формы и ориентации в случае поворота ветрового поля и рельефа, изображённых на рис. 20, на 45 градусов. Градации серого передают величину концентрации

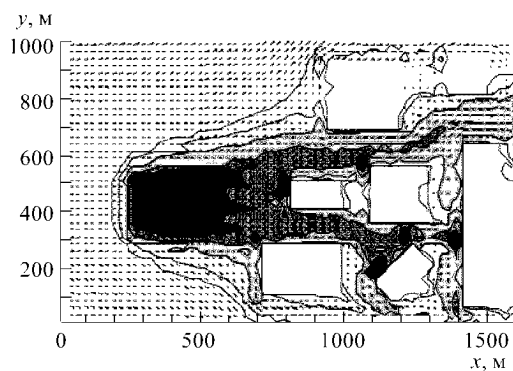


Рис. 22. Шлейф при работе непрерывно действующего площадного источника при наличии 7-ми домов различной формы и ориентации. Градации серого отражают величину концентрации

§ 4. Архитектура пакета

Архитектура пакета представляет собой блочную конструкцию.

В блоке начальных данных «regalign.dat» задаются типы вычислительных схем, в зависимости от которых изменяются методика работы модели и некоторые коэффициенты в исходном уравнении переноса. Варьируя их, можно проследить влияние вычислительных схем, размера шага сетки и т. д. на корректность получаемых результатов.

В блоке начальных данных «regfor.dat» задаются параметры сетки, некоторые коэффициенты в исходном уравнении переноса и граничные условия. Варьируя их, можно проследить влияние вычислительных схем на корректность получаемых результатов. Аналогично в блоке начальных данных «regmet.dat» задаются метеопараметры: скорость ветра, количество метеостанций, температурные градиенты и т. д.

В блоке начальных данных regdwini.dat задаются параметры, необходимые при использовании DWM-модели: критическое число Фруда и т. д.

В блоке начальных данных regurban.dat задаются параметры, необходимые при использовании URBAN-модели: координаты и высоты зданий.

В блоке начальных данных regtopo.dat задаётся массив высот рельефа, который может быть необходимым при использовании DWM- и URBAN-моделей.

В блоке начальных данных «regq.dat» задаются параметры источников.

В блоке начальных данных «reginit.dat» задаётся распределение концентрации в начальный момент времени.

Высоты рельефа отражаются массивами чисел размера double в середине ячеек. Порядок заполнения указан на рис. 23.

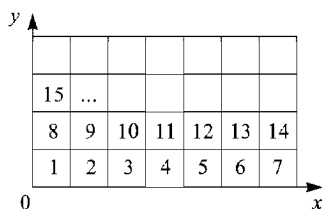


Рис. 23. Порядок заполнения массива в файле, описывающем рельеф

Обработка результатов из файлов производится любыми имеющимися средствами визуализации.

Последовательность действий при работе с пакетом:

1. Заполнить файлы данных «regfor.dat», «regmet.dat», «regurban.dat», «regalign.dat», «regq.dat», «reginit.dat».
2. Запустить poll.exe.

3. Обработать результаты — файлы «*.dat».

Модель TIMES имеет ограниченную область применимости по минимальному расстоянию от источника — 10 ячеек сетки, тогда как максимальное расстояние от источника — не ограничено.

§ 5. Исполнительные модули для интегрированной информационной системы

С целью повышения оперативности информирования высших органов власти о чрезвычайных ситуациях (ЧС) и их развитии была создана интегрированная информационная система моделирования. Она даёт возможность прогнозировать последствия ЧС и экологических бедствий. В эту систему вошли несколько исполнительных модулей на базе пакета «TIMES».

Моделирование для высокоскоростных оперативных расчётов осуществляется по аналитической формуле при упрощающих допущениях, приводимых выше в (2).

5.1. Программный комплекс «Атмосфера города». Программный комплекс «Атмосфера города» предназначен для прогноза распространения загрязнителей в атмосфере города и в масштабах производного региона на электронных географических картах при чрезвычайных ситуациях (ЧС). Он обеспечивает ввод исходных данных в расчётный модуль «TIMES» и расчёт и визуализацию полученных данных в виде виртуальной картины развития ситуации.

«Атмосфера города» создан средствами: Windows 95 OSR2 и выше; BDE версии 5.00 и выше, Borland C++ Builder 4 и Borland Delphi 5.

Для обеспечения функционирования Модели необходимы следующие технические средства:

- центральный процессор с тактовой частотой больше 233 МГц,
- объём оперативной памяти больше 24 Мб;
- объём дисковой памяти не менее 1 Гб.

Исполнительный модуль может функционировать в двух режимах: в работе общей системы и в автономном режиме.

В качестве исполнительного модуля системы модуль «Атмосфера города» обеспечивает принятие управления из подсистемы, формирует расчётное задание и осуществляет динамическую настройку параметров модели на входные данные.

Значения коэффициента горизонтальной диффузии вычисляются по формуле

$$K = \sigma_{\Theta}^2 \cdot \max(0,5, |w|) \cdot h,$$

где σ_{Θ} — угол горизонтальной флуктуации направления ветра в радианах, w — скорость ветра в м/с, H — высота слоя перемешивания в м, зависящая от атмосферной стабильности, α — некоторый эмпирический коэффициент ($\alpha = 0,01$). Оператор задаёт класс атмосферной стабильности в виде индекса, руководствуясь Таблицей 8.

Таблица 8. Классификатор атмосферной стабильности

Класс атмосферной стабильности	Соответствующий индекс в программе	Степень устойчивости	Погода	Высота слоя перемешивания H , м
A	1	Сильно неустойчивая	Очень солнечная летняя погода	2000 м
B	2	Средне неустойчивая	Солнечно и тепло	1500 м
C	3	Умеренно неустойчивая	Переменная облачность	1000 м
D	4	Нейтральная	Облачный день или ночь	750 м
E	5	Умеренно устойчивая	Переменная облачность ночью	300 м
F	6	Средне устойчивая	ясная ночь	250 м
G	7	Сильно устойчивая	ясная холодная ночь, ветер слабый	250 м

Значения коэффициента вертикальной турбулентности на произвольной высоте z вычисляются по следующим эмпирическим формулам:

$$K_z(z) = \begin{cases} \frac{z}{h} K_z(h), & z < h \\ K_z(h), & z \geq h, \end{cases}$$

где h — высота приземного слоя, $K_z(h)$ — значение коэффициента вертикальной турбулентности на некоторой высоте h . В табл. 9 приведены значения этих величин в зависимости от класса устойчивости атмосферы.

Таблица 9. Параметры для вычисления коэффициента вертикальной турбулентной диффузии

Класс атмосферной стабильности	A	B	C	D	E	F	G
K_z , м ² /с	160	100	70	15	5	1,5	0,13
h , м	100	100	100	150	250	250	250

Пользователь задаёт мощность точечного источника. Точечный источник всегда располагается в центре расчётной области. Пользователь задаёт величину скорости и азимут направления ветра. Азимут 0° соответствует направлению с юга на север, 90° — с запада на восток, 180° — с севера на юг, 270° — с востока на запад.

Информация о рельефе местности сообщается в двух одноимённых файлах с расширением «dbf» и «met».

Последовательность действий при работе с подсистемой:

1. Подготовить файлы с данными о рельефе.
2. Запустить программу.
3. Выбрать исполнительный модуль.
4. Выбрать в меню пункт «Вызов модели».
5. Задать входные данные в таблице в единицах, указанных в управляющем модуле. Имена файлов с рельефом и выходного файла задаются с полным путём, но без расширения.
6. Запустить на счёт нажатием кнопки «Расчёт». Ждать появления изображения на электронной географической карте.

В табл. 10 приведены входные параметры модели.

Таблица 10. Настройки входных параметров модели

Параметр	Единица измерения
Размер области по X	км
Размер области по Y	км
шаг сетки по X	км
шаг сетки по Y	км
Количество слоёв расчёта	шт
шаг по времени	ч
класс атм. стабильности	номер из табл. 8
Скорость ветра	м/с
направление ветра (азимут)	градусы
кол-во выброшенного вещества	тонны
газ	номер из табл. 11
точность расчёта	

Было проведено тестирование разработанного программного комплекса.

Исходные данные для расчёта:

размеры области — 90×90 км

шаг сетки — 2×2 км;

время наблюдения за распространением токсичного газа — 2 ч;

Изделие **ВОС** создано средствами: Windows 95 OSR2 и выше; BDE версии 5,00 и выше, Borland C++ Builder 4 и Borland Delphi 5.

Для обеспечения функционирования Модели необходимы следующие технические средства:

- центральный процессор с тактовой частотой больше 233 МГц, объём оперативной памяти больше 24 Мб;
- объём дисковой памяти не менее 1 Гб.

Модель **ВОС** может функционировать в двух режимах: в работе общей системы и в автономном режиме.

Программный комплекс предоставляет пользователям возможность:

- прогноза образования облака взрывоопасных и токсичных смесей веществ в атмосфере от мгновенных выбросов опасных промышленных объектов;
- получения прогностических карт образования в воздушной среде взрывоопасных и токсичных смесей веществ.

В соответствии со своим назначением модель реализует следующие функции:

- обеспечивает расчётные и игровые варианты моделирования;
- обеспечивает представление расчётной информации пользователям в графической форме.

В состав подсистемы моделирования входят:

- исполнительный программный модуль для расчёта зоны образования взрывоопасных смесей веществ с воздухом по данным о конкретных ЧС.
- база знаний с игровыми вариантами моделирования;
- блок формирования расчётного задания, позволяющий подготовить набор исходных данных для проведения расчётов на модели и обеспечивающий однозначное толкование запроса и выходного документа (меню);
- расчётный блок, обеспечивающий получение результатов моделирования;
- блок представления информации пользователю, обеспечивающий вид необходимой выходной информации.

Структура модели обеспечивает необходимые для реализации функциональных требований информационные и программные связи между её составляющими.

Информационное обеспечение системы включает:

- входную информацию;
- выходную информацию;
- средства ведения баз данных.

В табл. 11 приведены исходные данные для некоторых веществ, представляющих собой опасность образования таких смесей при распространении в атмосфере. Пределы взрываемости (нижний и верхний) выражены концентрациями в объёмн. % и в кг/м^3 паров растворителя в смеси с воздухом (ниже и выше этих пределов взрывчатые смеси обычно не образуются).

Таблица 11. Пределы взрываемости смесей паров органических растворителей, газов и некоторых паров с воздухом

Растворитель, газ, пар	Пределы взрываемости			
	объёмн. %		кг/м ³	
	низший	высший	низший	высший
Изо-амилацетат	2,2	10	0,119	0,541
Ацетон	2	13	0,0485	0,315
Бензин	1	7		
Бензол	1,4	9,5	0,45	0,308
н-Бутиловый спирт	3,7	10,2	0,114	0,314
Диоксан(1,4-)	2	22,2	0,0734	0,815
Дихлорэтан	6,2	15,9	0,256	0,680
Дихлорэтилен	9,7	12,8		
Ксилол	3,0	7,6	0,132	334
Метилацетат	2,2	15,6	0,127	0,431
Метиловый спирт	5,5	37	0,0734	0,493
Метилформиат	5	28,7		
Метилэтилкетон	1,8	9,5		
Пиридин	1,8	12,4		
Сероуглерод	1,0	50	0,0315	
Толуол	1,3	7	0,0498	0,268
Этилацетат	2,2	11,4	0,0824	0,403
Этиловый спирт	3,3	19	0,067	0,64
Этиловый эфир	1	40	0,365	1232
Этилформиат	3,5	6,4		
Аммиак	15,5	27		
Ацетальдегид	4	57		
Водяной пар	6	70		
Гексан	1,2	7		
Гептан	1	6		
Керосин	1	7,5		
Метил хлористый	8	20		
Пентан	1,4	7,8		
Полуводяной (генераторный) газ	20	75		
Пропиленовая окись	2	22		
Светильный газ	5	35		
Сероокись углерода	11,5	28,5		
Сероуглерод	1	50		
Толуол	1,3	7		
Циклогексан	1	9		
Этил бромистый	7	11		
Этил хлористый	3,5	14,8		
Этилена окись	3	80		

Нижние пределы взрываемости пылей некоторых веществ (c — концентрация твёрдого горючего вещества в мг на 1 л воздуха)

горючее вещество	c
Каменный уголь	17
Крахмал	7
Мука	10
Сахар	10
Сера	7

В Таблице 12 приведены параметры настроек входных данных в модель.

Таблица 12. Пример настроек входных параметров модели

Параметр	Поле входного файла	По умолчанию	Единица измерения
Размер области по X	PARAM01	1	км
Размер области по Y	PARAM02	1	км
Шаг сетки по X	PARAM03	0,025	км
Шаг сетки по Y	PARAM04	0,025	км
Количество слоёв расчёта	PARAM05	9	шт
Шаг по времени	PARAM07	0,05	ч
Класс атм. Стабильности	PARAM08	3	
Скорость ветра	PARAM09	5	м/с
Направление ветра (азимут)	PARAM10	90	градусы
Кол-во выброшенного вещества	PARAM11	50	тонны
Имя файла рельефа	PARAM12	relief	
Название вещества	PARAM13	toluol	
Точный/оперативный расчёт	PARAM14	1	

Входной информацией являются данные пользователей, введённые в процессе формирования расчётных заданий и данные из базы данных (см. таблицу настроек входных параметров модели).

Пользователь задаёт массу выброшенного взрывоопасного газа от точечного источника.

Точечный источник всегда располагается в центре расчётной области.

Пользователь задаёт величину скорости и азимут направления ветра.

Азимут 0° соответствует направлению с юга на север, 90° — с запада на восток, 180° — с севера на юг, 270° — с востока на запад.

Пользователь задаёт тип вещества (газ). Для каждого газа в табл. 13 задаются значения нижнего и верхнего предела концентраций образования взрывоопасных смесей. Таблица 13 содержится в информационном файле `pdk_mlx.dbf`.

Таблица 13. Структура файла с описанием веществ

Параметр	Поле входного файла	Единица измерения
Название газа	SUBSTANCE	
Нижний предел концентраций	PDK1	мг/м ³
Верхний предел концентраций	PDK2	мг/м ³

Информация о рельефе местности сообщается в 2-х одноимённых файлах с расширением «`dbf`» и «`met`», так же как и для модели «Атмосфера города».

Выходной информацией является визуализация проведённых модельных расчётов на электронной географической карте в виде изолиний нижнего и верхнего предела концентраций образования смесей.

Тестирование осуществлялось на основе проведения вычислительных экспериментов. На рис. 25 приведён один из результатов расчёта областей, где образуются взрывоопасные концентрации смесей. Показаны границы низших и высших взрывоопасных концентраций

5.3. Модель процессов распространения капельно-воздушной инфекции (модель КВИ). Модель представляет собой визуально-прогностический программный комплекс для исследования процессов распространения и зон заражения при действии источников инфекции. Капельно-воздушным путём распространяется инфекция таких болезней как грипп, ветрянка, корь, дифтерия и др.

«Модель КВИ» позволяет прогнозировать распределение больных людей, представляющих собой источник дальнейшего распространения инфекции по территории проживания, очерчивать зоны эпидемий во времени и пространстве средствами математического и компьютерного моделирования и выводить общее количество больных людей в этой зоне на данный момент времени.

«Модель КВИ» обеспечивает решение системы интегро-дифференциальных уравнений в частных производных, ввод исходных данных в расчётный модуль и вывод полученных результатов в виде виртуальной картины развития ситуации.

Решение системы уравнений «Модели» проводят с помощью численных методов на основе разностных аналогов дифференциальных уравнений для расчётной сетки.

«Модель» предоставляет пользователям возможность:

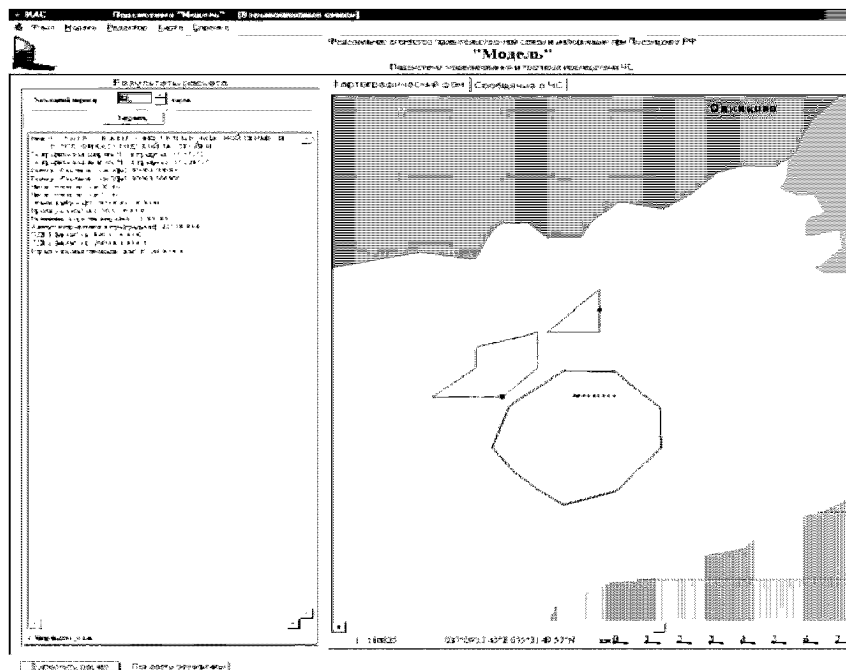


Рис. 25. Пример иллюстрации результатов расчёта по модели ВОС границы низших и высших взрывоопасных концентраций в системе ИАС

- прогноза распространения инфекций по территории проживания;
- получения прогностических карт распространения эпидемий.

В соответствии со своим назначением модель реализует следующие функции:

- обеспечивает расчётные и игровые варианты моделирования;
- обеспечивает представление расчётной информации пользователям в графической форме.

При известной эмиссии инфекции в центре моделируемой области с определённой плотностью населения, а также некоторых других известных медико-биологических величинах можно определить границы участков с эпидемической обстановкой во времени и пространстве и количество больных людей. Математическая модель в основном может быть представлена в следующем виде:

$$\begin{aligned} \frac{\partial}{\partial t} X &= \operatorname{div} (K \operatorname{grad} X) + rX(1 - b - m - n - p) + \operatorname{div} \mathbf{w}gX + \\ &\quad + X_0(x, y, t), \quad c(x, y, 2, t) \geq c_b, \\ X &= 0, \quad c(x, y, 2, t) < c_b, \end{aligned}$$

где $X = AN(x, y)c(x, y, 2m, t)$ — плотность больных в зонах инфицирования; $N(x, y)$ — плотность населения; $c(x, y, z, t)$ — концентрация инфицирующих бактерий в воздухе, причём, если $c(x, y, 2m, t) = c_b$ (граничное значение инфицирующей концентрации), то в точке $(x, y, 2m, t)$ начинается заболевание среднего человека (появляется некоторая плотность средних больных), если $c(x, y, 2m, t) < c_b$, то в точке $(x, y, 2m, t)$ заболевания нет; A — медико-биологический когнитивный коэффициент, характеризующий специфику данного вида инфекции; K — некоторый когнитивный коэффициент, зависящий от принятых допущений при рассмотрении процессов (в частности, это может быть коэффициент горизонтальной диффузии инфекции, если описывается процесс распространения инфекции в турбулентном поле атмосферы) и характеризующий скорость изменения градиента плотности больных; r — некоторый когнитивный медико-биологический коэффициент, характеризующий интенсивность эволюционного нарастания или убывания (наличие источников или стоков) концентрации инфекции от больных людей; $b(x, y, t)$, $m(x, y, t)$ и $n(x, y, t)$ — доли X , изменяющие на данный момент в точке (x, y, t) плотность больных людей: $b(x, y, t)$ — доля X людей, у которых инкубационный период развития болезни (они ещё не заражают), $m(x, y, t)$ — доля X выздоравливающих людей благодаря собственным защитным силам организма, $n(x, y, t)$ — доля X людей, ускоренно выздоравливающих благодаря принятым мерам борьбы с эпидемией, включая изоляцию больных, p — доля X , уменьшающая число новых больных и концентрацию вирусов из-за гибели инфекции в результате воздействия климатических условий (мороз, жара и т. п.) и конечной жизни вирусов; $\mathbf{w}$ — усреднённая скорость адвективного перемещения больных людей в определённом направлении или скорость перемещения инфекции в поле ветра на открытом воздухе в толпе людей; $X_0(x_0, y_0, 2m, t_0)$ — начальное значение плотности больных в точке $(x_0, y_0, 2m)$ с начальной концентрацией c_0 :

$$X_0(x, y, 2m, t) = \begin{cases} X_0 \leq N(x_0, y_0); & t = 0, x = x_0, y = y_0; \\ 0, & t \neq 0 \text{ or } x \neq x_0 \text{ or } y \neq y_0. \end{cases}$$

Если $X(x, y, t) \geq N(x, y)$, то полагается значение $X(x, y, t) = N(x, y)$. Результат счёта может быть представлен как изолинии, соответствующие 20% заболевших от общего числа людей, или как число заболевших в эпидемической зоне.

Диффузия инфекции в пространстве может моделироваться по-разному в зависимости от рассматриваемых областей. Если рассматриваются процессы распространения на открытом воздухе, то инфицирование можно моделировать как распространение вирусов в поле ветра. Тогда значения коэффициента горизонтальной диффузии вычисляются по формуле $K = \sigma_\Theta^2 \max(0, 5, |\mathbf{w}|)h$, где σ_Θ — угол горизонтальной флуктуации направления ветра в радианах, $\mathbf{w}$ — скорость ветра в м/с, H — высота слоя перемешивания в м, зависящая от атмосферной

стабильности, α — некоторый эмпирический коэффициент ($\alpha = 0,01$). Оператор задаёт класс атмосферной стабильности в виде индекса, руководствуясь Таблицей 7.

Если рассматриваемая область имеет сложную конфигурацию (загрязнение происходит в основном при близком контакте в зданиях и транспорте), тогда необходимо подбирать некоторый усреднённый коэффициент диффузии, характеризующий скорость распространения инфекции, например, в городских условиях.

Функционирование модели КВИ в системе в качестве исполнительного модуля обеспечивается специальным программным комплексом. Создана специальная техническая документация, позволяющая операторам производить запуск данного исполнительного модуля на счёт и выводить на экран дисплея виртуальную картину распространения и количественные характеристики реальных процессов.

Для обеспечения функционирования Модели необходима системная среда: Windows 95 OSR2 и выше; BDE версии 5,00 и выше.

Модель КВИ создана средствами Borland C++ Builder 4 и Borland Delphi 5. Она может функционировать как в работе общей системы, так и в автономном режиме.

В качестве исполнительного модуля системы модель обеспечивает:

- принятие управления из подсистемы;
- формирование расчётного задания;
- динамическую настройку параметров модели на входные данные.

Модель предоставляет пользователям возможность:

- прогноза распространения инфекции в пространстве от мгновенных источников инфекции;
- получения прогностических карт образования эпидемической обстановки при задании соответствующих параметров и прогноза необходимых мер борьбы для ликвидации болезни;
- подсчёт числа заболевших людей в этих зонах.

Целью разработки является оказание информационной поддержки пользователям при подготовке чрезвычайных мероприятий по ликвидации эпидемий, а также при проведении профилактических мер по снижению заболеваемости.

В соответствии со своим назначением модель реализует следующие функции:

- обеспечивает расчётные и игровые варианты моделирования;
- обеспечивает представление расчётной информации пользователям в графической форме.

Результатом счёта являются очерченные изолиниями области, где плотность больных соответствует некоторым условиям, включая эпидемии, а также количество больных людей в этих областях.

Структура программного обеспечения позволяет реализовать функциональные требования, информационные и программные связи между его составляющими.

Обеспечение системы включает:

- входную информацию;
- выходную информацию;
- средства численных расчётов.

Соответствующие поля параметров приведены в Таблице 14.

Таблица 14. Поля входных параметров

Параметр	Единица измерения
Размер области по X	км
Размер области по Y	км
Шаг сетки по X	км
Шаг сетки по Y	км
Шаг по времени	ч
Количество слоёв расчёта	шт
Класс атмосферной стабильности (I–V)	
Скорость ветра	м/с
Направление ветра	градусы
Скорость адвективного перемещения	м/с
Направление адвективного перемещения	градусы
Имя файла с плотностью населения	
Специфика инфекции	
Точный/оперативный расчёт	
Концентрация бактерий в начальный момент времени в очаге инфекции	шт/м ³
Коэффициент диффузии	м ² /с

Параметр «Специфика распространения» представляет собой когнитивный коэффициент и может быть задан непосредственно в поле значений. При наличии вместо числового значения в поле символа «х» данный параметр вычисляется в соответствующем файле по введённым значениям коэффициентов r, b, m, n, p .

При нулевом значении в поле коэффициента диффузии предполагается модель распространения инфекции на открытом воздухе. При этом коэффициент диффузии соответствует коэффициентам горизонтальной и вертикальной турбулентной диффузии в атмосфере, которые вычисляются и зависят от класса атмосферной стабильности.

При ненулевом значении коэффициента диффузии в поле скорости адвективного перемещения задаётся соответствующая величина, а скорость ветра полагается равной нулю. При нулевом значении в поле скорости адвективного перемещения используется значение скорости ветра, задаваемое в соответствующем поле, и предполагается распространение инфекции в горизонтальном поле ветра.

Таблица 15. Настройки входных параметров модели

Параметр	Поле входного файла	По умолчанию	Единица измерения
Размер области по X	PARAM01	1	км
Размер области по Y	PARAM02	1	км
Шаг сетки по X	PARAM03	0,025	км
Шаг сетки по Y	PARAM04	0,025	км
Количество слоёв расчёта	PARAM05	1	шт
Шаг по времени	PARAM07	0,05	ч
Класс атмосферной стабильности (I–V)	PARAM08	3	
Скорость ветра	PARAM09	5	м/с
Направление ветра	PARAM10	90	градусы
Концентрация бактерий в очаге инфекции	PARAM11	1 000 000	шт/м ³
Имя файла с распределением плотности населения	PARAM12	regpopul	
Специфика инфекции	PARAM13	influenza	
Точный/оперативный расчёт	PARAM14	1	
Коэффициент диффузии	PARAM15	0	м ² /с
Скорость адвективного перемещения	PARAM16	0	м/с
Направление адвективного перемещения	PARAM17	90	градусы

Вызов и загрузка осуществляется последовательными действиями оператора:

- 1) подготовка файлов с данными о распределении плотности населения;

- 2) из развёртываемого списка выбор соответствующего исполнительного модуля с названием «КВИ»;
- 3) выбрать в меню «Файл» пункт «Рабочая карта» и указать файл с плотностью населения на карте местности;
- 4) выбрать в меню «Файл» пункт «Настройка модели»;
- 5) в панели «Модель» заполнить соответствующие поля (в поле «Модуль» указать полный путь к Модели (файл infes.exe));
- 6) в панели «Параметры» заполнить соответствующие поля согласно Таблице 15;
- 7) выбрать в меню пункт «Вызов модели»;
- 8) задать входные данные в таблицах в единицах, указанных в управляющем модуле; имена файлов с плотностью и выходного файла задаются с полным путём, но без расширения;
- 9) запустить на счёт нажатием кнопки «Выполнить расчёт»; ждать появления изображения на электронной географической карте (показывается распределение % больных от общей плотности населения, измеряемой в чел./км²) и отчётной информации в окне. В случае необходимости нажать кнопку «Показать результаты».

Таблица 16. Структура файла со спецификой инфекции

Параметр	Поле входного файла	Единица измерения
Название инфекции	SUBSTANCE	
Коэффициент r	PARAMR	1/с
Коэффициент b	PARAMB	
Коэффициент m	PARAMM	
Коэффициент n	PARAMN	
Коэффициент p	PARAMP	
Значение выражения $r * (1 - b - m - n - p)$	PARAM_SUM	1/с
Доля заболевшего населения, при которой возникает эпидемия	PDK	
Значение параметра c_b	C_B	шт/ м ³
Коэффициент A	A	м ³ /шт
Минимальная плотность населения, для которой существует понятие эпидемии	POPUL_MIN	чел/км ²

Входной информацией являются данные пользователей, введённые в процессе формирования расчётных заданий и данные из базы дан-

информация картографических объектов. База данных должна иметь два обязательных поля и может иметь неограниченное количество полей атрибутивной информации. Второй файл имеет то же имя, что и D-base, но с расширением «**met**». Этот файл содержит координатную и топологическую информацию объектов.

Результаты тестирования приведены на рис. 26. Линия очерчивает область, где создавалась эпидемическая обстановка.

Список литературы

1. Huber A.H., Snyder W.H. «Building Wake effects on Short Stack Effluents», Preprint Volume for Triad Symposium Atmospheric Diffusion and Air Quality. American Meteorological Society, Boston, MA, 1976.
2. Hertel O., Berkowicz R., Larssen S. The operational street pollution model, «Air Pollut. Models and its Appl VIII: Proc 18th NATO / CCMS Int. Techn. Meet. Air Pollut. Models and its Appl. Vancouver. May 13–17 1990, N. Y. — London, 1991. P. 741–750.
3. Kamenetsky E., Viern N. «Model of the flow and air pollution concentration in urban canyons», «Boundary Layer Meteorol». 1995, V. 73, 1–2. P. 203.
4. Johson G., Hanter L. «A numerical study of dispersion passive scalars in city canyons», Boundary Layer Meteorol. 1995. V. 75. 3. P. 235–262.
5. «A Review of Development and Application of the Urban Airshed Model», Sheffe R.D., Morris R.E. Atmospheric Environment V. 278, № 1. P. 23–39, 1993.
6. Фатеев А.И., Клименко Е.Т. Моделирование концентрационного поля выбросов от автотранспорта в глухих уличных каньонах. Стационарная задача. «Новая технология в газовой промышленности // Конференция молодых учёных по проблемам газовой промышленности РФ. М., сент. 1995 г. Тезисы докладов. 1995. С. 288–289.
7. Murrey D., Rurmaster D. «Residential air exchange rates in the USA empirical and estimated parametric distributions by season and climatic region», Risk. Anal. 1995 V. 15. 4. P. 459–465.
8. Roth M., Oke T. «Сравнительная эффективность турбулентного переноса тепла, массы и количества движения над городской застройкой», J. Atmos. Sci. 1995. V. 52. 11. P. 1863–1874.
9. Hoydich W.G., Dabberdt W.F. «A fluid modeling study of concentration distributions at urban intersection», Sci / Total. Environ. 1994, 146–147. P. 425–432.
10. Колдоба А.В., Повещенко Ю.А., Самарская А.А., Тишкин В.Ф. Методы математического моделирования окружающей среды. — М.: Наука, 2000. — 256 с.
11. Ключкова Л.В., Повещенко Ю.А., Самарская Е.А. и др. Компьютерные технологии для наполнения информационных систем поддержки принятия решений о техногенном загрязнении различных сред. Труды II Всероссийского симпозиума «Математическое моделирование и компьютерные технологии». г. Kislovodsk (23–25 апреля 1998 г.). Т. 2. Изд-во КИЭП, 1998.

12. Белов И.В., Беспалов М.С., Ключкова Л.В. и др. Сравнительный анализ некоторых математических моделей для процессов распространения загрязнений в атмосфере. Журнал «Математическое моделирование». Т. 11, № 8. 1999. — 13 с.
13. Белов И.В., Беспалов М.С., Ключкова Л.В. и др. Транспортная модель процессов распространения газообразных примесей в атмосфере города. Журнал «Математическое моделирование». Т. 12, № 5. 2000. — 8 с.
14. Ключкова Л.В., Сузан Д.В., Тишкин В.Ф. Математическое моделирование для наполнения программных визуально-прогностических систем оценки качества воздуха городов и промышленных центров. Сборник Тезисов докладов на международной конференции «ENVIROMIS 2002», г. Томск, Изд-во ЦНТИ.
15. Беспалов М.С., Ключкова Л.В., Сузан Д.В., Тишкин В.Ф. Математическое моделирование процессов распространения загрязнений в ветровом поле атмосферы из воздушных коридоров траекторий движения летательных аппаратов. Сборник Тезисов докладов на 30-й школы-семинара «Математическое моделирование в проблемах рационального природопользования» и 10-го семинара «Экология. Экономика. Экспертиза. Информатика», п. Дюрсо, 2002.
16. Беспалов М.С., Ключкова Л.В., Сузан Д.В., Тишкин В.Ф. Распространение загрязняющих веществ из вертикального следа от выхлопных газов в различных слоях атмосферы при взлёте ракет. Сборник Тезисов докладов на 7-й международной конференции «Системный анализ и управление космическими комплексами», Украина, 2002 г.
17. Ключкова Л.В., Кулешов А.А., Сузан Д.В., Тишкин В.Ф. Математическое моделирование чрезвычайных ситуаций и промышленная безопасность. Проблемы прогнозирования чрезвычайных ситуаций и их источников; Научно-практическая конференция 26–27 июня 2001 г.; Доклады и выступления. Москва, Центр «Антистихия», 2002 г.
18. Израэль Ю.А., Назаров И.М., Прессман А.Я. и др. Кислотные дожди. — Л.: Гидрометеоиздат, 1983. — 206 с.

Глава II

МОДЕЛИ ВЫБОРА И ИХ ПРИМЕНЕНИЕ К АНАЛИЗУ СОЦИАЛЬНЫХ КОНФЛИКТОВ

В данной главе построены и исследованы математические модели поведения социальной общности в ситуации выбора ею той или иной позиции. Модели основаны на понятии социально-политического потенциала типичной личности, совершающей свой выбор на основе подражания в соответствии с законами социальной психологии. В математическом отношении модели представляют собой систему нелинейных обыкновенных дифференциальных уравнений, описывающую разнообразные временные режимы самоидентификации социальной общности, в том числе циклическую самоидентификацию. Приведены примеры использования моделей для анализа и прогноза этнополитической ситуации в Дагестане.

§ 1. Некоторые особенности моделирования социальных процессов

Методология математического моделирования завоевала прочные позиции в технологической и естественно-научной сферах, её прогресс существенно заметен также и в применениях к экономическим системам. Если же говорить о процессах с участием человеческого фактора (в первую очередь о социальных процессах), то успехи «третьего метода» познания, конструирования, проектирования в этой области гораздо скромнее. Существует немало глубоких и принципиально неустранимых причин такого отставания:

1. *Отсутствие в социологии фундаментальных законов*, т.е. связей между основными социальными переменными, допускающих точное количественное выражение. Впрочем, вспомним историю такого общеизвестного и не вызывающего ни у кого сомнения закона природы как сохранения энергии. Первый шаг в его официальном признании был сделан в конце 18-го века Французской Академией Наук, которая приняла своё знаменитое решение не рассматривать далее многочисленные проекты вечных двигателей ввиду их полной научной несостоятельности. Лишь столетие спустя этот закон обрёл, наконец, строгую математическую формулировку. Может быть, придёт время, когда соответствующие учреждения не станут всерьёз даже обсуждать

социальные проекты, противоречащие (пока ещё неизвестным) законам социологии.

2. *Неустраняемая неточность «измерений».* Даже в тех случаях, когда существует вполне удовлетворительная шкала для измерения (например, электоральных предпочтений) к результатам следует относиться с гораздо большей осторожностью, чем к результатам измерений в механике, астрономии, физике, химии. Социальные субъекты не слишком склонны к открытости, их мнения и поведение могут быстро изменяться, требуются немалые интеллектуальные ухищрения для приближения к истине с приемлемой точностью. Это удаётся далеко не всегда и не всюду.

3. *Разномасштабность, разнородность и нелинейность* изучаемых социологией явления и процессов. Например, рассматриваемые социальные общности могут насчитывать от нескольких человек до нескольких сотен миллионов человек (разброс в восемь порядков!), а наличие нелинейных и прямых обратных связей — неотъемлемая характеристика любого социума.

4. *Постоянное усложнение социальных объектов и их «рефлексивность»*, означающая, в частности, что новые знания рано или поздно сами становятся частью объекта, изменяя его характеристики и свойства. Примером служит судьба многих социальных прогнозов и проектов — даже будучи обоснованными по всем существующим меркам, они, вплетаясь в социальную ткань, становясь социальной реальностью, никогда не осуществляются в той степени, которая отвечала бы стандартам точных наук. Отсюда — самореализующиеся и (гораздо чаще) саморазрушающиеся социальные прогнозы и, как одно из следствий — бесконечная «гонка» математической модели за изучаемым объектом.

Поэтому к математическому моделированию социальных процессов необходимо предъявлять дополнительные «повышенные» методологические требования. Их соблюдение позволит сузить простор для интуитивного умозрительного «моделирования», расширить поле приложений рациональных методов. Моделирование присутствует почти во всех видах творческой активности людей различных специальностей — предпринимателей и военачальников, политиков и управленцев. Без привнесения в эти сферы точного знания невозможно рассчитывать на получение столь нужной нам высокотехнологичной, конкурентоспособной и разнообразной продукции. К общепризнанным [1] профессиональным требованиям математического моделирования относятся, например:

1. Чёткая формулировка основных понятий и предположений, апостериорный анализ адекватности используемых моделей, гарантированная точность вычислительных алгоритмов и компьютерных программ.

2. Аккуратное разграничивание математических и житейских терминов, звучащих одинаково, но имеющих, зачастую, разный смысл — хаос, порядок, бифуркация и т. д. Показательна в этом отношении пара

«хаос–порядок». В синергетике хаос означает полное равенство всех элементов рассматриваемой системы, а «порядок» — их упорядоченное «неравенство». В юриспруденции «равенство перед законом», наоборот, отвечает порядку, а неравенство — правовому хаосу.

3. Осторожное применение уже готового (и модного) математического аппарата к изучению реальных объектов, следование пути «от объекта к модели», но не наоборот. К примеру, нередки попытки чисто формальной трактовки ряда явлений с позиций широко известной теории катастроф, весьма повредившие репутации этого изящного и законченного раздела математики.

Эти и другие стандартные требования необходимо в случае анализа социальных процессов дополнить, по меньшей мере, следующими:

1. Не покушаться на твёрдо установленные факты и закономерности, избегать соблазна научных «революций», прикрытых математическим флёром. По мнению некоторых математиков, человечеству не следует особенно хлопотать в связи с наступлением третьего тысячелетия — мы ещё и первого не закончили (и в обсуждении этой темы участвуют немалые научные силы).

2. Не подменять собой предметников — социологов, философов, политологов (или, во всяком случае, ясно очерчивать границу между применяемой исследователем математической методологией и его собственными воззрениями на изучаемый предмет). Например, этнологи и социологи вряд ли признают существование такого понятия как *пассионарность* этноса (при всём уважении к автору этого термина) — его происхождение связывается с действием веземных сил, однако математические модели эволюции пассионарности уже строятся и изучаются.

3. По возможности придерживаться здравого позитивизма, трезво оценивая, какие из социальных процессов можно моделировать с достаточной степенью достоверности, а для каких ещё преждевременно говорить об исследовании с необходимой полнотой. Исследования в этой области можно условно подразделить на фундаментально-ориентированные (направленные на выяснение общих свойств объектов и, в то же время, связанные с крупной реальной проблемой) и конкретно-прикладные (опирающиеся на полученные ранее фундаментальные знания и дающие количественные характеристики).

В коллективах ИММ РАН, ГУИС ФАПСИ и соцфака МГУ ведутся работы как первого (изучение систем «Власть–Общество» и близких вопросов — краткая характеристика этих исследований содержится в данном параграфе), так и второго направлений (основное содержание данной главы). Причём, в этой деятельности исполнители работ стремятся, по мере возможностей, следовать сформулированным выше требованиям.

Почему так важно изучение именно вопросов функционирования нашей государственной машины, взаимодействующей с нарождающимся у нас гражданским обществом?

Потому что исторические задачи, стоящие перед новой российской государственностью, глубоки и долговременны:

1. Определение допустимых степеней «разгосударствления» жизни нации, вмешательства (невмешательства) власти в экономическую деятельность, в культуру, науку и другие сферы общественной активности.

2. Выбор разумных пределов децентрализации (централизации) властных полномочий в Федерации, оптимальных способов их раздела между центром и субъектами, субъектами и органами самоуправления.

3. Преодоление внутренних проблем властных структур — ограниченности средств и ресурсов, неинформированности, некомпетентности, коррумпированности.

Эти и другие не перечисленные здесь, но хорошо известные, задачи должны решаться в условиях непридуманной «русской специфики». К ней относятся, например, вековой (тысячелетний) опыт предыдущей российской государственности, недостаточно ещё широко известный, по-разному толкуемый и весьма противоречивый. На одной чаше весов — почти двухсотлетней давности, вывод Н. М. Карамзина о том, что только «мудрая, долговременная опытность научает людей благодетельному разделению властей...» т.е. мысль, достойная любого современного ему политического философа. А на другой — хотя и вполне объяснимая, но оттого ничуть не менее прискорбная массовая привычка обращаться «в Центр» почти по каждому поводу. Никуда не уйти и от наших геополитических реалий: евразийности, соседства с разнообразными цивилизациями, пространственной асимметрии, неравномерной заселённости, огромных расстояний. Беспрецедентен и факт появления, в одночасье, самой большой в мире русской «диаспоры».

Ко всем этим сложностям, сопровождающим работу нашей государственной машины, добавляются такие неотъемлемые требования, как сохранение не декларативной, а истинной целостности нации, заметное улучшение условий её существования, упрочение международного положения России, прежде всего, через осознание всеми её нынешней и грядущей роли в мире.

Приходит пора неизбежного обращения к более рациональным, более точным методам строительства российской государственности, разработке и анализу, если угодно, научных моделей этого сложнейшего процесса. Постепенное ограничение интуитивных подходов и импровизаций, политического мистицизма и туманных формулировок — насущная необходимость, ибо цена крупных ошибок в этой сфере давно стала неприемлемо высокой для сегодняшней России. Процесс демистификации власти уже начался, о чём говорит впечатляющий успех так называемых избирательных технологий, широко применявшихся в последние годы на общенациональных и региональных выборах. Данные технологии отнюдь не сводятся ко всеобщему «промыванию мозгов», результаты которого могут, кстати говоря, быть противопо-

ложны ожидаемым, а включают в себя весьма тонкие приёмы количественного анализа и даже элементы математического моделирования (имитации) поведения избирателей и политических соперников. В этом отношении мы, по-видимому, приближаемся к мировым стандартам, и претендующие на власть силы просто обязаны считаться с новыми подходами.

Однако как избирательные технологии, так и во многом близкие или связанные с ними методы социополитического мониторинга предназначены, в основном, для анализа политической конъюнктуры, для сопровождения текущей борьбы за власть. Изучение более глубокого пласта проблем государственного строительства, относящихся к фундаментальным вопросам политологии, требует создания более глубоких моделей, существенно использующих достижения точных наук. Нельзя сказать, что общая политология чужда математическому языку — многие политологические учения и теории оперируют с такими понятиями как «обратная связь», «потoki информации» и т. д., но только как с образами, служащими для краткой иллюстрации тех или иных положений.

В то же время наш пока ещё первоначальный опыт свидетельствует — при относительно небольшом числе допущений возможно построение весьма содержательных и поддающихся исследованию математических моделей, описывающих некоторые ключевые воздействия в системе «Государственная власть — Гражданское общество», а результаты моделирования допускают достаточно ясную интерпретацию в политологических терминах. Тем самым, методология математического моделирования может послужить более глубокому пониманию систем «Власть–Общество», многих важных вопросов «технологии» власти.

Рассматривая проблему шире, следовало бы иметь в виду, что традиционно сильная отечественная школа математического моделирования всегда отвечала мировому уровню (о некоторых её достижениях см., например [1]). Не будет большим риском утверждать: многим профессионалам точных наук по силам, не подменяя историков, социологов, политологов, внести заметную лепту в междисциплинарное дело разработки научных моделей становления и функционирования российской государственности. Вручив, тем самым, нашей власти дополнительный неполитизированный и неидеологизированный и потому универсальный инструмент для углублённого анализа и долгосрочного прогноза. Способствование власти, а значит, и всему обществу, «умом понять» Россию вовсе не противоречит желанию в неё «только верить». Скорее наоборот, выдвигаемый нашими выдающимися соотечественниками нравственный императив «жить не по лжи», может быть поддержан точными подходами, снимающими с власти излишний покров таинственности и непредсказуемости.

Перечислим некоторые направления исследований, развиваемые в рамках данной тематики:

1. Основные свойства математических моделей системы «Государственная власть — Гражданское общество» (проверка их на разумность и «адекватность», сопоставление результатов с реалиями традиционных демократий).

2. Катастрофические эволюции распределения власти в иерархиях и способы их парирования (тоталитарные и анархические «ловушки», сравнение с некоторыми крупномасштабными событиями из истории России и Германии XX-го века).

3. Коррупция в иерархиях и эффективные стратегии их ограничения.

4. Простейшие модели равновесия ветвей власти.

5. Другие актуальные проблемы (оптимальная конструкция властного дерева, вопросы «властометрии», децентрализация полномочий и федеральное устройство, последствия разгосударствления).

Результаты этих работ достаточно подробно опубликованы (см., например, [1–4]). Поэтому ограничимся лишь краткой иллюстрацией выводов, следующих из моделирования коррумпированных иерархий.

Коррупция давно стала огромной российской проблемой. Впрочем, не только российской, а и общемировой. На одном из недавних экономических форумов в Давосе работала секция с характерным названием «Взяктономика» (так что наша специфика заключается, прежде всего, в масштабах этого явления). Существуют даже всерьёз обсуждаемые точки зрения о полезности коррупции как своеобразной смазки колёс государственного аппарата. Хорошо известны апробированные способы и меры экономического и законодательного характера для если не полного истребления этой болезни (что, наверное, невозможно), то, по крайней мере, её ограничения. Думается, что анализ моделей властных структур мог бы помочь отысканию дополнительных, более тонких, инструментов подавления коррупции.

Задача формулируется примерно следующим образом. Пусть в единицу времени система «Власть–Общество» выделяет некий ресурс (часть бюджета) для борьбы со скрытым оплачиваемым использованием государственной власти в интересах частных групп (коррупций). Как распорядиться им более эффективно, т.е. так, чтобы при фиксированном антикоррупционном ресурсе (он, очевидно, не может быть сколь угодно большим) употребить деньги налогоплательщиков наилучшим образом и достичь максимального предотвращения ущерба от коррупции? В модели идеальных законопослушных структур власти вводится обусловленная коррупцией «порча», формализуются такие понятия, как степень коррумпированности, ущерб от коррупции, эффективность её подавления и т.д.

Вычислительные эксперименты с моделями показывают: из-за сложной топологии иерархий и их неоднородности существуют более и менее эффективные стратегии подавления коррупции (при выборе плохой стратегии эффект может быть и отрицательным — затраты превышают «прибыль»). Следовательно, появляется возможность

проводить антикоррупционные мероприятия более осмысленно, обращая основное внимание на наиболее чувствительные к ним участки заражённой иерархии (в данном случае имеются ввиду полицейские меры — поиск, осуждение и наказание коррупционеров, защита свидетелей и т.п.). Вытекающее из модели представление о существовании критических уровней в иерархии, через которые прокачиваются наибольшие суммарные потоки властных распоряжений (и это необязательно высшие звенья), даёт один из ключей для проведения кадровой политики — желательно правильно находить эти уровни, предвидеть их возможное смещение в иерархии и обеспечить, в первую очередь, именно эти уровни достаточно честными и компетентными чиновниками.

Подобный анализ даёт новые аргументы в пользу того, что различные несовершенства властных иерархий нельзя преодолеть наскоками, кампанией. Взаимодействия в системе «власть — общество», формируемое ими реальное распределение власти и «потоки» власти в иерархии зависят от поведенческих характеристик партнёров, разных в разных частях иерархии, от строения властной структуры, которые в свою очередь изменяются со временем и т.д. Конечно же, иерархия так или иначе приспосабливается к новым условиям, но во многом стихийно. В системе появляются излишние напряжения, которые можно было бы смягчить, принимая во внимание результаты моделирования тех или иных изменений. С моделями, в отличие от реальных объектов, дозволено проводить любые эксперименты (для более точного анализа необходимо, чтобы эти модели были оснащены, т.е. дополнены сведениями из пока ещё не существующей российской «властометрии»).

Сказанное в полной мере относится также к вопросам федерализма, разгосударствления и другим фундаментальным проблемам нашей государственности. Полезны даже сами постановки задач, следующие из изучения моделей системы «Власть—Общество». Например, федерализм тесно связан с децентрализацией власти, с наилучшими способами раздела полномочий между Центром и субъектами федерации. Эти способы и последствия их реализации могут предварительно изучаться по моделям. Кроме того, из моделей нетрудно усмотреть, что у федеральных властных структур имеется возможность, оставаясь в рамках уже утверждённых общих договорённостей, вести весьма тонкую «игру» с субъектами и достигать своих целей (и наоборот, такая же игра вполне доступна субъектам). Её сущность заключается в выборе подходящих пропорций между приказами «по команде» и приказами «через голову», причём расходуемый властный ресурс остаётся неизменным. Российская специфика такого рода деятельности должна, по-видимому, состоять в учёте пространственной и правовой асимметрии федерации, и, конечно же, в учёте устоявшейся в регионах психологии взаимоотношений с Центром.

Если же говорить о процессах разгосударствления (национализации), т.е. об уменьшении (увеличении) размера непосредственно кон-

тролируемой властью части национального пирога, то даже поверхностный анализ моделей наводит на следующие вопросы: надо ли при этом сокращать (увеличивать) штаты иерархии, а если да, то как — методом равномерной «прополки» или исключением целых участков властных структур, как изменится мера «ответственности» и условия работы остающихся (прибывающих) чиновников, в чём можно положиться на зарубежный опыт, а где нельзя игнорировать особенности родных бюрократов и граждан?

Отечественная школа математического моделирования имеет многолетние традиции, разнообразный потенциал и при соответствующей поддержке (организованной, финансовой, кадровой) способна профессионально послужить более глубокому пониманию важнейших социальных процессов, протекающих в российском обществе.

§ 2. Постановка проблемы, основные понятия и предположения

2.1. Ситуация выбора «электоратом» политической линии, участвующие стороны, основные величины, цели исследования. Рассматривается динамический, т.е. развивающийся во времени процесс выбора некоторой социальной общностью (этнической, конфессиональной, профессиональной, ...) позиции (линии поведения, стратегии, ...) по отношению к значимому для неё вопросу (политическому, экономическому, этическому, ...). В дальнейшем для определённости будем говорить о политическом выборе и ради краткости наряду со словами «социальная общность» будем употреблять слово «электорат».

Число дееспособных членов электората $N(t) > 0$ зависит, вообще говоря, от времени в силу как естественных причин (например, вступление новых членов в дееспособный возраст или выбытие из него), так и причин внешнего характера (например, благодаря наличию эмиграционных или иммиграционных потоков). По отношению к выбираемой позиции электорат состоит, очевидно, из трёх основных категорий

$$N(t) = Z(t) + X(t) + Y(t), \quad t \geq t_0, \quad (1)$$

где в балансном соотношении (1) через $Z(t) \geq 0$ обозначено число неучастников процесса (нейтралы), а через $X(t) \geq 0$ и $Y(t) \geq 0$ — число сторонников позиции «X» и «Y» соответственно (имеется в виду, что позиции «X» и «Y» взаимно исключают друг друга).

Численность категорий Z, X, Y в общем случае изменяется со временем не только из-за естественных и внешних факторов, но и благодаря внутренним причинам — переходам членов социальной общности из одной категории в другую.

Более детально структура числа неучастников, сторонников и противников описывается соотношениями ¹⁾

$$Z(t) = Z_0(t) + z(t), \quad X(t) = X_0(t) + x(t), \quad Y(t) = Y_0(t) + y(t), \quad (2)$$

где в балансах (2) через $Z_0 \geq 0$, $X_0 \geq 0$, $Y_0 \geq 0$, обозначено текущее число твёрдых нейтралов, сторонников и противников, не изменяющих свою позицию ни при каких обстоятельствах (во всяком случае на рассматриваемых промежутках времени), а через $z(t)$, $x(t)$, $y(t)$ — текущее число членов электората, способных относительно быстро (в сравнении с твёрдой его частью) либо занять определённую позицию ($z(t)$), либо стать нейтральными или перейти в противоположный лагерь ($x(t)$ и $y(t)$). Величина $n(t) = z(t) + x(t) + y(t)$ представляет собой, таким образом, численность подверженного колебаниям электорального «болота». Очевидно, что, если в какой-то момент времени $x(t) = 0$ ($y(t) = 0$), то $X(t) = X_0(t)$ ($Y(t) = Y_0(t)$) и текущее число сторонников позиции «X» («Y») — минимально возможное.

Сторонники и противники той или иной позиции отличаются от неучастников процесса тем, что совершают определённые действия в её поддержку — голосуют, агитируют, участвуют в митингах и пикетах, обращаются и выступают в СМИ, оказывают финансовую помощь, берутся за оружие и т. д. Эта активность («мощность» действий) может быть выражена через величины $E_X(t)$ и $E_Y(t)$ — общее число производимых в единицу времени равноценных (равных по силе) поступков в пользу той или иной позиции

$$E_X(t) = \sum_{i=1}^{X_0(t)} E_{0_X}^i(t) + \sum_{i=1}^{x(t)} E_x^i(t), \quad (3)$$

$$E_Y(t) = \sum_{i=1}^{Y_0(t)} E_{0_Y}^i(t) + \sum_{i=1}^{y(t)} E_y^i(t), \quad t \geq t_0,$$

где $E_{0_X}^i$, E_x^i , $E_{0_Y}^i$, E_y^i — индивидуальные интенсивности действий участников из категорий $X_0(t)$, $x(t)$, $Y_0(t)$, $y(t)$ соответственно (i — номер участника).

Целью работы является построение и анализ математической модели для определения функций $E_X(t)$, $E_Y(t)$ — интегральных величин (вообще говоря наблюдаемых и измеряемых), количественно описывающих временную динамику процесса выбора электоратом политической линии. Эта динамика, очевидно, зависит от начального состояния социальной общности, её внутренних характеристик (исторического опыта, уровня культуры, ...), внешних воздействий на неё (например, со стороны других электоратов) и от многих других факторов и процес-

¹⁾ конкретные примеры приводятся в § 4 (см. также [5], [6]).

сов. Поэтому для получения достаточно обозримой модели необходимо упростить первоначальную общую проблему выбора.

2.2. Редукция общей проблемы и понятие социально-политического потенциала. В дальнейшем будем рассматривать ситуацию выбора в относительно краткосрочной перспективе, когда справедливы следующие предположения.

Предположение 1. Число твёрдых сторонников и противников не изменяется со временем, т. е.

$$X_0(t) = X_0 \geq 0, \quad Y_0(t) = Y_0 \geq 0, \quad (4)$$

Предположение 2. Индивидуальные интенсивности поступков членов категорий X_0, Y_0 постоянны во времени, т. е.

$$E_{0X}^i(t) = E_{0X}^i \quad (1 \leq i \leq X_0); \quad E_{0Y}^i(t) = E_{0Y}^i \quad (1 \leq i \leq Y_0). \quad (5)$$

Тогда суммарные активности категорий X_0, Y_0 (первые члены в правых частях (3)) равны величинам

$$E_{0X} = X_0 \sum_{i=1}^{X_0} E_{0X}^i, \quad E_{0Y} = Y_0 \sum_{i=1}^{Y_0} E_{0Y}^i, \quad (6)$$

их сумма

$$E_{X+Y}^0 = E_{0X} + E_{0Y} \quad (7)$$

представляет собой неизменную (или слабо изменяющуюся) во времени компоненту («фон») общей активности противоборствующих сторон, а величина

$$E_{X-Y}^0 = E_{0X} - E_{0Y} \quad (8)$$

характеризует постоянную разницу в активности сторонников и противников.

Замечание 1. Величина (7) формируется из двух разнонаправленных составляющих E_{0X} и E_{0Y} («за» и «против») и для постороннего наблюдателя (т. е. для наблюдателя, не интересующегося знаком компонент противоборства) предстаёт как суммарная характеристика общего фона политической активности электората, а величина (8) важна для наблюдателя, интересующегося, прежде всего, фоновой степенью превосходства одной позиции над другой.

В отличие от категорий X_0, Y_0 активности категорий $x(t), y(t)$ (вторые члены в правых частях (3))

$$E_x(t) = \sum_{i=1}^{x(t)} E_x^i(t), \quad E_y(t) = \sum_{i=1}^{y(t)} E_y^i(t), \quad (9)$$

так же, как и их сумма

$$E_{x+y}(t) = E_x(t) + E_y(t) \quad (10)$$

и разность

$$E_{x-y}(t) = E_x(t) - E_y(t) \quad (11)$$

зависят от времени. Вместе с ними зависят от времени итоговые активности категорий X и Y

$$E_X(t) = E_{0_X} + E_x(t), \quad E_Y(t) = E_{0_Y} + E_y(t), \quad (12)$$

а также их сумма

$$\begin{aligned} E_{X+Y}(t) &= E_X(t) + E_Y(t) = E_{0_X} + E_{0_Y} + E_x(t) + E_y(t) = \\ &= E_{X+Y}^0 + E_{x+y}(t) \end{aligned} \quad (13)$$

и разность

$$\begin{aligned} E_{X-Y}(t) &= E_X(t) - E_Y(t) = E_{0_X} - E_{0_Y} + E_x(t) - E_y(t) = \\ &= E_{X-Y}^0 + E_{x-y}(t). \end{aligned} \quad (14)$$

Замечание 1 справедливо также и относительно величин (10), (11) и (13), (14).

В силу Предположений 1, 2 исходная задача сводится к изучению динамики поведения членов электорального болота — категории $n(t)$ (как уже отмечалось $n(t) = z(t) + x(t) + y(t)$). Первый шаг при построении математической модели, описывающей поведение данной категории, состоит в формулировке понятия *социально-политического потенциала (СПП)* [5, 6] отдельной личности из электората (точнее, из его действующей части $x(t) + y(t)$):

Под СПП члена электората будем понимать его готовность осуществлять поступки в пользу выбранной позиции (соответственно, величина СПП представляет собой отнесённую к единице времени меру этой готовности). СПП относится к потенциалам, так сказать, гуманитарной группы, в отличие от, например, экономического, научно-технического или военного потенциалов (более подробно о потенциалах см. [7–9]).

Представители категорий $x(t)$ или $y(t)$, очевидно, не всегда имеют условия для полной реализации своего СПП (например, существующая в какой-то момент времени готовность электората сменить своё политическое руководство «замораживается» до наступления ближайших выборов). Поэтому индивидуальные интенсивности их действий (см. (3)) связаны с их индивидуальными потенциалами $P_x^i(t)$, $P_y^i(t)$ соотношениями

$$\begin{aligned} E_x^i(t) &= \alpha_x^i(t) P_x^i(t) \quad (1 \leq i \leq x(t)), \\ E_y^i(t) &= \alpha_y^i(t) P_y^i(t) \quad (1 \leq i \leq y(t)), \end{aligned} \quad (15)$$

где коэффициенты $0 \leq \alpha_x^i(t) \leq 1$, $0 \leq \alpha_y^i(t) \leq 1$ характеризуют возможность реализации существующего потенциала. Будем считать справедливым

Предположение 3. *Возможности для полной реализации СПП существуют* (т.е. в (15) коэффициенты $\alpha_x^i(t) = \alpha_y^i(t) = 1$).

Тогда *основная задача* данной работы сводится к построению математической модели для динамики численностей $x(t), y(t)$ и индивидуальных потенциалов $P_x^i(t), P_y^i(t)$. Если они найдены, то по формулам (15) находятся величины (9), величина (10), определяющая переменную во времени составляющую общей активности противоборства на статическом фоне борьбы (формула (7)) и величина (11) — разница в активности категорий $x(t), y(t)$ на неизменном фоне преобладания одной позиции над другой (величина (8)), а при известных величинах (4), (5) — величины (6), (12), (3) (т.е. все искомые величины).

Замечание 2. Предположение 3 вводится для упрощения. Если из каких-то соображений величины $\alpha_x^i(t), \alpha_y^i(t)$ известны, то основная задача остаётся прежней.

Очевидно, что динамика СПП (так же как и динамика суммарных активностей (3) в исходной задаче) определяется сложной совокупностью разнообразных факторов и процессов. Будем подразделять их на внутренние и внешние. Первые связаны с собственными поведенческими характеристиками электората, т.е. отвечают идеальной, мысленной ситуации, когда электорат предоставлен самому себе — является собой коллективного «робинзона». Ко вторым относятся влияния «окружающей среды» — других электоратов, СМИ и прочих внешних по отношению к рассматриваемому электорату влияний (финансовых, экономических, религиозных, культурных, ...).

В дальнейшем будем считать, что обусловленное внутренними (собственными) причинами поведение членов электората (и формируемая ими часть СПП) основано на *подражании* (см. работы [10–12] и цитируемую в них литературу) и описывается следующей *социально-психологической* Гипотезой:

1. *Член электората осуществляет свой выбор (самоидентификацию) путём подражания, считывая позитивные и негативные стимулы в ходе общения со своими соседями по электорату («сканируя» круг своего общения);*

2. *Стимулами, определяющими выбор, служат как величина абсолютной разницы между сторонниками и противниками данного курса*

$$\begin{aligned} Q(t) &= X(t) - Y(t) = X_0 - Y_0 + x(t) - y(t) = Q_0 + q(t), \\ Q_0 &= X_0 - Y_0, \quad q(t) = x(t) - y(t), \quad t \geq t_0, \end{aligned} \quad (16)$$

так и величина скорости изменения со временем (производной) этой разницы

$$S(t) = \dot{Q}(t) = \dot{X}(t) - \dot{Y}(t) = \dot{x}(t) - \dot{y}(t) = \dot{q}(t), \quad (17)$$

где $\dot{f} \equiv df/dt$, $t \geq t_0$. Тем самым собственная часть СПП формируется статической (определяемой функцией $Q(t)$) и ажиотажной (определяемой функцией $S(t)$) составляющих.

Обозначая через $F_x^i(t)$, $F_y^i(t)$ силу влияния внешних факторов (например, активности СМИ) на формирование СПП (функции $F_x^i(t)$, $F_y^i(t)$ считаются известными) и принимая во внимание Гипотезу, приходим к следующим связям между введёнными величинами (между величинами индивидуальных СПП и аргументами, от которых они зависят):

$$P_x^i(t) = P_x^i(Q, S, F_x^i(t)) = P_x^i(X_0 - Y_0 + x(t) - y(t), \dot{x}(t) - \dot{y}(t), F_x^i(t)), \\ 1 \leq i \leq x(t), \quad (18)$$

$$P_y^i(t) = P_y^i(Q, S, F_y^i(t)) = P_y^i(X_0 - Y_0 + x(t) - y(t), \dot{x}(t) - \dot{y}(t), F_y^i(t)), \\ 1 \leq i \leq y(t).$$

Число неизвестных в (18) больше, чем количество связей. Поэтому для получения замкнутой модели необходимо сформулировать предположения и упрощения, дополнительные к тем, что были сделаны ранее.

Прежде всего, учитывая практическую невозможность определения индивидуальных функциональных зависимостей в связях (18) для каждого члена из категорий $x(t)$, $y(t)$ (вообще говоря, разных для разных представителей электората) введём понятие *средней типической личности*, характеризующей рассматриваемую социальную общность. Другими словами, вместо множества функций P_x^i , P_y^i , F_x^i , F_y^i будем рассматривать лишь функции $P(t)$ (СПП типической личности) и $F(t)$ (силу воздействия на неё внешних факторов).

Важно отметить, что при таком подходе *типическая личность трактуется как амбивалентная социальная единица, совершающая, вообще говоря, поступки как в пользу позиции X, так и в пользу позиции Y*. В реальности в категориях $x(t)$, $y(t)$ могут находиться члены, совершающие в данный момент времени поступки только «за», только «против» и смешанные действия — как «за», так и «против». Поясним также, что для целей данного исследования понятие средней типической личности плодотворно лишь при условии, если в электорате работают механизмы, синхронизирующие во времени действия составляющих его членов (или, хотя бы, значительного числа — подробнее см. § 4). Только в этом случае по поведению средней типической можно судить, как это и делается ниже, о поведении электората в целом.

Понятие типической личности агрегирует разнообразие поведения членов категории $x(t) + y(t)$ и, оперируя с ним, вместо (18) можно ограничиться лишь единственной связью

$$P(t) = P(Q, S, F(t)) = P(X_0 - Y_0 + x(t) - y(t), \dot{x}(t) - \dot{y}(t), F(t)). \quad (19)$$

Однако, и после сделанного упрощения связь (19) ещё не является замкнутой моделью, так как она содержит две искомые величины: $P(t)$ и $Q(t)$. Связь замыкается, очевидно, в двух случаях:

а) На рассматриваемом промежутке времени известны величины $Q(t)$ и $\dot{Q}(t)$ (например, получены из измерений). Тогда из (19) непосредственно находится потенциал $P(t)$, затем величина (14), а затем, если известны численности X_0, Y_0 и величины E_{0X}^0, E_{0Y}^0 , находятся величины $E_{0X}, E_{0Y}, E_{X+Y}^0, E_{X-Y}, E_X, E_Y, E_{X+Y}$ (формулы (6), (7), (8));

б) На рассматриваемом промежутке времени известен из каких-либо соображений потенциал $P(t)$. Тогда из (19) (в этом случае связь (19) представляет собой неявное обыкновенное дифференциальное уравнение первого порядка относительно $Q(t)$ при известном начальном значении $Q(t_0)$) находится функция $Q(t)$ на всём промежутке времени, а вместе с ней и все те же величины, что и в случае а).

Очевидно, что в обоих случаях для нахождения остальных характеризующих процесс величин ($E_x, E_y, E_{x+y}, E_{x-y}, E_X, E_Y, E_{X+Y}$ — формулы (9)–(13)) нужна дополнительная связь между численностями $X(t), Y(t)$ (например, знание общей численности колеблющегося электората, т. е. величины $x(t) + y(t)$).

Случай а) ввиду его математической тривиальности интереса для моделирования не представляет, являясь, прежде всего, предметом соответствующего социологического анализа. Поэтому в дальнейшем будем изучать случай б). Удобно, продифференцировав (19) по времени

$$\frac{dP}{dt} = \frac{dP}{dQ} \frac{dQ}{dt} + \frac{dP}{dS} \frac{dS}{dt} + \frac{dP}{dF} \frac{dF}{dt}$$

и переписав последнее выражение в эквивалентном виде

$$\frac{dS}{dt} = \left(-\frac{dP}{dQ} \frac{dQ}{dt} + \frac{dP}{dt} - \frac{dP}{dF} \frac{dF}{dt} \right) \bigg/ \frac{dP}{dS},$$

перейти, с учётом равенства $dQ/dt = S$, к системе двух обыкновенных дифференциальных уравнений, разрешённых относительно величин Q и S

$$\frac{dS}{dt} = \frac{-P_Q S + P_t - P_F F_t}{P_S} \frac{dQ}{dt} = S, \quad t > t_0. \quad (20)$$

Динамическая система 2-го порядка (20) вместе с поведенческими характеристиками

$$\begin{aligned} P_Q &= P_Q(Q, S, F(t)) = \frac{dP}{dQ}, \\ P_S &= P_S(Q, S, F(t)) = \frac{dP}{dS}, \\ P_F &= P_F(Q, S, F(t)) = \frac{dP}{dF}, \end{aligned} \quad (21)$$

заданными функциями времени

$$P_t = \frac{dP}{dt}, \quad F_t = \frac{dF}{dt} \quad (22)$$

и начальными значениями (*задача Коши*)

$$Q(t_0) = Q^0, \quad S(t_0) = S^0 \quad (23)$$

представляет собой замкнутую математическую модель рассматриваемого процесса, полученную в рамках *Предположений* 1–3, социально-психологической *Гипотезы* и представления о *средней типичской личности*. Из уравнений (20) с входными данными (21)–(23) однозначно (при соответствующих условиях гладкости) находится искомое решение — функции $Q(t)$, $S(t)$ в любой момент времени $t > t_0$.

Замечание 3. Как следует из вида связи (19), помимо случаев а) и б), возможно рассмотрение случая в), когда по известным потенциалу $P(t)$ и численности $Q(t)$ находится интенсивность внешнего воздействия $F(t)$ на электорат (своего рода «обратная задача социальной психологии»).

2.3. Зависимость СПП от основных переменных. Свойства модели (20)–(23) могут изучаться с общих позиций (при входных данных, сформулированных в достаточно общем виде). Однако, для получения конкретных результатов требуется задать конкретные входные данные (желательно, в явном виде), прежде всего, задать вид функции $P = P(Q, S, F(t))$, определяющей СПП. Для этого требуется сделать ряд дополнительных упрощений.

Упрощение 1. СПП либо не зависит от внешнего воздействия $F(t)$, т.е.

$$P = P(Q, S), \quad P_F \equiv 0,$$

либо внешнее воздействие постоянно во времени, т.е.

$$F(t) = \text{const}, \quad F_t \equiv 0, \quad P = P(Q, S, C). \quad (24)$$

либо оба этих свойства выполняются одновременно. Для простоты в дальнейшем будем (во втором случае) опускать параметрическую зависимость потенциала от постоянной C и записывать его (как и в первом случае) в виде $P = P(Q, S)$.

Упрощение 2. СПП является аддитивной функцией своих аргументов (зависит от величин Q, S по отдельности):

$$P(Q, S) = \chi_1 P_1(Q) + \chi_2 P_2(S), \quad (25)$$

где $\chi_1 \geq 0$, $\chi_2 \geq 0$ — «веса» статичной и динамической компонент СПП ($\chi_1 + \chi_2 = 1$).

Для окончательной формулировки вида функций $P_1(Q)$, $P_2(S)$ обратимся к психофизиологическим закономерностям восприятия, ис-

следовавшихся в работах [13–16]. Основываясь на их результатах, запишем функции $P_1(Q)$, $P_2(S)$ следующим образом

$$P_1(Q) = \frac{\alpha Q^2}{(Q^2 + \beta)}, \quad (26)$$

$$P_2(S) = \mu \sqrt{\frac{2}{\pi}} \frac{S^2}{\rho^3} \exp\left(-\frac{S^2}{2\rho^2}\right), \quad (27)$$

где $\alpha > 0$, $\beta > 0$, $\rho > 0$, $\mu > 0$ — параметры, численно характеризующие особенности восприятия (вообще говоря, разные для разных объектов) и служащие для «привязки» функций к эмпирически установленным конкретным зависимостям (законам Вебера–Фехнера, Бугера и Стивенса).

Дадим краткий комментарий к формулам (26), (27) (см. также рис. 1).

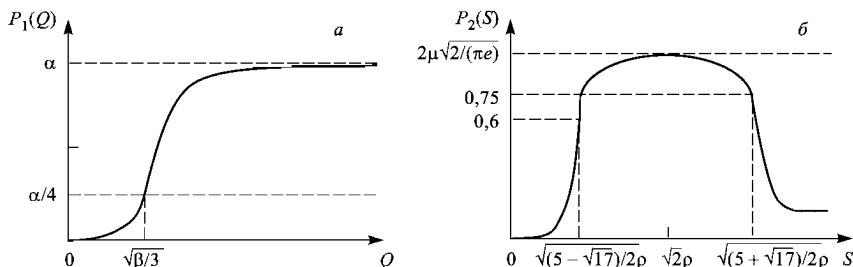


Рис. 1. Зависимость статической компоненты СПП а) от разницы между числом сторонников и противников Q ; б) от скорости изменения разницы между числом сторонников и противников S

Чётная функция (26) (рис. 1 а) монотонно растёт с ростом $|Q|$ и выходит при $Q \rightarrow \pm\infty$ на *асимптотическое значение* $P_1(\infty) = \alpha$, определяемое параметром α . При этом производная (нечётная функция)

$$\frac{dP_1}{dQ} = \frac{2\alpha\beta Q}{(Q^2 + \beta)^2}$$

положительна при $Q > 0$, отрицательна при $Q < 0$, равна нулю при $Q = 0$, стремится к нулю при $Q \rightarrow \pm\infty$ и достигает своего *максимального по модулю значения*, равного $\pm(\sqrt{3}/2)^3 \alpha/\sqrt{\beta}$ в точках $Q = \pm\sqrt{\beta/3}$ (в силу симметрии функции $P_1(Q)$ относительно точки $Q = 0$ её график приведён лишь при $Q \geq 0$).

Другими словами, с увеличением стимула (аргумента $|Q|$) отклик на стимул (функция $P_1(Q)$) увеличивается с «насыщением» изменяясь (при $|Q| > \sqrt{\beta/3}$) всё слабее и слабее. Этот эффект называется *привыканием с утомлением*.

Чётная функция $P_2(S)$ (рис. 1б) монотонно растёт с ростом $|S|$ вплоть до точки с координатой $|S| = \sqrt{2} \rho$, где она *максимальна* и равна $P_2(S)_{\max} = 2\mu \sqrt{2/(\pi\epsilon)}$, для $|S| > \sqrt{2} \rho$ она монотонна и уменьшается до нуля при $|S| \rightarrow \infty$ (аналогично рис. 1а график функции $P_2(S)$ приведён лишь в области $S \geq 0$). Производная (нечётная функция)

$$\frac{dP_2}{dS} = P_2(S) \rho^{-2} \frac{2\rho^2 - S^2}{S}$$

равна нулю при $S = 0$ и $|S| = \sqrt{2} \rho$ и стремится к нулю при $|S| \rightarrow \infty$, достигая своих *максимальных по модулю значений* в точках

$$S = \pm \sqrt{\frac{5 + \sqrt{17}}{2}} \rho \quad \text{и} \quad S = \pm \sqrt{\frac{5 - \sqrt{17}}{2}} \rho.$$

В отличие от функции $P_1(Q)$, динамическая компонента потенциала $P_2(S)$ увеличивается с увеличением стимула (аргумента $|S|$) лишь до достижения определённого *порога восприятия* (определяемого значением параметра ρ), а при дальнейшем его нарастании стремится к нулю. В [13–16] этот эффект интерпретируется как *восприятие растущей по объёму информации* (при превышении пропускной способности воспринимающего) *в качестве растущего по амплитуде «шума»* (всё более заглушающего полезный сигнал).

Заметим, что определяемые параметрами α, β, μ, ρ значения аргументов функций (26), (27), самих функций и их производных в характерных точках имеют вполне ясную интерпретацию (максимальный отклик на стимул, максимальная скорость роста отклика на стимул, порог восприятия и т. д.).

§ 3. Математическая формулировка базовой модели и описываемые ею режимы эволюции

3.1. Постоянство СПП и уравнения базовой модели. В дальнейшем модель, получающуюся из модели (20)–(23) при *Упрощениях* 1, 2 (п. 3, § 2), зависимостях (26), (27) и в *Предположении* 4:

$$P_t = \frac{dP}{dt} = 0, \quad (28)$$

т. е. при *постоянстве СПП во времени*, будем называть *базовой моделью* электорального поведения (базовая модель была впервые предложена и предварительно изучена в работах [5, 6]).

Из (20), (26), (27) с учётом сделанных *Упрощений* и *Предположения* 4 следуют уравнения базовой модели

$$\frac{dS}{dt} = -S \frac{P_Q}{P_S} = -\frac{2\chi_1 \alpha \beta \rho^5 Q \exp \frac{S^2}{2\rho^2}}{\sqrt{2/\pi} \chi_2 \mu (Q^2 + \beta^2) (2\rho^2 - S^2)}, \quad \frac{dQ}{dt} = S, \quad t > t_0, \quad (29)$$

дополняемые начальными условиями

$$Q(t_0) = Q^0, \quad S(t_0) = S^0. \quad (30)$$

Динамическая система (29) автономна и её удобно исследовать в фазовой плоскости Q, S , т. е. изучать методами качественной теории ОДУ поле интегральных кривых нелинейного уравнения первого порядка

$$\frac{dS}{dQ} = - \frac{2\chi_1\alpha\beta\rho^5 Q \exp \frac{S^2}{2\rho^2}}{\sqrt{2/\pi} \chi_2\mu (Q^2 + \beta)^2 (2\rho^2 - S^2) S}. \quad (31)$$

В содержательном отношении базовая модель описывает достаточно широкий и типичный круг ситуаций, условно говоря, *мобилизационного* («военного») характера, когда СПП социальной общности «возбуждён» и, в течение некоторого времени, поддерживается на определённом уровне.

Заметим, что система (20) автономна не только в случае постоянства СПП, но и тогда, когда СПП является линейной функцией времени

$$P(t) = C_0 + Ct \geq 0, \quad C_0 \geq 0, \quad t \geq t_0. \quad (32)$$

Тогда уравнения (20) принимают вид

$$\begin{aligned} \frac{dS}{dt} &= \frac{-P_Q S + C}{P_S}, \\ \frac{dQ}{dt} &= S, \quad t > t_0, \end{aligned}$$

откуда следуют уравнения для *расширенной базовой модели* (аналоги (29) при условии (32))

$$\begin{aligned} \frac{dS}{dt} &= \frac{\rho^5}{\sqrt{2/\pi} \chi_2\mu} \left(-\chi_1 \frac{2\alpha\beta Q}{(Q^2 + \beta)^2} S + C \right) \frac{\exp \frac{S^2}{2\rho^2}}{S(2\rho^2 - S^2)}, \\ \frac{dQ}{dt} &= S, \quad t > t_0 \end{aligned} \quad (33)$$

и уравнение (аналог (31))

$$\frac{dS}{dQ} = \frac{\rho^5}{\sqrt{2/\pi} \chi_2\mu} \left(-\frac{2\chi_1\alpha\beta Q S}{(Q^2 + \beta)^2} + C \right) \frac{\exp \frac{S^2}{2\rho^2}}{S^2(2\rho^2 - S^2)}. \quad (34)$$

3.2. Монотонные режимы эволюции «электората». Для изучения временной динамики величин Q, S исследуются фазовые траектории уравнения (31) в плоскости Q, S . Фазовые портреты несколько различаются в случаях $a = a_{кр}$, $a > a_{кр}$ и $a < a_{кр}$, где $a = 2\chi_1\alpha\beta\rho^5(\sqrt{2/\pi} \chi_2\mu)^{-1}$, $a_{кр} = 4e\rho^4/\beta$. Результаты подробного математического исследования фазовых траекторий уравнений базовой модели

(допускающих интегрирование в квадратуре) и расширенной базовой модели содержатся в [17]. Здесь мы ограничимся изложением для промежуточного значения $a = a_{кр}$, кратко пояснив, в чём состоит отличие данного случая от остальных.

Уравнение (31) имеет три особые точки: $(0, 0)$, $(0, \sqrt{2} \rho)$, $(0, -\sqrt{2} \rho)$ — точки пересечения изоклины нуля ($Q = 0$) с изоклинами бесконечности ($S = 0$, $S = \pm\sqrt{2} \rho$). Первая из них является точкой типа центра, вторая и третья имеют тип седла. Картина поля интегральных кривых уравнения приведена на рис. 2, стрелками показано направление движения по кривым при увеличении t (напомним, что движение осуществляется при постоянстве СПП, определяемого значениями Q^0 , S^0 и разного для разных кривых), пунктирные линии — изоклины бесконечности $S = \pm\sqrt{2} \rho$, штрихпунктирные линии А и В — сепаратрисы седел.

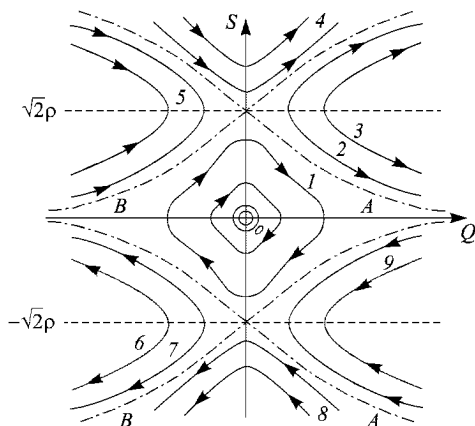


Рис. 2. Фазовый портрет динамической системы (29) в плоскости Q, S :

$$S = \frac{dQ}{dt} = \frac{dq}{dt}, \quad Q = Q_0 + q$$

Замечание 4. Анализ в плоскости (Q, S) может трактоваться также и как анализ в плоскости переменных $q(t) = Q(t) - Q_0$ и $S = dQ/dt = dq/dt$ (см. (16), (17)). Это обстоятельство отражено в обозначениях осей на рис. 2 (при $Q_0 = 0$ в трактовке величин $Q(t)$ и $q(t)$, очевидно, различий нет). В реальности величины Q_0 и $q(t)$ ограничены. Поэтому содержательная часть плоскости (Q, S) заведомо ограничена областью

$$Q_0 - n_M \leq Q \leq Q_0 + n_M, \quad (35)$$

где $n_M = (z(t) + x(t) + y(t))_M$ — максимально возможная численность электорального болота (см. 2).

После этих предварительных замечаний перейдём к описанию допускаемых моделью режимов эволюции электората и их интерпретации (разным режимам отвечают разные типы интегральных кривых I–IX).

I) Замкнутые кривые этого типа описывают периодический процесс и подробно рассматриваются в п. 2, 3.

II) В данном режиме величина Q неограниченно растёт с течением времени — *весь электорат в итоге занимает позицию «X»*. При этом скорость перехода, так же как и ажиотажная часть потенциала, уменьшается до нуля (значения скорости лежат в допороговой области $0 < S \leq \sqrt{2} \rho$). Асимптотическое поведение величин $S(Q)$ и $Q(t)$ даётся формулами:

$$S = \frac{b}{Q} + \dots, \quad Q = \sqrt{bt + (Q^0)^2} + \dots, \quad b = \sqrt{\frac{a}{2\rho^2}}, \quad (36)$$

где точками обозначены члены более высокого порядка малости и из которых видно, что *переход всего электората на единую позицию происходит за бесконечное время*. Реальное время перехода, с учётом (35), разумеется, конечно и может быть рассчитано, если известны значения Q^0 , S^0 и параметры a и ρ . Заметим, что явные выражения (36) могут, при известных условиях, использоваться для качественного анализа динамики соответствующих процессов (в частности, видно, что при больших Q она определяется лишь одним параметром b , представляющим собой комбинацию всех параметров задачи).

III) Как и в случае II, величина Q неограниченно растёт. Однако (находящаяся в диапазоне $\sqrt{2}\rho < S < \infty$) *скорость перехода на единую позицию не уменьшается, а неограниченно растёт со временем* (при этом ажиотажная часть потенциала уменьшается до нуля). Асимптотика функции $S(Q)$ даётся неявной формулой

$$Q = b \cdot \frac{e^{S^2/4\rho^2}}{S} + \dots, \quad (37)$$

анализ которой показывает, что переход всего электората на позицию «X» происходит, как и в предыдущем случае, за бесконечное время (асимптотика процесса также определяется лишь одним параметром b).

Заметим, что при значениях S , близких к величине $\sqrt{2}\rho$, *эволюция электората «неустойчива»* в том смысле, что малые изменения величины S могут привести к переходу с режима II на режим III (и наоборот).

IV) Как и в случаях II, III, величина Q неограниченно растёт при $t \rightarrow \infty$. Скорость S , как и в случае III, также неограниченно растёт при $t \rightarrow \infty$ (для интегральных кривых IV справедлива асимптотика (37)).

Данный режим отличается от режима III тем, что он *может осуществляться* при начальных значениях $Q^0 < 0$, т.е. *при начальном преобладании противников позиции «X» над её сторонниками*. Однако, достаточно большое значение начальной скорости перехо-

да $S^0 > 0$ обуславливает итоговую победу позиции «Х» (при этом скорость изменяется немонотонно, достигая минимума в некоторый момент времени).

Итак, режимы II–IV, различаясь в деталях, отвечают *необратимой эволюции всего электората в пользу одной позиции*. Для осуществления такой эволюции начальные значения величин Q^0, S^0 должны лежать в области плоскости Q, S , расположенной выше сепаратрисы A (см. рис. 2). Другими словами, *должна быть достаточно большой либо величина $Q^0 > 0$, либо (при $Q^0 < 0$) величина S^0* . Причём, *начальная скорость перехода всегда положительна* (конкретные численные значения величин, характеризующих динамику эволюции легко находятся из модели, если известны конкретные входные данные).

V) Интегральные кривые V (лежащие между сепаратрисами A и B при $Q^0 < 0$) описывают *процесс стабилизации числа сторонников и противников позиции «Х»*. Величина Q растёт и достигает некоторого конечного отрицательного значения $Q = Q_s$, приближаясь к нему либо из области $S < \sqrt{2} \rho$, либо из области $S > \sqrt{2} \rho$, т.е. *стабилизация происходит при достижении порога восприятия $S = \sqrt{2} \rho$* . Данная эволюция *устойчива*, в том смысле, что любое изменение величин Q^0, S^0 (в описываемом диапазоне) не изменяет характера процесса — стабилизация в конце концов осуществляется, но при других значениях величины Q_s (ср. со случаем III).

В окрестности линии $S = \sqrt{2} \rho$ для кривых V справедливы асимптотики

$$\begin{aligned} S &= \sqrt{2} \rho \pm \sqrt{2b_1(Q - Q_s)} + \dots, \\ Q &= Q_s + \sqrt{2} \rho(t - t_s) + \dots, \quad b_1 = \frac{aQ_s e}{4\rho^2(Q_s + \beta)^2}, \end{aligned} \quad (38)$$

из которых следует, что стабилизация достигается в некоторый конечный момент времени $t = t_s$ (рассчитываемый, как и другие характеристики процесса, из модели по известным входным данным).

Как ясно следует из анализа кривых V , *стабилизация достигается, если в начальный момент противники позиции «Х» преобладают над сторонниками, а начальная скорость перехода на позицию «Х» не слишком велика, но и не слишком мала* (при этом в стабильном состоянии $Q_s < 0$, т.е. число сторонников позиции «Х» меньше числа противников).

Анализ кривых типа VI, VII, VIII, IX и свойства описываемых ими режимов (включая формулы (36)–(38)) аналогичны в силу симметрии модели и, соответственно, симметрии фазовой плоскости Q, S , анализу случаев II, III, IV, V соответственно. Разница лишь в том, что электорат полностью переходит на позицию «Y» — случаи VI, VII, VIII (ср. с II, III, IV), а стабилизация (случай IX) реализуется с преобладанием числа сторонников позиции «Х»

3.3. Циклы самоидентификации электората. Интегральные кривые I располагаются в части области $-\sqrt{2}\rho < S < \sqrt{2}\rho$ плоскости Q, S , ограниченной сверху и снизу сепаратрисами A и B (рис. 2). Они описывают процесс периодического изменения величин Q и S со временем. Амплитуда колебаний скорости $S_{\max}$ ограничена ($|S_{\max}| < \sqrt{2}\rho$, допороговая область — см. рис. 1 б). Причём максимальное значение величины $S = S_M$ на замкнутой траектории достигается при $Q = 0$ и связано с максимальным значением величины $Q = Q_M$ (достигаемом при $S = 0$) соотношением, следующим из формул (25), (26), (27) и интеграла системы (постоянства СПП — см. (28))

$$\sqrt{\frac{2}{\pi}} \chi_2 \mu \rho^{-3} S_M^2 \exp\left(-\frac{S_M^2}{2\rho^2}\right) = \frac{\alpha \chi_1 Q_M^2}{Q_M^2 + \beta}.$$

Колебания численности Q могут иметь любую амплитуду (с учётом, разумеется, (34)). Из вида кривых I ясно следует, что циклическая эволюция электората осуществляется в условиях, когда начальные значения скорости S^0 и численности Q^0 не слишком велики (в отличие от рассмотренных в п. 2 режимов неограниченного роста и стабилизации числа сторонников и противников позиции «X» или «Y»). Конкретные численные характеристики колебаний нетрудно найти из модели по известным входным данным.

В окрестности особой точки (0, 0), т.е. при малых амплитудах ($Q \ll 1$, $S \ll 1$) процесс описывается классическим линейным уравнением колебаний

$$\frac{d^2 Q}{dt^2} = -b_2 Q, \quad b_2 = \frac{a}{2\beta^2 \rho^2}$$

с периодом

$$T_\ell = 2\pi \sqrt{\frac{2}{\pi}} \sqrt{\frac{\chi_2 \mu \beta}{\chi_1 \alpha \rho^3}}, \quad (39)$$

определяемым комбинацией параметров модели.

В случае колебаний (нелинейных) произвольной амплитуды их период зависит также и от амплитуды (Q_M) и выражается формулой

$$T = T_\ell \frac{2}{\pi} \left(\frac{Q_M^2}{\beta} + 1 \right) E_i \left(\frac{Q_M}{(Q_M^2 + \beta)^{1/2}} \right), \quad (40)$$

получаемой двукратной квадратурой из уравнений (29) (с учётом существования интеграла (28)). Значения эллиптического интеграла E_i изменяются в пределах $1 < E_i < 1,57$ (т.е. слабо), его конкретные численные значения находятся по известному аргументу.

Замечание 5. При $a > a_{\text{кр}}$ поведение фазовых траекторий уравнения (31) аналогично, за тем исключением, что на сепаратрисах A и B величина $|S|$ стремится при $|Q| \rightarrow \infty$ не к нулю, а к конечному значению $\pm S_a = S_a(\alpha, \beta, \rho, \dots)$. Особенность случая $a < a_{\text{кр}}$ в том, что сепаратрисы A и B не разомкнуты, а замкнуты и пересекают

ось Q в конечных точках $\pm Q_{\max} = Q_{\max}(\alpha, \beta, \rho, \dots)$ т. е. колебания ограничены как величиной $S = S_{\max}$, так и величиной $Q = Q_{\max}$.

Дадим краткий комментарий к формулам (39), (40). Наиболее сильно период колебаний зависит от ρ (пропорционален $\rho^{-3/2}$) — величины порога восприятия, входящей в ажиотажную часть потенциала, уменьшаясь с её увеличением. Зависимость периода от величин α , χ_1 (максимального значения статичной части потенциала и его «веса» в общем потенциале) качественно аналогична, но заметно более слабая. Также относительно слабо (как корень квадратный) период зависит от величин χ_2 , β , μ (их смысл ясен из предыдущих описаний), увеличиваясь с их увеличением. Описанные зависимости могут быть интерпретированы следующим образом.

Выражение (25) для потенциала $P(Q, S)$ запишем, с учётом (26), (27), в эквивалентном виде

$$P(Q, S) = \chi_1 P_1 + \chi_2 P_2 = \varkappa_1 \frac{Q^2}{Q^2/\beta + 1} + \varkappa_2 2\pi S^2 \exp\left(-\frac{S^2}{2\rho^2}\right). \quad (41)$$

Из (41) ясно видно, что коэффициенты

$$\varkappa_1 = \frac{\chi_1 \alpha}{\beta}, \quad \varkappa_2 = \frac{\chi_2 \mu}{\rho^3}$$

характеризуют (при заданных значениях Q и S) вклад статической и «кинетической» (связанной со скоростью) компонент в общий потенциал P . В терминах коэффициентов $\varkappa_1$, $\varkappa_2$ выражение (39) для T_ℓ даётся формулой

$$T_\ell = 2\pi \sqrt[4]{\frac{2}{\pi}} \sqrt{\frac{\varkappa_2}{\varkappa_1}}, \quad (42)$$

аналогичной классической формуле

$$T = \sqrt{\frac{m}{k}}$$

для периода колебаний шарика массы m , присоединённого к пружине с жёсткостью k . Напомним, что в системе «шарик — пружина» величина массы m характеризует её инерционные (кинетические) свойства, а величина жёсткости k — её силовые (статические) свойства.

Конечно, аналогия между колебаниями электората и колебаниями механической системы весьма условна, но, тем не менее, она позволяет дать более ясную интерпретацию зависимости (42).

1. Чем больше удельный вес статической части потенциала в общем постоянном потенциале, тем сравнительно больше «сила принуждения» личности со стороны большинства и тем быстрее она совершает переходы колебания (период уменьшается с ростом $\varkappa_1$).

2. Чем больше удельный вес ажиотажной (кинетической) части потенциала, тем сравнительно больше инерционность личности (тем сравнительно меньше, в условиях постоянства потенциала, сила при-

нуждения со стороны большинства) и тем больше времени требуется для изменения её позиции (период растёт с ростом κ_2).

В случае нелинейных колебаний период (40) весьма сильно (квадратично) зависит от максимального значения (Q_M) разницы между сторонниками и противниками позиции «X» или «Y», увеличиваясь с её ростом. Это естественно, так как при прочих равных условиях разворот большей по численности части электората в сторону противоположной позиции требует большего времени.

Тем самым изучение модели (29), (30) показывает, что существует принципиальная *возможность* (до определённой степени) *осмысленного управления циклическими режимами эволюции электората* (то же самое относится к режимам, описанным в п. 2).

Заметим также, что для электоратов, обладающих разными свойствами (разными значениями величин χ_1 , χ_2 , α , β , μ , ρ , Q_M), периоды (39), (40) могут быть близкими, т.е. *может существовать количественное подобие электорального поведения*.

Изученные в данном параграфе режимы эволюции со временем числа сторонников и противников какой-либо из позиций естественно рассматривать как *режимы самоидентификации электората*, описывающие преимущественный выбор электорального болота в пользу той или иной позиции. Модель допускает три типа таких режимов: полный переход электората на позицию «X» или «Y» — *однозначная самоидентификация*; переход к стабильному соотношению числа сторонников и противников — *поляризованная самоидентификация*; периодическое преобладание числа сторонников и противников — *циклическая самоидентификация*. Периоды колебаний (39), (40) естественно трактовать как *периоды самоидентификации электората*. В дальнейшем основное внимание будет уделено третьему типу эволюции электората.

§ 4. Циклы политической самоидентификации дагестанского этноса — модельные прогнозы и сопоставление с реальностью

4.1. Описание ситуации, связь наблюдаемых величин с переменными модели. Общеизвестно, что северокавказские субъекты РФ ещё в советское время относились к числу наиболее неблагополучных в экономическом, социальном и политическом отношениях. Трудоизбыточность и особое геополитическое положение, сосуществование разных конфессий и традиций, многоэтничность и драматическая история — вот далеко не самый полный перечень причин, постоянно порождающих повышенную, по сравнению со среднероссийским уровнем, социально — политическую напряжённость в этом регионе. После распада СССР, прекращения идеологической монополии одной партии в России, начала экономических и политических реформ проблемы, таившиеся ранее «под спудом», обнаружались с особой остротой и силой.

В заметной части северокавказских этносов явно или завуалированно обозначилась тенденция решать проблемы своего региона сепаратно от России. При этом выдвигаются различные обоснования, прежде всего этнического, а также религиозного и исторического характера, рассматриваются различные степени политической независимости от российской государственности, вплоть до создания новых самостоятельных государств, и применяются различные методы борьбы, вплоть до жёсткого вооружённого противостояния. Тем самым, в ряде этих регионов имеют место этнополитические конфликты между сторонниками сепаратизма и противниками этой линии, считающими, что решение основных проблем этноса следует искать, оставаясь в составе единого государства.

Одним из весьма проблемных в этом отношении регионов является Дагестан. Причём, сепаратистские настроения (согласно социологическим данным затрагивающие значительную часть населения) зачастую проявляются скрытно, принимая форму тех или иных действий, дестабилизирующих социально — политическую ситуацию в регионе. Поэтому выраженная количественно интенсивность этих действий служит (опосредовано) текущей наблюдаемой мерой сепаратистских проявлений (или шире — социально-политической напряжённости (СПН)) в регионе. В терминах рассматриваемой в § 1 модели речь идёт о наблюдаемой величине E_x (см. формулы (3) и (12)), где позиция «X» — позиция в пользу сепаратизма.

Применительно к Дагестану (и к большей части северокавказских этносов) электорат разбивается на четыре группы:

- твёрдые сторонники (противники) позиции «X» или «Y» или нейтралы (по данным социальной психологии их число не превышает нескольких процентов);

- действующие в пользу позиции «X» («Y», «Z») в соответствии с абсолютным послушанием (например, так называемые мюриды или берущие на себя обязательства подчинения за оплату, как «рекрутированные ваххабиты»);

- находящиеся под действием пропаганды или обоснованных социальных чувств, например, родственники и друзья пострадавших от боевых действий (они составляют от 20% до 40% от численности этноса);

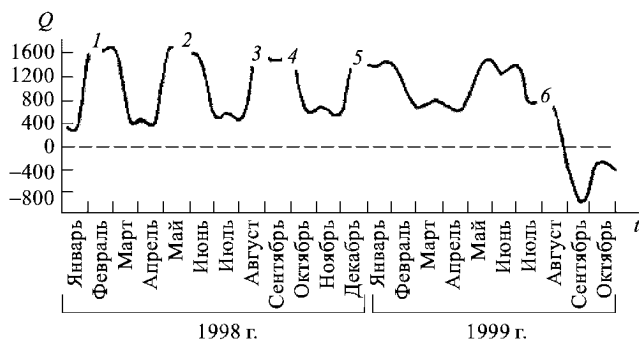
- выбирающие свою позицию в результате подражания, «заряженности» настроением, преобладающие в их социальной группе.

Из указанных групп и формируются категории $X(t)$, $Y(t)$, $Z(t)$, модели для динамики численности которых построены и изучены в §§ 3, 4. Напомним, что модели предусматривают синхронизацию действий категорий «X» и «Y» (в данном случае такими «маркирующими датами» могут служить день рождения имама Шамиля, выборы в органы власти и т.д.). Заметим также, что, согласно этнографическим представлениям и данным (см., например, [18] и цитируемую там литературу), модели, рассмотренные в § 3, хорошо соответствуют

поведению жителей крупных сёл и небольших городов республики, в которых проживает основная часть населения (причём, в условиях жёсткого противостояния). Дальнейшие обобщения этих моделей учитывают поведение горожан, межрегиональные отношения и другие значимые процессы. После этих предварительных замечаний опишем некоторые результаты моделирования.

4.2. Прогнозы по динамической модели и сравнение с наблюдениями. Рассматривалась временная динамика СПН в Дагестане с начала 1998 г. Использовалась модифицированная базовая модель (с учётом реальных «перетоков» между категориями X , Y , Z). Начальное соотношение сторонников и противников позиции « X » (сепаратизма) определялась из данных социологических опросов, в качестве маркирующего события были взяты февральские выборы мэра Махачкалы, начальное значение скорости изменения разницы между числом сторонников и противников (наиболее трудно определяемая стартовая величина подбиралась полуэмпирически).

На рис. 3 показана полученная из вычислительных экспериментов временная динамика СПН (по оси абсцисс отложено время в месяцах, по оси ординат — разница между числом сторонников и противников сепаратизма, номерам на кривой СПН отвечают хорошо известные



- 1 — конец января — первая декада февраля — борьба местных элит за кресло мэра Махачкалы;
- 2 — 21 мая — захват братьями Хачипаевыми здания Госсовета;
- 3 — 21 августа — убийство муфтия;
- 4 — конец сентября — поход на Махачкалу;
- 5 — 21 декабря — убийство директора Центрального рынка Махачкалы;
- 6 — начало боевых действий

Рис. 3. Временная динамика социально-политической напряжённости в Дагестане. По оси абсцисс отложено время в месяцах, по оси ординат — разница в числе сторонников и противников сепаратизма в тыс. человек

события, приходящиеся на очередной период обострения). Первый же прогноз обострения (начало февраля) оказался весьма точным, что позволило не только убедиться в адекватности модели, но и оснастить

её дополнительными реальными данными. В дальнейшем прогнозы по модели также хорошо совпадали с реальными событиями. Динамика СПН в течение почти полутора лет носит циклический характер, повторяющиеся циклы политической самоидентификации имеют стабильный «основной» период равный 120–130 дням. После проведения боевых действий в августе–сентябре 1999 г. динамика СПН заметно изменяется — сторонников сепаратизма становится значительно меньше числа его противников и исчезает ярко выраженная циклическая динамика СПН (амплитуда колебаний уменьшается, периодическая структура становится гораздо менее ощутимой). Другими словами, наступает относительная стабилизация политической обстановки в регионе (сильное внешнее вмешательство перевело систему на другую траекторию). Данный пример хорошо демонстрирует описываемые моделью режимы самоидентификации (см. § 4), применительно к конкретному периоду эволюции дагестанского этноса.

С целью дополнительной проверки и подтверждения гипотез модели динамики группового противоборства сторонников и противников определённого курса действий и получаемых из неё выводов и прогнозов изучалась хронологическая последовательность полученных из различных источников данных о социально-политических конфликтах в регионе. Предполагалось, что такой анализ динамики конфликтов позволил бы выявить временные закономерности поведения социума (в том числе циклы самоидентификации), которые затем могли бы быть сопоставлены с модельными прогнозами. Следует отметить и самостоятельную ценность математической обработки временного событийного ряда для нахождения тенденции в поведении изучаемого объекта (и получения на этой основе хотя бы краткосрочного дополнительного прогноза).

Для решения указанной задачи необходимо выбрать некую единую меру оценки конфликтного поведения социума, т.е. решать задачу в рамках общей проблемы снижения размерности многомерного признака (см. [19] и цитируемую там литературу). Действительно, типичны ситуации, когда приходится сравнивать между собой и упорядочивать по некоторому не поддающемуся непосредственному измерению свойству ряд явлений (в том числе социальных). При этом общее представление о степени проявления анализируемого латентного (т.е. не поддающегося непосредственному измерению) свойства складывается как результат определённого суммирования ряда частных (и поддающихся измерению) характеристик, от которых зависит в конечном счёте это свойство.

В качестве такой меры был принят уровень СПН, понимаемый как интегральный показатель опасности конфликтов, произошедших в регионе в течение дня. Более опасному конфликту присваивалось по методике, изложенной ниже, более высокое численное значение в баллах, а затем эти значения суммировались по всем конфликтам, произошедшим за день (шкалирование).

Методика шкалирования заключается в следующем. Информация о каждом событии (конflikте) структурируется по ряду характеристик. В свою очередь, каждая характеристика разбивается по показателям значимости, каждому из которых присваивается своё значение в баллах. Некоторые характеристики конфликта имеют прямую количественную меру (например, число людей, участвующих в конфликте, или продолжительность конфликта в днях) и шкалирование этих характеристик проводится интервальным образом, т.е. диапазон возможных числовых значений характеристики разбивался на интервалы и каждому интервалу чисел присваивалась своя балльная оценка, тем большая, чем выше границы интервала. Характеристики конфликта, не имеющие прямой количественной меры, подразделялись по показателям значимости методом экспертной оценки.

Для примера в табл. 1 представлена упрощённая схема структуризации k -го конфликта из N произошедших за день только по двум характеристикам: количественной (длительность конфликта в днях) и экспертной (вооружённость конфликта). Предположим, конфликт продолжался 3 дня и в ходе его применялось лёгкое стрелковое оружие. В этом случае в таблице выделяется показатель значимости № 3 для длительности и показатель значимости № 2 для вооружённости.

Таблица 1. Пример упрощённой схемы структуризации и количественной оценки опасности конфликта

Конфликт № k				
№ п/п	Длительность	Баллы	Вооружённость (оружие)	Баллы
1	Несколько часов	1	Невооруженный	1
2	1 день	2	Лёгкое стрелковое оружие	2
3	2–7 дней	3	Автоматическое стрелковое оружие	3
4	>7	4	Тяжёлое стрелковое оружие. Взрывные устройства	4

Соответственно, данному конфликту присваивается 3 балла по длительности и 2 балла по вооружённости. Суммарный показатель опасности — 5 баллов. В данном примере применялась линейная шкала баллов по показателям значимости.

После проведения таким способом структуризации всех конфликтов за день и суммирования их показателей опасности получается интегральный показатель уровня СПН за день.

В табл. 2 представлены характеристики конфликта, учитываемые при реальной обработке массива информации по региону в соответствии с принятой методикой. Примером применения служит многодневный митинг-поход в поддержку местного неформального лидера

(Таблица 3). Суммарный балл (значение уровня СПН) для данного события равен 19 единицам.

Таблица 2. Характеристики конфликта, учитываемые при реальной обработке массивов информации

№ п/п	Уровень конфликта	Причина конфликта	Вооружённость	Массовость	Продолжительность	Локализация	Последствия (число жертв)	Баллы
1	Федеральный	Политическая	Тяжёлое стрелковое	Свыше 1000 чел.	Свыше 1 недели	Межрегиональная	Свыше 100 чел.	4
2	Республиканский	Международная	Автоматическое	100–1000 чел.	2–7 дней	Региональная	10–100 чел.	3
3	Районный	Социально-экономическая	Лёгкое стрелковое	10–100 чел.	1 день	Районная	1–10 чел.	2
4	Локальный	Криминальная	Невооружённый	1–10 чел.	Несколько часов	Локальная	Без жертв	1

Таблица 3. Пример количественной оценки конкретного конфликта

Характеристика конфликта	Показатель значимости	Баллы
Уровень конфликта	Республиканский	3
Причина конфликта	Политическая	4
Вооружённость	Невооружённый	1
Массовость	100–1000 чел.	3
Продолжительность	Свыше 1 недели	4
Локализация	Региональная	3
Последствия (число жертв)	Без жертв	1

Основным источником информации служили ежедневные оперативные сводки региональных управлений МВД. Для целей данного исследования этот источник оказался наиболее значимым (сводки содержат подробное ежедневное описание событий и относительно хорошо структурированы). После их обработки по вышеуказанной методике времени получается временной ряд уровня СПН за определённый период. Графическое представление такого ряда в виде гистограммы приведено на рис. 4 на примере Дагестана. Время наблюдения — с 1 июля 1998 г. по 6 августа 2000 г. По оси абсцисс отложены дни, отсчитанные от 1,07,98 г. (цифра 1 соответствует 1,07,1998 г.). По оси ординат отложено значение уровня СПН в баллах.

Сопоставим полученные результаты с событийным рядом сводок. Зачёрнённый участок на гистограмме над числом 77, показываю-

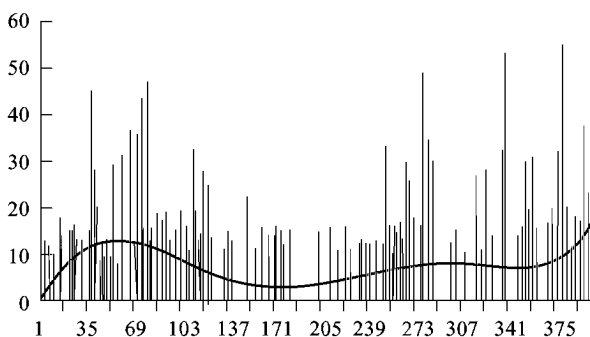


Рис. 4. Временной ряд (гистограмма) уровня социально-политической напряжённости в Дагестане с 1 июля 1998 г. по 6 августа 2000 г. По оси абсцисс отложено время в днях, по оси ординат уровень СПН в баллах

ший высокую временную «плотность» уровня СПН, соответствует многодневному маршу-походу в сентябре 1998 г. группы в несколько сот человек в защиту местного лидера Хачилаева. Одновременно в это же время происходили массовые многодневные выступления шофёров-«дальнобойщиков» с перекрытием автомобильной дороги федерального значения в знак протеста против поборов со стороны таможенников. Участок сгущения (увеличения интегрального уровня СПН) над числом 267 отвечает массовым выступлениям протеста в марте 1999 г. против результатов выборов в Народное собрание республики. Наконец, резкое увеличение уровня СПН в конце рассматриваемого периода отвечает ситуации, складывавшейся в Дагестане перед началом вооружённых действий в августе-сентябре 1999 г.

Роль шкалирования, построения и анализа временных рядов не ограничивается описанием (постфактум) эволюции конфликтов во времени и частичном прогнозировании нарастания их остроты (хотя в сопоставление этих результатов с прогнозами по динамической модели уже весьма важно для их обоснования и повышения их точности). Соответствующие математические процедуры обработки временных рядов более подробно проясняют их внутреннюю структуру — временные тренды (например, сплошной жирной линией на рис. 4 представлен тренд, основанный на аппроксимации ряда полиномом 5-й степени и характеризующий относительно медленные изменения уровня СПН) и различные спектральные характеристики (например, с помощью пакета STADIA и других аналогичных средств определяются частота основных гармоник процесса и их фаза, что позволяет оперативно перенастраивать динамическую модель в соответствии с изменениями ситуации).

Тем самым, комплексное использование различных подходов, моделей, расчётных методик, программных средств и информационных ре-

сурсов даёт возможность моделировать и прогнозировать, фактически в режиме реального времени, протекание реальных этнополитических конфликтов.

§ 5. Некоторые обобщения моделей выбора

5.1. Учёт изменения со временем социально-политического потенциала [17]. В общем случае СПП среднего типического члена электората не является постоянным, а в силу каких-то причин может изменяться со временем. Ограничимся здесь рассмотрением ситуации, когда $P(t)$ является заданной линейной функцией времени, а именно

$$P(t) = P^0 + Ct \geq 0, \quad P^0 \geq 0, \quad t \geq t_0, \quad C = \text{const} \neq 0.$$

Система

$$\frac{dS}{dt} = \frac{P_t - P_Q S}{P_S}, \quad \frac{dQ}{dt} = S$$

является в этом случае также автономной и её уравнения принимают вид

$$\frac{dS}{dt} = \frac{-P_Q S + C}{P_S}, \quad \frac{dQ}{dt} = S.$$

Отсюда следуют уравнения

$$\frac{dS}{dt} = f(S, Q), \quad \frac{dQ}{dt} = S,$$

$$f(S, Q) \equiv \frac{\sqrt{\pi/2} \rho^5}{\chi_2 \mu} \left(-\frac{2\chi_1 \alpha \beta Q}{(Q^2 + \beta)^2} S + C \right) \frac{\exp \frac{S^2}{2\rho^2}}{S(2\rho^2 - S^2)} \quad (43)$$

при $t \geq t_0$ и условии $Q(t_0) = Q^0$, $S(t_0) = S^0$. Эта модель, называемая расширенной базовой моделью электорального поведения, учитывает влияние скорости равномерного изменения СПП на поведение электората в процессе выбора.

Аналогом уравнения

$$\frac{dS}{dQ} = -\frac{g_1(S, Q)}{g_2(S, Q)},$$

где

$$g_1(S, Q) \equiv \sqrt{2\pi} \chi_1 \alpha \beta \rho^5 Q \exp \frac{S^2}{2\rho^2},$$

$$g_2(S, Q) \equiv \chi_2 \mu (Q^2 + \beta)^2 (2\rho^2 - S^2) S$$

является в этом случае уравнение

$$\frac{dS}{dQ} = \frac{f(S, Q)}{S}. \quad (44)$$

Это уравнение «хорошо соответствует» модифицированной автономной системе уравнений

$$\frac{dS}{dt} = S(2\rho^2 - S^2)f(S, Q), \quad \frac{dQ}{dt} = S^2(2\rho^2 - S^2), \quad (45)$$

определённой во всей плоскости S, Q . Система (45) может быть получена из (43) введением на каждом из решений нового времени

$$d\tilde{t} = \frac{dt}{S(2\rho^2 - S^2)}.$$

Уравнение (44) не решается, поэтому построение фазового портрета в этом случае можно проводить лишь численно, а система (45) удобна для непосредственного численного решения.

Из уравнения (44) следует, что фазовый портрет при всех значениях параметров центрально-симметричен относительно начала координат плоскости S, Q (при одновременном изменении знаков Q и S уравнение не изменяется).

Исследуем особые точки уравнения (44). Отметим, что точка $(0, 0)$ для этого уравнения не является особой. Очевидно, что особые точки, если они есть, могут лежать только на прямых $S = \pm\sqrt{2}\rho$. Для нахождения соответствующих значений Q требуется решать алгебраическое уравнение четвёртого порядка, что можно сделать тоже численно. Это уравнение имеет вид

$$(Q^2 + \beta)^2 - p \left(\frac{S_0}{C} \right) Q = 0, \quad (46)$$

где $p = 2\chi_1\alpha\beta > 0$, $S_0 = \pm\sqrt{2}\rho$.

Очевидно, что для существования вещественных корней (46) величина S_0Q/C должна быть положительной. Это обеспечивает следующее расположение особых точек, если они есть: при $C > 0$ особые точки лежат на прямой $S = \sqrt{2}\rho$ в полуплоскости $Q > 0$ и центрально-симметрично им — на прямой $S = -\sqrt{2}\rho$ в полуплоскости $Q < 0$; при $C < 0$ особые точки лежат на прямой $S = \sqrt{2}\rho$ в полуплоскости $Q < 0$ и соответственно симметрично им — на прямой $S = -\sqrt{2}\rho$ в полуплоскости $Q > 0$.

Выведем условие существования особых точек, т.е. условие наличия вещественных корней уравнения (46). Геометрически особые точки — это точки пересечения выпуклой кривой $y(Q) = (Q^2 + \beta)^2$ и прямой $y(Q) = kQ$, где $k = pS_0/C$. Следовательно, их может быть не больше двух. Из условия касания прямой и рассматриваемой кривой имеем $4Q(Q^2 + \beta) - k = 0$. Используя (46), получаем $Q^2 = \beta/3$. Для определённости (что не умаляет общности рассуждений) будем считать величины S_0 , Q , k положительными. Таким образом, при $k = k_0 = 16\beta\sqrt{\beta}/(3\sqrt{3})$ имеется единственный корень уравнения (46) $Q_0 = \sqrt{\beta/3}$ (отметим, что значение Q_0 совпадает с точкой максимума

функции $P_1(Q)$). При $k > k_0$ у этого уравнения имеются два корня Q_1 и Q_2 такие, что $Q_1 < Q_0 < Q_2$. При $k < k_0$ корней нет.

Отсюда, если считать все параметры модели заданными, следует такая зависимость существования особых точек от величины C : при $|C| = C_0 = pS_0/k_0$ имеются две особые точки (Q_0, S_0) и $(-Q_0, -S_0)$; при $|C| < C_0$ — четыре особые точки (Q_1, S_0) , (Q_2, S_0) , $(-Q_1, -S_0)$ и $(-Q_2, -S_0)$; при $|C| > C_0$ особых точек нет.

Таким образом, иллюстрируя влияние введенной в базовую модель величины C (влияние производной от функции $P(t)$) на поведение интегральных кривых в фазовой плоскости (здесь и далее рассматривается, ввиду симметрии, её верхняя полуплоскость), можно сделать следующие выводы. При $|C| = 0$ имеем базовую модель, у которой на прямой $S = \sqrt{2}\rho$ одна особая точка $Q = 0$. С ростом $|C|$, пока сохраняется соотношение $|C| < C_0$, эта точка сдвигается с оси $Q = 0$ вдоль прямой $S = \sqrt{2}\rho$ и, кроме того, на этой прямой возникает ещё одна особая точка (она как бы «приходит» из бесконечности). Когда достигается равенство $|C| = C_0$, эти особые точки сливаются и при дальнейшем росте $|C|$, $|C| > C_0$, пропадают совсем.

Пример. Рассмотрим вариант численного расчёта, который наглядно подтверждает сделанные выводы. При фиксированных параметрах уравнения (46) $\beta = 2$, $p = S_0 = 1$ зависимость Q_1 и Q_2 от величины C имеет вид, приведённый в следующей табл. 4.

Таблица 4. Зависимость величин Q_1 и Q_2 от величины C

C	10^{-8}	10^{-6}	10^{-3}	10^{-2}	0,1	0,1148
Q_1	$4 \cdot 10^{-8}$	$4 \cdot 10^{-6}$	$4 \cdot 10^{-3}$	$4 \cdot 10^{-2}$	0,515	0,8165
Q_2	464,15	99,992	9,865	4,340	1,224	0,8166

При $C > C_0 = 0,1148$ особых точек нет (это расчётное значение C_0 совпадает с полученным по формуле $C_0 = pS_0/k_0$).

Для определения типов особых точек нужно в каждой из них найти собственные значения матрицы вида

$$\begin{vmatrix} \frac{\partial g_1}{\partial S} & \frac{\partial g_1}{\partial Q} \\ \frac{\partial g_2}{\partial S} & \frac{\partial g_2}{\partial Q} \end{vmatrix}$$

для уравнения (44), что сводится к нахождению спектра матрицы

$$\begin{vmatrix} -\frac{pQr}{c} & \left(4Q(Q^2 + \beta) - \frac{pS}{C}\right)r \\ -2S & 0 \end{vmatrix}, \quad r = \frac{\sqrt{\pi/2}\rho^5}{\chi_2\mu} \frac{\exp \frac{S^2}{2\rho^2}}{S^2(Q^2 + \beta)^2}.$$

Фиксируем значение $S = S_0$ и будем исследовать спектр этой матрицы в зависимости от величины C . Опуская громоздкие выкладки, приведём полученные результаты проведённых исследований.

1. При $|C| < C_0$, когда имеются две особые точки с координатами Q_1 и Q_2 , $|Q_1| < |Q_2|$, особая точка с координатой $Q = Q_1$ имеет тип седла (её собственные значения вещественные и такие, что $\lambda_1 < 0 < \lambda_2$) при любых параметрах модели (в этом состоит некая преемственность между базовой и расширенной базовой моделями).

2. Особая точка с координатой $Q = Q_2$ может иметь любой тип (λ_1 и λ_2 могут быть любыми) в зависимости от значений параметров модели. В частности, он может быть вырожденным, так что для исходной нелинейной системы уравнений поведение решений в окрестности этой особой точки (и фазовый портрет в целом) может быть весьма экзотическим. Это же положение имеет место и для «слипшихся» особых точек при $|C| = C_0$.

3. Можно показать, что при любых параметрах модели существует такое значение C^* , $C^* \leq C_0$, что при $|C| < C^*$ особая точка с координатой $Q = Q_2$ имеет тип фокуса ($\text{Re } \lambda_1 = \text{Re } \lambda_2 \neq 0$, $\text{Im } \lambda_1 = -\text{Im } \lambda_2 \neq 0$).

Эти теоретические результаты подтверждены непосредственным численным решением системы (45) при различных начальных данных и некоторых наборах параметров.

Так, в частности, при значениях параметров задачи, взятых в примере, для $C = 0, 1$ (см. таблицу) имеются две особые точки с координатами $(0,515, 1)$ и $(1,224, 1)$. Первая из них является седлом, а вторая — фокусом. Фазовый портрет системы (43) при этих параметрах, составленный на основе расчётов и геометрических соображений, имеет вид, изображённый на рис. 5.

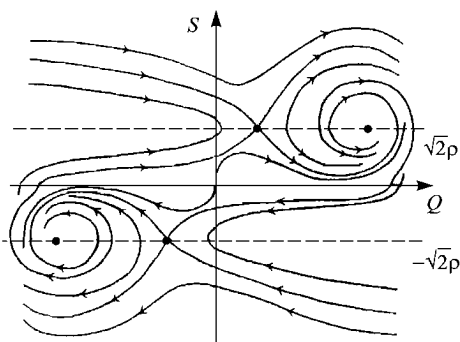


Рис. 5. Фазовый портрет динамической системы (43)

Из рис. 5 видно, что так же, как и в случае базовой модели, для этой модели имеются режимы, описывающие, в частности, поведение в

окрестности фокусов, ведущие себя так, что за конечное время модель приходит в состояние, в котором она не определена.

Этих недостатков лишена модифицированная система уравнений (45), дающая описание более естественных возможных вариантов процессов эволюции. Так, если на рис. 5 расставить стрелки в соответствии со вторым уравнением системы (45) (т.е. считать, что внутри полосы $-\sqrt{2}\rho < S < \sqrt{2}\rho$ величина Q растёт по времени, а вне этой полосы убывает), то нетрудно видеть, что особая точка в первом квадранте плоскости Q, S превращается в закручивающийся (устойчивый) фокус, а в третьем квадранте — в раскручивающийся (неустойчивый) фокус. Отметим возникающие при этом интересные режимы, которые, «сматываясь» с нижнего фокуса, «наматываются» на верхний. Напомним, что этот эскиз сделан при $C > 0$. Если $C < 0$, то получим закручивающийся фокус в четвёртом квадранте фазовой плоскости и раскручивающийся — во втором.

Таким образом, для системы уравнений (45) при ограниченной скорости изменения СПП могут реализовываться колебательные режимы как с растущей, так и с убывающей амплитудой колебаний, причём при $C > 0$ растущие колебания возможны, если $Q < 0, S < 0$, а убывающие — если $Q > 0, S > 0$. При $C < 0$ растущие колебания возникают при $Q < 0, S > 0$, а убывающие — при $Q > 0, S < 0$. Для достаточно быстрого изменения СПП колебательные режимы невозможны.

5.2. Влияние внешнего информационного воздействия. В исходной модели (в отличие от базовой) имеется возможность учитывать влияние внешнего информационного воздействия (см. формулы (20)–(23)). Воспользуемся данной возможностью. Для этого сформулируем функциональную зависимость потенциала от внешнего воздействия.

Будем основываться на следующем представлении о механизме проникновения информации в социум. Некоторая часть социума восприимчива к внешнему воздействию. Впитав в себя новую информацию, эти люди начинают распространять её в социуме так же, как и свою собственную. В результате информация распространяется на весь социум, воздействуя на принятие решений.

Функционально данную зависимость можно записать в виде:

$$P_3(t) = \gamma QF, \quad (47)$$

то есть потенциал пропорционален как величине внешнего информационного воздействия, так и количеству человек, подверженных воздействию (их часть от общего числа Q обозначим как γ). Здесь, как и раньше, Q и F могут быть как положительными, так и отрицательными — отрицательный знак указывает соответственно на преобладание численности сторонников позиции «Y» и на информационное воздействие в поддержку позиции «Y».

В данном разделе будет рассматриваться постоянное информационное воздействие. Это упрощение позволяет понять качественное влияние информационного воздействия на модель в зависимости от величины воздействия, не смешивая это влияние с эффектами, которые могут быть вызваны непостоянством воздействия. Внешнее информационное воздействие, зависящее от времени, имеет смысл рассматривать после рассмотрения постоянного воздействия.

С учётом дополнительного слагаемого, представляющего внешнее информационное воздействие, потенциал выглядит следующим образом:

$$P(Q, S, F) = \chi_1 P_1(Q) + \chi_2 P_2(S) + \chi_3 P_3(Q, F). \quad (48)$$

Подставляя конкретные функциональные зависимости соответственно и предполагая, как и в базовой модели, что потенциал постоянен во времени, получаем уравнения модели:

$$\begin{cases} \frac{dS}{dt} = - \frac{\left(\frac{2\chi_1\alpha\beta Q}{(Q^2 + \beta)^2} + \chi_3\gamma F \right) \rho^5 \exp \frac{S^2}{2\rho^2}}{\chi_2 \sqrt{2/\pi} \mu (2\rho^2 - S^2)}, \\ \frac{dQ}{dt} = S, \end{cases} \quad (49)$$

и соответствующее им уравнение, связывающее Q и S на фазовой плоскости:

$$\frac{dS}{dQ} = - \frac{(2\chi_1\alpha\beta Q + \chi_3\gamma F(Q^2 + \beta)^2) \rho^5 \exp \frac{S^2}{2\rho^2}}{\chi_2 \sqrt{2/\pi} \mu (Q^2 + \beta)^2 (2\rho^2 - S^2) S}. \quad (50)$$

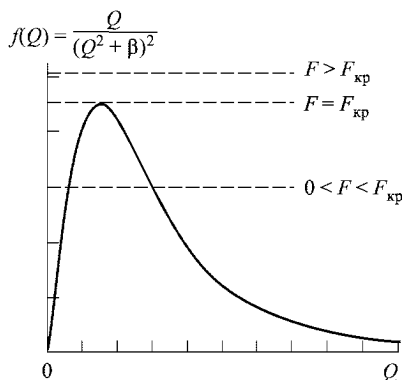
Для исследования модели на фазовой плоскости найдём особые точки системы (38). Сразу следует отметить, что, как и в базовой модели, система не определена на прямых $S^2 = 2\rho^2$. Поэтому точки, лежащие на данных прямых и составляющие равновесие уравнения (50), не могут рассматриваться как особые точки, так как они не подпадают под определение особой точки. Однако траектории, проходящие через такие точки (точнее, уходящие в них или исходящие из них), играют существенную роль в анализе портрета системы на фазовой плоскости. Поэтому они будут исследованы отдельно.

Для нахождения нулей числителя правой части первого уравнения системы (49) необходимо решить следующее уравнение:

$$\frac{Q}{(Q^2 + \beta)^2} = - \frac{\chi_3\gamma F}{2\chi_1\alpha\beta} \quad (51)$$

Правая часть уравнения (51) — константа (горизонтальная прямая на рис. 6), значение которой линейно зависит от F . Отсюда ясно (см. рис. 6), что возможны **три случая**.

1. Если F относительно невелико ($F < F_{кр}$), то уравнение (51) имеет два корня Q_1, Q_2 и у системы существуют две особые точки — $(Q_1, 0)$ и $(Q_2, 0)$.

Рис. 6. График зависимости левой части уравнения (50) от Q

2. При некотором значении $F = F_{\text{кр}}$ существует только один корень Q_3 (имеются ввиду вещественные корни, без учёта кратности) и одна особая точка $(Q_3, 0)$.

3. При больших значениях F ($F > F_{\text{кр}}$) особых точек нет (уравнение (51) не имеет корней).

Исследуем более детально каждый из случаев. Сразу заметим, что, хотя решение уравнения (51) относительно Q (нахождение Q как функции F) проблематично (требуется решить алгебраическое уравнение четвёртого порядка), критическое значение F , соответствующее максимуму левой части уравнения, находится легко и равно

$$F_{\text{кр}} = \frac{3\sqrt{3}\chi_1\alpha}{8\chi_3\gamma\sqrt{\beta}}. \quad (52)$$

Для определения типов особых точек в первом и втором случаях нужно линеаризовать систему (49) вблизи каждого положения равновесия и посмотреть собственные значения матрицы коэффициентов при линейных членах. Если обозначить правые части системы уравнений (49) как Φ_1 и Φ_2 , то эта матрица выглядит следующим образом:

$$\begin{vmatrix} \frac{\partial \Phi_1}{\partial S} & \frac{\partial \Phi_1}{\partial Q} \\ \frac{\partial \Phi_2}{\partial S} & \frac{\partial \Phi_2}{\partial Q} \end{vmatrix} \quad (53)$$

Подставляя конкретные функции системы (49) Φ_1 и Φ_2 :

$$\frac{\partial \Phi_1}{\partial S} = \left(\frac{2\chi_1\alpha\beta Q}{(Q^2 + \beta^2)^2} + \chi_3\gamma F \right) \rho^5 \frac{\partial}{\partial S} \left(\frac{\exp \frac{S^2}{2\rho^2}}{\chi_2 \sqrt{2/\pi} \mu (2\rho^2 - S^2)} \right),$$

$$\frac{\partial \Phi_1}{\partial Q} = \left(\frac{\exp \frac{S^2}{2\rho^2}}{\chi_2 \sqrt{2/\pi} \mu (2\rho^2 - S^2)} \right) \rho^5 \frac{\partial}{\partial Q} \left(\frac{2\chi_1 \alpha \beta Q}{(Q^2 + \beta)^2} + \chi_3 \gamma F \right), \quad (54)$$

$$\frac{\partial \Phi_2}{\partial S} = 1, \quad \frac{\partial \Phi_2}{\partial Q} = 0$$

в матрицу (53) в точках, $(Q_i, 0)$, $i=1, \dots, 3$, получаем уравнение для собственных значений:

$$\det \begin{pmatrix} -\lambda & \left(\frac{\rho^5}{\chi_2 \sqrt{2/\pi} \mu 2\rho^2} \right) \frac{\partial}{\partial Q} \left(\frac{2\chi_1 \alpha \beta Q_i}{(Q_i^2 + \beta)^2} + \chi_3 \gamma F \right) \\ 1 & -\lambda \end{pmatrix} = 0 \quad (55)$$

или (раскрывая определитель)

$$\lambda^2 = \left(\frac{\rho^5}{\chi_2 \sqrt{2/\pi} \mu 2\rho^2} \right) \frac{\partial}{\partial Q} \left(\frac{2\chi_1 \alpha \beta Q_i}{(Q_i^2 + \beta)^2} + \chi_3 \gamma F \right). \quad (56)$$

Обозначим корни этого уравнения как λ_1 и λ_2 . Тогда тип особой точки определяется положительностью/отрицательностью величины $\frac{\partial}{\partial Q} \left(\frac{2\chi_1 \alpha \beta Q_i}{(Q_i^2 + \beta)^2} + \chi_3 \gamma F \right)$, которая наглядно видна из рис. 6.

В результате:

— особая точка $(Q_1, 0)$ имеет тип центр (λ_1, λ_2 — ненулевые чисто мнимые)

— точка $(Q_2, 0)$ имеет тип седла (λ_1, λ_2 — вещественные, ненулевые и разных знаков).

— в точке $(Q_3, 0)$ матрица (53) является вырожденной, поэтому данная особая точка не относится ни к какому типу.

В первом случае (при $0 < F < F_{\text{кр}}$) в системе существуют две особые точки: $(Q_1, 0)$ и $(Q_2, 0)$, где (Q_1, Q_2) — решения уравнения (51). Первая из этих точек является центром, вторая — седлом.

Характерный портрет на фазовой плоскости в этом случае имеет вид, изображённый на рис. 7 а.

Как и в базовой модели, вокруг центра (особой точки типа центр) возникает область колебаний. Однако внешнее воздействие изменяет картину (по сравнению с базовой моделью), деформируя её в сторону воздействия. В итоге у колебаний есть две амплитуды: одна в сторону внешнего воздействия (т. е., численность $(Q - Q_1)$ и воздействие F — одного знака, либо оба положительны, что соответствует позиции «Х», либо оба отрицательны, в сторону позиции «Y») и вторая — в противоположную (когда $(Q - Q_1)$ и F — разных знаков), причём первая заметно превосходит вторую. В некотором смысле это ожидаемый результат — количество людей, занимающих позицию, поддерживаемую внешним воздействием, больше количества людей, позиция которых противоположна внешнему воздействию.

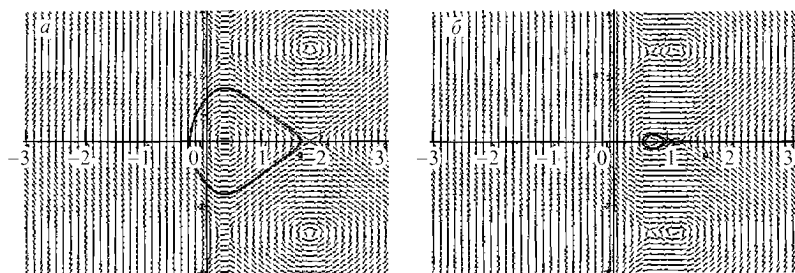


Рис. 7. Фазовый портрет системы (49) при относительно слабом (а) и сильном (б) внешних воздействиях. По осям абсцисс и ординат отложены величины Q и S соответственно

При увеличении величины внешнего воздействия особые точки сближаются (это видно из рис. 6), амплитуды колебаний становятся всё меньше. На рис. 7 б приведён портрет системы с более высоким (по сравнению с рис. 7 а) уровнем внешнего воздействия (но также, как и в прошлом случае, не превосходящего $F_{кр}$):

Из сравнения рисунков 7 а и 7 б видно, как область колебаний уменьшается. Фактически эту область ограничивает сепаратриса, проходящая через седло $(Q_2, 0)$ вокруг центра. Эта траектория определяет как максимальную амплитуду, так и максимальное значение скорости в колебательных режимах.

Вне области колебаний возможны только монотонные режимы эволюции социума. Под монотонными в данном контексте понимаются режимы, в которых в конечном итоге вся социальная общность переходит в какое-то определённое состояние.

Поведение монотонных траекторий удобнее понять, анализируя выражение для потенциала

$$P(Q, S, F) = \chi_1 \frac{\alpha Q^2}{Q^2 + \beta} + \chi_2 \mu \sqrt{\frac{2}{\pi}} \frac{S^2}{\rho^3} \exp\left(-\frac{S^2}{2\rho^2}\right) + \chi_3 \gamma QF = C, \quad (57)$$

где C — постоянное значение потенциала. Выражение (57), очевидно, есть решение системы уравнений (49), поскольку эта система была получена дифференцированием этого выражения (с целью анализа траекторий на фазовой плоскости переменных Q, S).

Зафиксировав C и устремив S к бесконечности или к нулю, получаем, что Q к бесконечности не стремится. Поэтому все монотонные траектории заведомо ограничены по Q .

В случае $0 < F < F_{кр}$ в полосе $S^2 < 2\rho^2$ траектории различаются для областей, на которые эта полоса делится сепаратрисами точек $(Q_2, 0)$, $(Q_1, \sqrt{2}\rho)$ и $(Q_1, -\sqrt{2}\rho)$. Однако все эти траектории так или иначе (см. рис. 6 в) стремятся к точкам (для каждой траектории — своя точка) прямой $S = \pm\sqrt{2}\rho$, обрываясь в этих точках (напомним, что на

этих прямых система не определена). В областях $S > \sqrt{2}\rho$ и $S < -\sqrt{2}\rho$ траектории также стремятся к определённым значениям по оси Q .

При достижении внешним воздействием критического значения $F = F_{кр}$ две особые точки $(Q_1, 0)$ и $(Q_2, 0)$ сливаются в одну — $(Q_3, 0)$ (см. рис. 6). В этом случае, а также при больших значениях величины внешнего воздействия ($F > F_{кр}$) особой точки типа «центр» у системы не существует, и колебаний нет.

Характерный портрет фазовой плоскости при $F > F_{кр}$ приведён на рис. 8. В той области, где при меньших воздействиях ($0 < F < F_p$) происходили колебания, заметны некоторые неоднородности общей картины, связанные с малым значением производной dS/dt . Однако качественно эти траектории ведут себя точно так же, как и все остальные траектории в полосе $S^2 < 2\rho^2$.

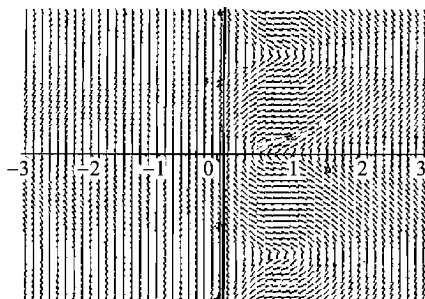


Рис. 8. Фазовый портрет системы (49) при $F > F_{кр}$. По осям абсцисс и ординат отложены величины Q и S соответственно

Таким образом, все монотонные траектории стремятся по оси Q к конкретному конечному значению. Более детальное исследование этих траекторий в рамках данной работы не производится. Более важны (как в рассматриваемой, так и в исходной модели) именно колебательные режимы.

Рассмотрим теперь зависимость максимальной амплитуды колебаний от величины внешнего воздействия. Эта зависимость получается из уравнения (51). Для этого надо выразить больший (по модулю) корень уравнения как функцию от F . Графически эта зависимость изображена на рис. 9.

При очень маленьких значениях F величина Q_2 , ограничивающая амплитуду колебаний в сторону внешнего воздействия, может быть очень большой, бесконечно увеличиваясь с уменьшением F . С увеличением F до критического значения максимальная амплитуда колебаний падает до нулевого значения.

Период колебаний также зависит от величины внешнего воздействия. Для получения этой зависимости линеаризуем правую часть системы (49) вблизи центра $(Q_1, 0)$. Отбрасывая члены более высокого

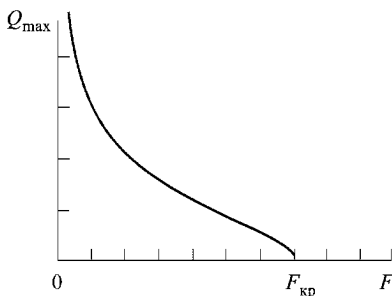


Рис. 9. Амплитуда колебаний как функция внешнего воздействия

порядка малости по Q и S , получаем в окрестности центра уравнение:

$$\frac{d^2 Q}{dt^2} = -(Q - Q_1) \frac{\chi_1 \alpha \beta \rho^3}{\sqrt{2/\pi} \chi_2 \mu} \frac{\beta - 3Q_1^2}{(Q_1^2 + \beta)^3}. \quad (58)$$

Для нахождения периода как функции от величины информационного воздействия (т.е. $T = T(F)$) нужно знать зависимость $Q_1(F)$. Эта зависимость задаётся уравнением (51) и в явном виде не выписывается. Однако с очень хорошей степенью точности её можно приблизить линейной функцией:

$$Q_1 = \frac{8\beta\gamma\chi_3}{9\alpha\chi_1} F = \eta F. \quad (59)$$

С учётом этого для периода колебаний вблизи центра получается формула:

$$T(F) = 2\pi \left(\frac{\sqrt{2/\pi} \chi_2 \mu}{\chi_1 \alpha \beta^3} \right)^{1/2} \frac{((\eta F)^2 + \beta)^{3/2}}{\sqrt{\beta - 3} (\eta F)^2}, \quad (60)$$

переходящая в формулу

$$T_e = 2\pi \sqrt[4]{\frac{2}{\pi}} \sqrt{\frac{\chi_2 \mu \beta}{\chi_1 \alpha \rho^3}}$$

при $F = 0$ (т.е. при отсутствии внешнего воздействия). Положительность подкоренного выражения в знаменателе гарантируется условием $0 < F < F_{\text{кр}}$. График зависимости (60) приведён на рис. 10.

С ростом внешнего воздействия период колебаний возрастает, увеличиваясь при приближении к критическому значению информационного воздействия. Напомним, что амплитуда вблизи точки $F = F_{\text{кр}}$ стремится к нулю, то есть колебания исчезают при приближении к этой точке.

Итак, при увеличении величины внешнего воздействия социум всё меньше склонен к колебаниям, и, если колебания происходят, то они становятся более длительными, то есть члены общности всё менее склонны к изменению своих позиций.

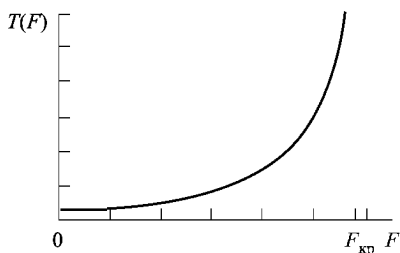


Рис. 10. Период колебаний как функция внешнего воздействия

Представленные в § 2–5 гл. II результаты свидетельствуют о том, что комплексное использование различных подходов, моделей, расчётных методик, программных средств и информационных ресурсов даёт возможность моделировать и прогнозировать, фактически в режиме реального времени, протекание сложных социальных конфликтов, в том числе этнополитических.

Список литературы

1. Самарский А. А., Михайлов А. П. Математическое моделирование (идеи, методы, примеры). — М.: Наука, 1997. — 320 с.
2. Михайлов А. П. Моделирование эволюции распределения власти в государственных иерархиях // Вестник фонда «Российский общественно-политический центр». 1996. № 2. С. 26–39.
3. Mikhailov A. P. Efficient Strategies of Corruption Suppression in State Power Hierarchies // Proceedings of 15th IMACS World Congress, Berlin. 1997. V. III. P. 727–732.
4. Михайлов А. П. Модель коррумпированных властных иерархий // Математическое моделирование. 1999. Т. 11. № 1. С. 3–18.
5. Шведовский В. А. Динамическая модель электорального поведения // Математическое моделирование. 2000. Т. 12. № 8. С. 46–56.
6. Михайлов А. П., Шведовский В. А., Маслов А. Н., Ковалёв В. Ф. Обобщённая модель электорального поведения и её применение к изучению этнополитических конфликтов // Математическое моделирование. 2003. Т. 15, № 8. С. 39–56.
7. Федоренко Н. П., Дементьев В. Е., Гофман К. Г. и др. Общественное богатство: проблемы эффективной реализации экономического потенциала // ЭиММ. Т. 27, вып. 2. 1991. С. 229.
8. Статистический словарь / Под ред. А. И. Ежова. — М.: Статистика, 1965.
9. Комплексная программа научно-технического прогресса / Под ред. акад. В. А. Котельникова. — 1989.
10. Рашевский Н. Две модели: подражательное поведение и распределение статуса // Математические методы в современной буржуазной социологии. — М.: Прогресс, 1966. — С. 175–198.

11. Шведовский В. А. О возникновении случайных флуктуаций в модели подражательного поведения // Вопросы моделирования социально-экономических объектов. — М.: ЦЭМИ АН СССР, 1978. — С. 87–99.
12. Краснощёков П. С. Об одной простейшей модели коллективного поведения. Материалы учредительной конференции Российского научного общества исследования операций. — М.: Вычислительный центр РАН, 1997.
13. Психологические измерения / Под ред. Л. Д. М. Мешалкина. — М.: Мир, 1967. — С. 195.
14. Психофизика дискретных и непрерывных задач / Под ред. Ю. М. Забродина, А. П. Пахомова. — М.: Наука, 1985. — С. 214.
15. Психофизиологические закономерности восприятия и памяти / Отв. ред. д.б.н., проф. А. Н. Лебедев. Институт психологии АН СССР. — М.: Наука, 1985.
16. Данилова Н. Н. Психофизиология. — М.: АСПЕКТ ПРЕСС, 1998. — С. 67.
17. Михайлов А. П., Юхно Л. Ф. Исследование некоторых математических моделей эволюции электората // Дифференциальные уравнения. Т. 23, № 7. 2002.
18. Арутюнян Ю. В., Дробижева Л. М., Сусоколов А. А. Этносоциология. — М.: АСПЕКТ ПРЕСС, 1998.
19. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. — М.: Юнити, 1998.

Глава III

КОМПЬЮТЕРНЫЕ МЕТОДЫ КЛИРИНГОВЫХ РАСЧЁТОВ

§ 1. Денежные расчёты: от курьера до электронной сети

Традиционным способом финансовых расчётов между юридическими лицами является перевод денежных средств банковского счёта одного лица на банковский счёт другого. На Западе этот способ широко используется и при расчётах физических лиц. Для этого: а) первое лицо должно иметь деньги на своём счёте; б) первый банк должен иметь свободные средства; в) банки должны быть связаны какой-то сетью, технически позволяющей производить перечисления денег. Работа этой сети требует определённого времени, так что платёж производится с запаздыванием.

В условиях хорошо отлаженной западной экономики эта система неплохо работает. Лишь при биржевых крахах система сбивается, когда техническая скорость банковских переводов не обеспечивает резко возросший темп расчётов. Тогда биржи закрывают на несколько дней, и за это время успокаивают ситуацию. Но в России эта система была расшатана реформами 1992 г. Тогда возник кризис взаимных неплатежей и продолжается до сих пор, хотя постепенно ослабевает.

В данной главе предложены компьютерные методы клиринговых расчётов, которые позволяют успешно решать эту проблему даже в кризисных ситуациях. Они основаны на создании единого расчётного центра, обслуживающего всю межбанковскую сеть. Этот расчётный центр рассматривает требования всех клиентов, находит способы их взаимного погашения даже при отсутствии свободных средств как у отдельных лиц, так и у банков. Для этого разработаны следующие математические методы.

Во-первых, метод фракционирования долгов. Он сводит задачу взаимного погашения долгов юридических лиц к частному случаю задачи линейного программирования. Для неё был построен итерационный алгоритм, сходящийся тем быстрее, чем запутаннее сетка взаимных долгов. Его скорость такова, что на современных компьютерах можно за минуты решать задачу для миллиона предприятий (подобного быстрого действия нет у известных методов — они применимы лишь

к системам тысячи предприятий). Этот метод опробован на модельных и реальных примерах и позволял на них погасить до 80–90% взаимных долгов.

Во-вторых, метод долгов, имитирующий вексельное обращение. Он позволяет зачесть ещё около половины долгов, остающихся после применения метода фракционирования. Таким образом, остаётся только 5–10% первоначальных долгов.

В третьих, метод банковских проводок. Он основан на фракционировании пакетов платёжных поручений и позволяет проводить текущие расчёты предприятий через банковскую сеть в тех условиях, когда у самих банков мало оборотных средств.

Разработанные методы имеют ещё ряд важных применений. Они позволяют проводить мониторинг экономики страны, отслеживая монополий и выявляя предприятия, скатывающиеся к банкротству. С их помощью можно определить, каким предприятиям целесообразно выдать кредиты для расширения взаимных неплатежей, чтобы инициировать наиболее длинные цепочки выплат. Они позволяют определить разумную стартовую цену при аукционной продаже долгов.

Быстродействие всех предложенных методов таково, что они пригодны для создания общегосударственной системы, работающей в режиме «боевого дежурства» и производящей погашение всех задолженностей раз в сутки, чтобы не допускать накопления долгов. Такая система предотвращала бы большинство кризисов и смягчала оставшиеся.

1.1. Проблемы взаимозачёта.

1.1.1. Денежные расчёты. Человечество до сих пор пользуется древним способом делать покупки: открыть кошелёк, достать деньги и расплатиться. Если кошелёк опустел, то покупки прекращаются до тех пор, пока не удастся снова наполнить кошелёк.

С течением времени этот способ модернизировался. Так, в 18 веке появились бумажные деньги — ассигнации. Они весили много меньше металлических денег и давали возможность носить или возить с собой гораздо большие суммы. По мере развития достаточно густой сети банков с отделениями появились банковские чековые книжки, позволявшие расплачиваться ещё более крупными суммами.

В конце XX века развитие электронной техники позволило связать банки и магазины густой сетью коммуникаций: на этой основе были введены электронные карточки различных банков, получившие широкое распространение, эти модернизации не изменили принципа оплаты. И чековая книжка, и электронная карточка означают, что человек открыл в банке счёт, то есть внёс в банк определённую сумму денег. Это и есть его кошелёк, только он лежит в банке. Если на счёте кончились деньги, то покупки прекращаются до тех пор, пока счёт не будет пополнен.

Однако ещё в средневековой Европе начал вырабатываться другой способ оплаты — клиринговые расчёты. Путешественнику или купцу при поездках в далёкий город или другую страну было невыгодно и даже опасно возить с собой деньги: на дорогах разбойничали, на таможнях грабили (вывоз золота в другую страну нередко облагался пошлиной, а иногда даже воспрещался), а деньги одной страны могли не принимать в другой стране. Поэтому купец вносил нужную сумму денег в банк и получал от него заёмное письмо на эту сумму в банк другого города или страны. Прибыв на место, он получал свои деньги по этому письму.

Для самого купца это был всё тот же принцип кошелька. Новизна состояла в другом. Известные банки городов и стран заключали между собой соглашения. Купцы и путешественники ездили во встречных направлениях. Поэтому в банке А скапливались заёмные письма из банка Б, а в банке Б — заёмные письма из А. Большая часть этих встречных заёмных писем гасила друг друга, и банки открывали свой кошелёк только для расчётов по остатку. Такой клиринговый способ расчёта требовал гораздо меньше наличных денег. В этом их огромное преимущество.

Клиринговые расчёты, в отличие от кошелька, нуждаются в чётко налаженной связи между банками. Современная электронная техника предоставляет такие возможности. Казалось бы, сейчас такие расчёты должны занять основное место в хозяйственной практике. Однако этого не происходит. Далее будет показано, что причина — в отсутствии математических методов, адекватных сложности нынешних задач.

В данной главе предлагаются методы, позволяющие решить наиболее актуальную задачу — взаимозачёт долгов в системе большого числа предприятий, а также ряд примыкающих к ней задач.

1.1.2. Расчёты предприятий. Уже капитализм привёл к появлению большого числа предприятий, связанных сложной системой денежно-хозяйственных отношений. Предприятие покупает сырьё у многих поставщиков, а свою продукцию поставляет многим потребителям. У крупных предприятий число таких связей исчисляется тысячами. Иногда предприятие расплачивается за сырьё немедленно по получении, или даже авансом, но гораздо чаще оно оплачивает поставки лишь спустя некоторое время, определённое договором. Так образуется кредиторская задолженность предприятия (оно должно поставщикам). Если же оно само поставляет свою продукцию в кредит, эти суммы образуют дебиторскую задолженность (ему должны потребители).

Таким образом, предприятие выполняет некоторые банковские функции: выдаёт и получает кредиты. Банки, в смысле выполнения этих функций, также можно называть предприятиями. Государство вступает в хозяйственные отношения — выдаёт оборонный, продуктовый, топливный и другие заказы, а также собирает налоги, акцизы и таможенный платежи. В этом смысле государство и ряд фондов (пенсионный и т. п.) также являются предприятиями. Полное число

предприятий огромно. В России к середине 2003 года было примерно 3 млн действующих предприятий (подававших ненулевые годовые балансовые отчёты в налоговую инспекцию).

Предприятия ведут друг с другом денежные расчёты — отдают долги и получают долги. До сих пор они отдают долги по традиционному способу кошелька. Кошельком являются банковские счета предприятий, на которых хранятся их оборотные средства. Предприятие может оплачивать долги, пока его оборотные средства не иссякнут (это часто называют расплатой «живыми» деньгами). Пополняются оборотные средства за счёт возврата кредитов, выданных другим предприятиям. Если возврат кредитов задержится, предприятие быстро окажется неплатёжеспособным. Часто ли такое будет происходить в действительности?

В нормальной экономике основная масса предприятий должна быть умеренно прибыльной. У таких предприятий платёжное сальдо положительно: полученные кредиты меньше выданных. Считаем, что все предприятия добросовестны, то есть при наличии оборотных средств своевременно выплачивают долги. На первый взгляд кажется, что тогда все предприятия останутся платёжеспособными.

Однако выполнение платежа — это проводка определённых документов через банковскую систему (отправка платёжного поручения в свой банк, связь своего банка с банком получателя через расчётно-кассовый центр, зачисление денег на корреспондентский счёт получателя и уведомление получателя). Эти операции занимают определённое время; даже в лучших банковских системах мира оно составляет 2–3 часа, а в прочих может быть много больше. Вдобавок получатель денег ещё должен решить, какие из своих долгов он должен оплатить полученными деньгами. С учётом этого реальный темп движения денег в долговых цепочках составляет не более одного платежа в рабочий день. Современные цепочки хозяйственных связей предприятий очень сложные, длинные и разветвлённые. Поэтому долг, который предприятие сегодня выплатило, вернётся к нему от какого-нибудь должника лишь пройдя много звеньев цепочки. Это может потребовать несколько недель и даже месяцев. Но предприятие не может стоять, оно должно работать, то есть покупать сырьё и выплачивать зарплату, что приводит к нарастанию его долгов. Если темп нарастания новых долгов окажется выше, чем темп возврата выданных кредитов, то предприятие окажется неплатёжеспособным даже при положительном сальдо.

Таким образом, активная и сложная современная экономика может устойчиво функционировать лишь при достаточно быстром техническом способе выплаты долгов. Если эта операция выполняется медленно, то долги каждого предприятия лавинообразно нарастают и *разражается кризис неплатежей*.

Оценим границу перехода к такому кризису. Пусть время одного платежа составляет τ , среднее число звеньев долговой цепочки равно n . Тогда долги возвращаются предприятию через время $n\tau$. При

нормальной работе предприятие расходует в единицу времени сумму денег Q и чуть больше получает за проданную продукцию. Однако поступление денег задерживается на время nt , и предприятие должно проводить закупки на свои оборотные средства. Таким образом, ему нужны оборотные средства не менее ntQ . Например, если брать $t = 1$ день (современный технический предел) и $n = 25-50$, то требуются оборотные средства не менее одно — двухмесячного полного дохода предприятия. Реально же нужно ещё больше для страховки от неизбежных осложнений. Омертвлять такие крупные средства для предприятия весьма накладно. Но не выделять их опасно — возможно банкротство, за которым последует цепочка банкротств у кредиторов этого предприятия, и начнётся кризис.

Такие кризисы известны давно. Так, в конце 1800-х годов состояние экономики США было вполне благополучным. Но Тихоокеанская железнодорожная компания приостановила платежи. Телеграф уже существовал и мгновенно оповестил всю страну. За считанные часы в стране возникла паника. Все требовали возврата долгов, но не могли выплатить свои долги (техническая скорость выполнения платежа была тогда низкой — требовалось вручить наличные деньги или высоколиквидные ценные бумаги, что занимало один или несколько дней). В итоге обанкротилось много процветающих предприятий. Даже современные скорости банковских расчётов недостаточны для предотвращения таких кризисов. В последнее время на Западе с ними пытаются бороться следующим образом; закрывают все фондовые биржи на несколько дней, а за это время стараются произвести выплаты долгов существующими способами. Это существенно уменьшает число банкротств и финансовые потери, однако они остаются значительными.

Вспомним, что все эти расплаты основаны на методе кошелька. На Западе твёрдо придерживаются убеждения, что платить долги следует «живыми» деньгами. А что, если отказаться от этой традиции и перейти на клиринговые расчёты? В этой главе построены такие математические методы, которые делают это возможным на современной электронной технике.

1.1.3. Кризисы в России. С момента распада СССР в России начался хронический кризис неплатежей с периодическими обострениями. Начало этому было положено в конце 1991 г. Тогда государство, во-первых, прекратило регулировать большинство цен. Во-вторых, возникло множество частных банков вместо нескольких государственных. В-третьих, Госбанк СССР непосредственно перед своей ликвидацией выдал предприятиям большие кредиты (очевидно, такое вливание денег «из воздуха» в экономику усилило инфляцию).

Предприятия немедленно повысили цены. Спрос на их продукцию сразу упал, но полученные кредиты позволили им в первые месяцы удержаться на плаву. Взаимные задолженности быстро накапливались. Положение усугублялось тем, что резко снизилась скорость прохождения платежей. Ранее большинство предприятий держало свои обо-

ротные средства на счетах в Промстройбанке СССР. При этом платёж заключался в переводе денег с одного корреспондентского счёта на другой в Промстройбанке, что выполнялось в тот же день. Теперь же счета предприятий находились во многих банках; платёж стал включать перевод денег из одного банка в другой. Межбанковские связи лишь начинали налаживаться, и техническая скорость платежа стала очень низкой. Платёж даже в пределах одного города мог занимать неделю, а платёж в другой конец страны — месяц.

Поэтому уже к концу первого квартала 1992 года возник острейший кризис неплатежей, полностью парализовавший экономику. Для его ликвидации Центробанк РФ выдал огромные кредиты (хотя много меньше, чем общая сумма задолженностей), причём в основном тем предприятиям, которые надо было спасать в первую очередь. Обстановка разрядилась. Однако уже вторая выдача денег «из воздуха» резко подстегнула инфляцию, которая в 1992 году была 30-кратной (заметим, что выдача таких кредитов очень рискованна, так как большая их часть никогда не возвращается, или возвращается сильно обесценившимися из-за инфляции деньгами).

Выдача кредитов сбила температуру, но не ликвидировала болезнь. Правда, межбанковские связи стали улучшаться, и скорость платежей понемногу возрастала. Однако взаимная задолженность снова нарастала, и периодически возникали очередные кризисы. Каждый раз Правительство и Центробанк РФ ликвидировали их очередной порцией кредитов, подстёгивая инфляцию (в 1993 г. инфляция была 10-кратной, а в 1994 г. — 3-кратной). Тем не менее, в конце 1994 года сообщалось о 50000 предприятий с просроченной задолженностью. По-видимому, речь шла только о предприятиях, которыми интересовалось Правительство, а большим количеством мелких предприятий никто не интересовался.

В последующие годы официальная статистика стала полнее. Так, по официальным данным на 1 июля 1998 г. (накануне дефолта) у 1,5 млн предприятий имелась просроченная задолженность: кредиторская на сумму 1,5 трлн рублей и дебиторская на 0,75 трлн руб. Заметим, что отсюда видна скудность официальной информации: в замкнутой системе суммы кредиторской и дебиторской задолженностей точно равны. Значит, информация была не о всех предприятиях или не о всех их долгах. По-видимому, здесь учтены долги предприятий налоговому ведомству и различным фондам, но «забыты» долги Госбюджета предприятиям за оборонные заказы и т.п. Но даже эта неполная задолженность огромна — в среднем 1 млн руб. (200 тыс. долл. по курсу того дня) долгов у одного предприятия, большинство из которых некрупные.

В последний раз процедура взаимозачёта готовилась в 1999 г. кабинетом Е.М. Примакова. При Министерстве экономики РФ была создана Межведомственная рабочая группа по урегулированию просроченной задолженности. Эта группа сформировала комиссию для сравнения и отбора алгоритмов решения данной задачи. В группу

вошли представители следующих организаций: ФАПСИ, АОЗТ МФД, МИИП, НИИ экономики Минавиапрома, ИММ РАН. Работа группы была прекращена из-за потери интереса Правительства к этой теме. Процедура была плохо автоматизированной, полуручной и полуинтуитивной, очень трудоёмкой. Её подготовка требовала нескольких месяцев, поэтому проводить её можно было не чаще раза в год. Эффект же от её проведения исчезал тоже за несколько месяцев. Это был тупиковый путь и требовалось создание высокоэффективной системы взаимозачёта долгов, причём без выдачи значительных кредитов.

Сейчас проблема по-прежнему остаётся острой. В 2003 году в России насчитывалось 3 млн действующих предприятий, из которых 69% имели просроченную задолженность. В «Российской газете» публиковались большие списки долгов, выставленных на аукционную продажу; среди них были крупные единичные долги (наибольший составлял 172 млн рублей!).

1.2. Клиринговые методы. Выше отмечалось, что даже при современных технических средствах связи платёж «живыми» деньгами фактически занимает рабочий день (так как включает человеческий фактор — принятие решения, какой долг следует заплатить только что полученными деньгами). Поэтому для хорошего решения проблемы взаимозачёта долгов надо отказаться от платежей только живыми? деньгами и включить принципиально другой подход — клиринговые методы. Такие подходы уже предлагались в последнее десятилетие и кое-что применялось.

1.2.1. Долговые цепочки. Рассмотрим идею метода на несложном примере четырёх предприятий А, Б, В, Г (рис. 1). Их долговые связи изображены стрелками, указывающими направление долга. Около каждой стрелки указана величина долга в некоторых единицах. Сумма этих величин, то есть суммарный долг в системе, приведена в табл. 1. Около каждого кружка поставлено сальдо предприятия.

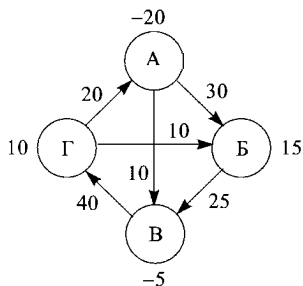


Рис. 1. Схема долгов четырёх предприятий; стрелки с цифрами — долги, около кружков проставлены сальдо предприятий

Идея метода проста. Постараемся отыскать среди предприятий замкнутую цепочку предприятий-должников, в которой все долги направлены в одну сторону. Например, на рис. 1 имеется такая цепочка БВГБ. В ней наименьший $\Gamma \rightarrow \text{Б}$ долг имеет величину 10. На эту величину можно уменьшить все долги в этой цепочке (долг $\Gamma \rightarrow \text{Б}$ становится при этом нулевым). В результате получается схема долгов, показанная на рис. 2а. Суммарный долг в системе при этом уменьшается (табл. 1).

В новой схеме долгов видна ещё цепочка АБВГА. Сокращение её наименьшего звена приводит к рис. 2б. Суммарный долг ещё уменьшается (табл. 1). В полученной схеме видна ещё одна цепочка АВГА. Сокращение её наименьшего звена даёт рис. 2в с ещё меньшей суммой остаточного долга. В последней схеме улучшение уже невозможно — замкнутых цепочек нет.

Таблица 1. Уменьшение суммарного долга при разных способах взаимозачёта

Способ взаимозачёта	Суммарный долг
Исходная система (рис. 1)	135
Долговые цепочки (рис. 2 а,б,в) (рис. 3 а,б)	105; 45; 30 55; 40
Переадресация (рис. 4 а,б); Долговой Центр	25;

Однако процесс исключения цепочек неоднозначен. Например, на рис. 1 видна цепочка АБВГА. Исключение этой цепочки приводит к рис. 3а. На нём видна цепочка БВГБ. Её исключение приводит к рис. 3б и на этом останавливается. Эта схема отлична от рис. 2в, и для неё остаточная сумма долга больше (табл. 1). Такая неоднозначность является очевидным недостатком метода.

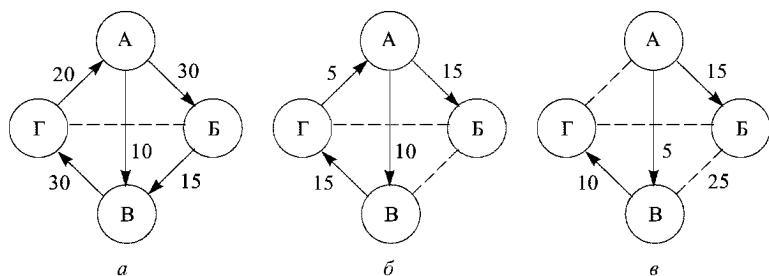


Рис. 2. Последовательное исключение цепочек долгов: а — БВГБ, б — АБВГА, в — АВГА

Ещё неприятнее другой недостаток метода. Если число предприятий есть N , то число возможных долговых связей равно $N(N-1)/2$. Число же всевозможных цепочек растёт быстрее чем e^N , и астрономически велико уже при умеренных N . Поэтому применять метод можно лишь к небольшим группам предприятий, но отнюдь не в масштабах региона, не говоря уже о стране в целом.

Кроме того, далее будет показано, что метод цепочек не минимизирует суммарный долг. На практике он обычно позволяет погасить лишь небольшую часть взаимных долгов.

Однако в середине 1990-х годов проблема неплатежей была настолько острой, что даже этот метод был востребован. Его развивала

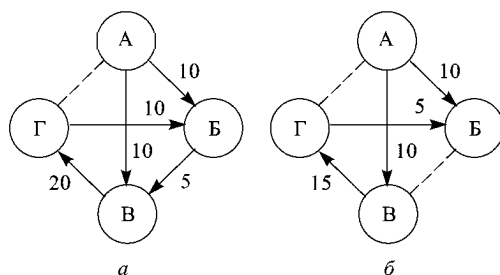


Рис. 3. Иной вариант цепочек: *а* — АБВГА, *б* — БВГБ. Результат хуже, чем на рис. 2

и применяла группа специалистов бывшего НИИ экономики Минавиапрома СССР.

1.2.2. Переадресация. Этот метод развивался авторами с 1995 года [1–4]. Метод моделирует обращение векселей. Если в счёт долга выдан переводной вексель, то он может многократно перейти из рук в руки. Поэтому предъявить его к оплате может любой участник системы, не связанный с исходным предприятием никакой долговой цепочкой.

Здесь очевидно оптимальное решение: предприятия с положительными сальдо собирают все вексели и предъявляют их предприятиям с отрицательными сальдо. Таким образом, у последних остаются только долги, а у первых — только выданные кредиты. Сумма остаточного долга в системе при этом минимальна и равна сумме всех положительных сальдо.

Такое оптимальное решение неоднозначно. Это хорошо видно из рис. 4 *а, б*, где показаны примеры решений для системы рис. 1. Все такие решения имеют одинаковую минимальную сумму остаточного долга (табл. 1). Такая неоднозначность не опасна; наоборот, она предоставляет дополнительные возможности (например, можно оставить связи между предприятиями одного региона или одной отрасли).

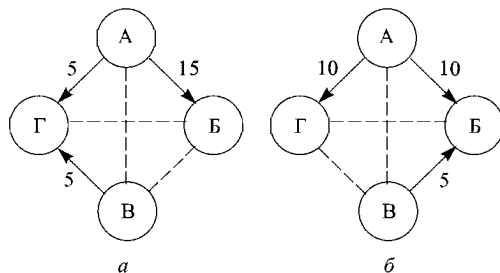


Рис. 4. Метод переадресации для рис. 1; *а, б* — разные варианты. Результат лучше, чем на рис. 2 или рис. 3

Метод переадресации не требует привлечения кредитов. Единственный недостаток метода переадресации — пробелы в законодательстве, не позволяющие считать любой долг переводным векселем.

1.2.3. Долговой центр. Этот метод был предложен Отделением экономики РАН в 1998 году. Его суть в том, что создаётся Центр, который получает от Правительства или Центробанка РФ кредиты, равные суммарному долгу системы. Центр раздаёт их всем предприятиям, имеющим долги. Тогда предприятия уже на другой день могут выплатить все свои долги. Предполагается, что они это сделают, а все суммы, полученные от других предприятий, ещё через день честно вернут Центру. Тогда только предприятия с отрицательными сальдо окажутся должниками Центра, а все остальные кредиты Центр сможет вернуть Правительству.

В идеале остаточный долг минимизируется, как и в методе переадресации. Но даже в этом случае остаточный кредит Правительства остаётся большим, что увеличит инфляцию. Кроме того, рассчитывать на честность всех предприятий наивно. Эти два обстоятельства являются серьёзными недостатками метода долгового Центра.

1.2.4. Фракционирование. Этот метод был предложен авторами в 2001 году [5]. Ниже предложен, обоснован и проиллюстрирован на примерах новый метод погашения взаимных долгов системы предприятий — метод фракционирования. Он сводит проблему к задаче линейного программирования и обеспечивает погашение максимально возможной (без переадресации) доли долгов. Разработан алгоритм настолько малой трудоёмкости, что возможен расчёт для системы $\approx 10^6$ предприятий. При этом метод соответствует нынешнему законодательству и не требует привлечения кредитов. Такой метод пригоден для создания общегосударственной системы финансовых расчётов [5].

Данный подход позволяет создать автоматизированную общегосударственную и региональные системы со столь высоким быстродействием, что взаимозачёт долгов можно будет проводить ежесуточно. Это не позволит долгам накапливаться, принесёт немало выгод каждому предприятию и экономике страны. Попутно такие системы осуществляют мониторинг экономики (например, легко выявляют факты монопольного завышения цен). С их помощью можно также рассчитывать эффективность предоставления кредитов и определять, каким именно предприятиям целесообразно дать эти кредиты.

§ 2. Взаимозачёт долгов методом фракционирования

2.1. Построение метода фракционирования.

2.1.1. Основные формулы.

1. *Постановка задачи.* Пусть имеется система N предприятий; на данный момент n -е предприятие должно некоторым m -м предприятиям суммы $c_{nm} > 0$ (кредиторская задолженность), другие предприятия

должны ему суммы $d_{nm} > 0$ (дебиторская задолженность), а его оборотные средства равны $b_n \geq 0$. Очевидно, $c_{nm} = d_{mn}$.

Введём полную кредиторскую C_n и дебиторскую D_n задолженности предприятия и его сальдо S_n :

$$C_n = \sum_m c_{nm}, \quad D_n = \sum_m d_{nm}, \quad S_n = b_n + D_n - C_n. \quad (1)$$

Внешние связи предприятий не учитываются, ограничиваясь замкнутой системой. Для неё суммарная задолженность

$$C \equiv \sum_n C_n = D \equiv \sum_n D_n. \quad (2)$$

Математически долги описываются матрицей $N \times N$ с элементами a_{nm} , которыми являются величины d_{nm} , $-c_{nm}$ или нули при тех n и m , для которых долги отсутствуют (в больших системах много нулевых долгов). Матрица долгов кососимметрична. Взаимозачёт долгов является такое изменение этой матрицы, которое оставляет её кососимметричной и сохраняет всё сальдо S_n , $1 \leq n \leq N$ [1]. Введём

Определение 1. Взаимозачёт называется разрешённым, если все элементы матрицы долгов сохраняют знаки, а их модули не возрастают.

Это означает, что если между двумя предприятиями долга не было, он не появится; если был, то не увеличится, но может уменьшиться. Разрешённый взаимозачёт соответствует как правилам экономики, так и действующему законодательству: с банковского счёта любого предприятия могут быть в беспорядном порядке списаны средства на уплату его долгов, а за счёт какой именно части его дебиторской задолженности — не ему решать.

Если оборотные средства всех предприятий настолько велики, что все сальдо $S_n \geq 0$, то задача взаимозачёта имеет тривиальное решение: все долги выплачены и $d'_{nm} = 0$, $c'_{nm} = 0$, $b'_n = S_n$. Интересна другая ситуация, когда некоторые $S_n < 0$ и выплата всех долгов невозможна. Нужно найти разрешённый взаимозачёт, минимизирующий остаточную суммарную задолженность C' .

Рассмотрим такой подкласс фракционных взаимозачётов, когда n -е предприятие выплачивает всем своим кредиторам одинаковую долю долгов p_n ($0 \leq p_n \leq 1$). Такой взаимозачёт является разрешённым и приводит к следующим результатам:

$$\begin{aligned} c'_{nm} &= (1 - p_n)c_{nm}, \quad d'_{nm} = (1 - p_m)d_{nm}, \\ C'_n &= (1 - p_n)C_n, \quad D'_n = \sum_m d'_{nm}, \\ b'_n &= S_n - D'_n + C'_n, \quad C' = \sum_n C'_n, \quad D' = \sum_n D'_n. \end{aligned} \quad (3)$$

Согласно правилам экономики, его возможно провести, если конечные оборотные средства каждого предприятия неотрицательны: $b'_n \geq 0$, $1 \leq n \leq N$. Нетрудно показать, что $C' = D'$, так как система остаётся замкнутой.

Определение 2. Назовём фракционный взаимозачёт (3) оптимальным, если он минимизирует остаточную суммарную задолженность C' .

Нахождение оптимального фракционного взаимозачёта сводится к следующей задаче линейного программирования с неизвестными p_n , $1 \leq n \leq N$:

$$C - C' \equiv \sum_n p_n C_n = \max, \quad (4a)$$

$$b'_n \equiv b_n - p_n C_n + \sum_m p_m d_{nm} \geq 0, \quad (4б)$$

$$0 \leq p_n \leq 1. \quad (4в)$$

Эффективность такого взаимозачёта удобно измерять коэффициентом погашения задолженности

$$k_0 = 1 - \frac{C'}{C}, \quad 0 \leq k_0 \leq 1. \quad (5)$$

Сравним его с аналогичным коэффициентом для идеального взаимозачёта методом переадресации долгов [2]:

$$k_n = 1 - \sum_n \frac{|S_n| - S_n}{2C}, \quad k_0 \leq k_n \leq 1. \quad (6)$$

Метод переадресации выгоднее, однако при существующем законодательстве не может применяться. Интересно, сильно ли проигрывает фракционный метод.

З а м е ч а н и е. Можно рассмотреть наиболее общий случай экстремального взаимозачёта с $N(N-1)$ переменными p_{nm} ($n \neq m$). В нём коэффициент погашения k_3 больше и лежит в пределах $k_n \leq k_3 \leq k_0$. Но такая постановка вряд ли целесообразна по экономическим соображениям: возможны решения, при которых все деньги пойдут на выплаты крупных долгов, а мелкие кредиторы окажутся несправедливо обделёнными.

2. Решение. Неравенства (4в) описывают единичный куб в N -мерном пространстве вероятностей. Каждое неравенство (4б) описывает полупространство, ограниченное n -й гиперплоскостью. Гиперплоскость *ограничений* пересекает положительную полуось n -й координаты в точке

$$\tilde{p}_n = \frac{b_n}{C_n} \geq 0 \quad (7)$$

(возможно даже $\tilde{p}_n > 1$); для всех остальных координат она пересекает отрицательные полуоси (рис. 5 для случая $N = 2$ и рис. 6 для $N = 3$). Видно, что эти гиперплоскости вырезают из единичного куба непустой и конечный многогранник допустимых решений p_n . Если все сальдо

неотрицательны $S_n \geq 0$, то вершина единичного куба ($p_n = 1, 1 \leq n \leq N$) принадлежит этому многограннику; если хотя бы одно $S_n < 0$, эта вершина лежит вне многогранника допустимых решений.

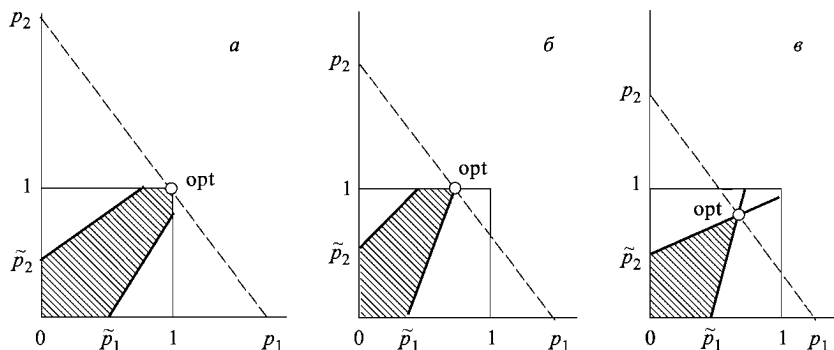


Рис. 5. Случай $N = 2$: сплошные линии — гиперплоскости ограничений, штриховая линия — гиперплоскость экстремума, заштрихованы многогранники допустимых решений, кружками показаны оптимальные решения. Случай *а* соответствует положительности всех сальдо, в случае *б* часть сальдо отрицательны, в случае *в* все сальдо отрицательны

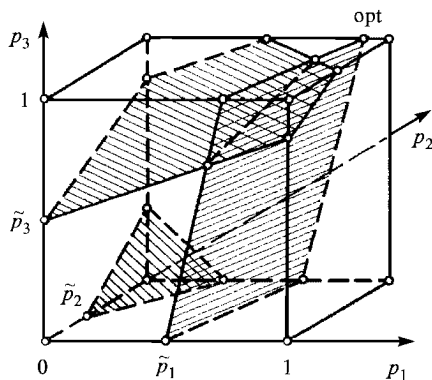


Рис. 6. Случай $N = 3$ (здесь часть сальдо положительны, что соответствует рис. 5 б); заштрихованы гиперплоскости ограничений (4б), показано оптимальное решение; гиперплоскость экстремума не изображена

Гиперплоскость *экстремума*, соответствующая (4а), при $C' = C$ проходит через начало координат. При уменьшении C' она параллельно перемещается, удаляясь от начала координат так, что пересекает только положительные полуоси. Геометрически очевидно, что при $C' \approx C$ эта гиперплоскость пересекает многогранник допустимых решений, но при достаточном уменьшении C' выходит из него. Следовательно,

имеется наименьшее C' , при котором гиперплоскость (4а) имеет общую точку с этим многогранником. Значит справедливо

Утверждение. Оптимальный фракционный взаимозачёт долгов (4) существует.

Оно является частным случаем известной теоремы о существовании конечного решения задачи линейного программирования в ограниченной области.

Очевидно, таким решением является одна из вершин многогранника допустимых решений. Однако, не будет ли экстремальное значение C' достигаться на ребре или грани этого многогранника, то есть будет ли оптимальное решение единственным? Имеется тривиальный случай такой неединственности. Пусть некоторое l -е предприятие не имеет долгов: $C_l = 0$, $c_{lm} = 0$ при всех m . Тогда любое значение $0 \leq p_l \leq 1$ ни на что не влияет. Значит, если найти некоторое оптимальное решение $\{p_n, 1 \leq n \leq N\}$ и изменять в нём значение p_l , полученное решение также оптимально. Одно предприятие без долгов, то есть изменение одного параметра p_l , соответствует ребру многогранника; r таких предприятий соответствуют r -мерной плоской грани.

Отбросим эти тривиальные случаи и рассмотрим углы наклона гиперплоскости экстремума к гиперплоскостям ограничений и граням единичного куба (рис. 5 и 6). Можно установить следующую теорему.

Теорема. Нетривиальная часть оптимального решения единственна.

Для нахождения этого решения перепишем уравнение (4б) в следующем виде:

$$p_n \leq \frac{1}{C_n} \left(b_n + \sum_m p_m d_{nm} \right). \quad (8)$$

Учтём также (4в) и построим следующий итерационный процесс:

$$p_n^{(0)} = 1, p_n^{(q)} = \min \left[1, \frac{1}{C_n} \left(b_n + \sum_m p_m^{(q-1)} d_{nm} \right) \right] \quad \text{при } C_n > 0, \quad (9)$$

$$p_n^{(q)} = 1 \quad \text{при } C_n = 0 \quad \text{для } q = 1, 2, \dots$$

При этом из тривиальной неединственности выбирается решение, в котором нулевые долги выплачиваются с вероятностью 1. Трудоемкость одной итерации не превышает $\approx 2N^2$ операций. Фактически она много меньше, ибо в матрице долгов большинство элементов — нули, и их не надо включать в сумму по m . Алгоритм (9) прост и легко программируется.

3. Сходимость. Процесс (9) относится к классу простых итераций $x^{(q)} = \varphi(x^{(q-1)})$ с непрерывной и кусочно-гладкой правой частью. Из (9) видно, что элементы матрицы Якоби $\partial \varphi / \partial x$ равны

$$\frac{\partial p_n^{(q)}}{\partial p_m^{(q-1)}} = 0 \quad \text{или} \quad = \frac{d_{nm}}{C_n} \geq 0. \quad (10)$$

Для линейной сходимости процесса (9) достаточно, чтобы какая-нибудь норма этой матрицы была меньше 1. Знаменатель этой сходимости $\varkappa$ не превышает наименьшей из норм матрицы — спектральной. Спектр найти практически невозможно, но нетрудно мажорировать эту норму другими, например евклидовой (сферической). С учётом (10) это даёт оценку

$$0 < \varkappa \leq \left\| \frac{\partial \varphi}{\partial x} \right\|_{\lambda} \leq \left\| \frac{\partial \varphi}{\partial x} \right\|_E \leq \left[\sum_{n,m} \left(\frac{\partial p_n^{(q)}}{\partial p_m^{(q-1)}} \right)^2 \right]^{1/2} \leq \left(\sum_{n,m} \frac{d_{nm}^2}{C_n^2} \right)^{1/2}. \quad (11)$$

На практике обычно эти неравенства далеки от равенств, так что оценка $\varkappa$ по конечной сумме цепочки (11) оказывается существенно завышенной. Фактическая скорость сходимости оказалась очень высокой на модельных примерах для системы предприятий с большим числом взаимных связей. Зачастую 10 итераций давали 6 верных знаков. Это на порядки быстрее, чем классические методы решения общей задачи линейного программирования. Для системы почти не связанных предприятий скорость сходимости существенно замедлялась, однако трудоёмкость расчётов оставалась малой.

2.1.2. Примеры. Реальная экономическая информация труднодоступна. Поэтому свойства метода фракционирования были подробно изучены на модели системы предприятий, среди которых нет монополистов. Последнее означает, что разница в финансовом благополучии предприятий носит чисто случайный характер.

1. Модельный пример. Финансовая «мощность» предприятий, то есть характерные величины денежных сумм их оборотных средств и долговых обязательств, описывалась весовой функцией

$$\rho(n) = \sqrt{\frac{N}{N-n+1}}, \quad \rho(1) = 1, \quad \rho(N) = \sqrt{N}; \quad (12a)$$

при этом малых предприятий много больше, чем крупных, а финансовая мощность самого крупного предприятия в $\sqrt{N}$ раз больше, чем у малого; это качественно описывает реальную ситуацию. Верхняя половина матрицы долгов формировалась с помощью псевдослучайных чисел $\{\gamma_j\}$ по формуле:

$$a_{nm} = \gamma_j \rho(n) \cdot \begin{cases} 1 & \text{при } \gamma_{j+1} > 1 - N^{-1/2}, \\ 0 & \text{при } N^{-1/2} < \gamma_{j+1} < 1 - N^{-1/2}, \\ -1 & \text{при } \gamma_{j+1} < N^{-1/2}. \end{cases} \quad (12б)$$

Положительные элементы становились дебиторскими задолженностями d_{nm} , модули отрицательных — кредиторскими c_{nm} ; число ненулевых элементов в строке, то есть связей одного предприятия с другими, составляло $\sim \sqrt{N}$. Нижняя половина матрицы a_{nm} дополнялась кососимметрично. Для экономического правдоподобия такого способа

построения матрицы долгов существенно то, что предприятия упорядочены по возрастанию мощности.

Оборотные средства определялись по аналогичной формуле:

$$b_n = bN^{1/2}\rho(n)\gamma_j. \quad (12в)$$

Здесь $b \geq 0$ коэффициент, характеризующий общее состояние экономики: малые значения b соответствуют дефициту оборотных средств. Варьируя N и b , можно провести исследование как свойств предложенного метода, так и состояния системы.

Результат конкретного расчёта зависит от выбора отрезка последовательности $\{\gamma_j\}$. Поэтому при каждом N , b расчёт повторялся 10 раз с различными неперекрывающимися участками последовательности. Затем вычислялись средние значения и дисперсии исследуемых величин. При увеличении N дисперсии довольно быстро убывали; при $N \geq 40$ они становились незначительными.

2. *Результаты.* Расчёты проводились в диапазонах $0 \leq b \leq 1,6$ и $5 \leq N \leq 640$; последнее соответствует количеству предприятий в небольшом районе. При этом уже чётко выявляются основные закономерности, так что дальнейшего увеличения N не требуется. На рис. 7 и 8 представлены важнейшие результаты — коэффициенты k_0 , k_n (для удобства в процентах) и число итераций q , нужное для сходимости с точностью 10^{-6} . Различными значками представлены усреднённые данные расчётов, линиями — их аппроксимации.

Наиболее интересный результат — быстрая сходимость. Даже при $b = 0$ (отсутствие оборотных средств) достаточно $q < 100$ итераций; при увеличении b число итераций уменьшается $\sim 1 + \text{const}/b$ (рис. 7), быстро стремясь к 1. При небольших $b > 0$ зависимость $q(N)$ логарифмически слабая, и обычно достаточно ≈ 10 итераций. Это обеспечивает быстрое действие метода: даже если матрица долгов плотно заполнена, общий объём вычислений составит всего $\approx 10N^2$ операций (в общих задачах линейного программирования это от $\sim N^3$ до $\sim N^4$). Для разреженной матрицы долгов объём вычислений пропорционально уменьшается. Поэтому для реальных экономических задач он вряд ли превысит $10N^{3/2}$.

При увеличении N коэффициенты погашения долгов k_0 , $k_n \rightarrow 1$ довольно быстро (рис. 8). Закон стремления близок к $1 - \text{const} \cdot N^{-\alpha}$, $\alpha \approx 0,3-0,5$. Разница между k_0 и k_n становится небольшой уже при $N \geq 40$, так что оптимальный фракционный взаимозачёт лишь немного уступает идеальному (переадресации долгов) в практически интересных случаях. Разница между ними значительна только при очень малых $N \leq 10$ (рис. 8), что неинтересно, или при $b \approx 0$ (рис. 7), что соответствует полному развалу экономики.

Разумеется, эти результаты относятся к примерам рассмотренного типа. Если среди предприятий окажутся монополисты, взвинчивающие цены, то матрицу долгов нельзя строить чисто случайно по (12). Од-

нако в такой ситуации даже значение k_n может оказаться небольшим, так что фракционный метод и здесь полезен.

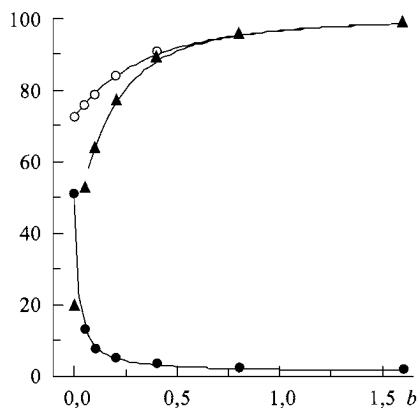


Рис. 7. Расчёты при $N = 20$: q — ●, k_n (%) — ○, k_o (%) — ▲; линии — аппроксимации

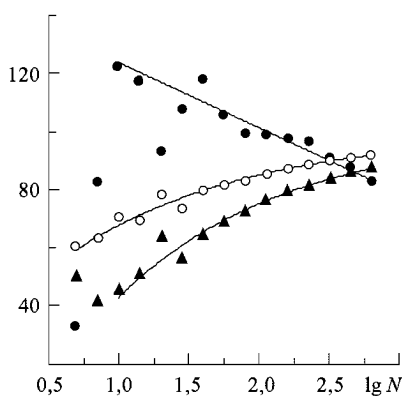


Рис. 8. Расчёты при $b = 0,1$: $10q$ — ●, k_n (%) — ○, k_o (%) — ▲; линии — аппроксимации

3. Вариант модели. Пример (12) имеет один качественный недостаток: согласно (12а), число долговых связей $\approx 2N^{-1/2}$ одинаково у всех предприятий, как крупных, так и мелких. На самом деле крупные предприятия имеют много больше связей. Это можно моделировать следующей матрицей долгов:

$$a_{nm} = \gamma_j \rho(n) \cdot \begin{cases} 1 & \text{при } \gamma_{j+1} > 1 - \alpha \rho(n)/N, \\ 0 & \text{при } \alpha \rho(n)/N < \gamma_{j+1} < 1 - \alpha \rho(n)/N, \\ -1 & \text{при } \gamma_{j+1} < \alpha \rho(n)/N, \end{cases} \quad (13)$$

при тех же весовой функции (12а) и оборотных средствах (12в). Это означает, что наименьшие предприятия имеют в среднем по α дебиторских и кредиторских задолженностей, а наибольшее — по $\alpha N^{1/2}$; разумны значения $\alpha \approx 1-2$. Данная модель лучше, чем предыдущая, описывает большую систему предприятий. Расчётов по данной модели не производилось. Однако и эта, и предыдущая модель представляют самостоятельный интерес как тесты для сравнения различных методов клиринговых расчётов.

4. Пример региона. В ходе подготовки взаимозачёта долгов по России в 1999 г. различные алгоритмы тестировались на реальных списках задолженностей, составленных для небольшого числа предприятий Башкирии по состоянию на 1997 г. Этот пример подробно рассмотрен ниже. Тогда тестирование не было доведено до конца из-за отмены решения о взаимозачёте. Предварительные результаты оказались нечёткими и не были серьёзно проанализированы.

Сейчас вернулись к тому списку, который содержал наибольшее число предприятий, подавших сведения о своих долгах, и детально изучили его. Результаты оказались неожиданными и интересными.

Сведения подали $N = 120$ предприятий региона. Суммарный список состоял из 13 584 отдельных обязательств; на одно исходное предприятие приходилось в среднем 113 обязательств. Среди кредиторов и дебиторов были не только исходные 120 предприятий, но более 7600 новых из различных регионов. Другие предприятия не подавали списков своих долгов, так что для них классический взаимозачёт невозможен (применим лишь метод переадресации). Поэтому из общего списка был выделен список *внутренних* долгов, когда и кредитор и дебитор были исходными предприятиями. Этот список содержал 1040 кредиторских и 580 дебиторских задолженностей.

Но исходные предприятия и их долги образуют замкнутую систему. В ней число кредиторских и дебиторских записей должно совпадать сами денежные величины должны удовлетворять соотношениям $c_{nm} = d_{mn}$, а суммарное сальдо всех предприятий должно обращаться в нуль. Всё это не выполняется (см. табл. 1, где исходные $C \neq D$). Такого материала совершенно недостаточно для взаимозачёта долгов. Так что сравнения методов в 1999 г. и не могли ничего выявить; надо было сначала привести в порядок исходный материал.

Первоначально были опасения в достоверности, но некоторая часть перекрёстных обязательств совпадала. Это указывало, что предприятия дали достоверные записи, но не полные, а частичные. Кроме того, между одной и той же парой предприятий в ряде случаев было несколько обязательств.

Поэтому, во-первых, были просуммированы все обязательства каждой пары; во-вторых, дополнен список, считая бухгалтерии предприятий добросовестными, но неаккуратными. Если о долговом обязательстве сообщала только одна сторона, оно включалось в список. Если сообщали обе стороны, но суммы не совпадали, бралась большая из сумм. После этого список составил по 444 строки как кредиторской, так и дебиторской задолженности при совпадении $c_{nm} = d_{mn}$. При этом у 9 предприятий не оказалось никаких внутренних обязательств, так что этот список фактически включал 111 предприятий — в среднем по 4 связи у каждого предприятия.

5. *Расчёт.* Результаты расчёта представлены в табл. 2. Из неё видно, что в данном реальном примере метод фракционирования позволил погасить 80% долгов системы. Ещё примерно 42% оставшейся части долгов (т. е. 8% от начальной величины) гасится методом переадресации. Это доводит полный коэффициент взаимозачёта до 88%, в то время как методы нахождения цепочек во время тестирования в 1999 г. давали лишь от 2% до 11%.

Более тщательный анализ показал ещё одну трудность. Среди погашенных долгов основную часть составили встречные долги 62 пар предприятий друг другу (заметим, что эти встречные долги не связаны

Таблица 2. Взаимозачёт фракционированием и переадресацией для 120 предприятий Башкирии по 1997 г.; суммы даны в неденоминированных рублях

Суммарная внутренняя задолженность по начальной неполной информации	кредиторская C	28044,1 млн руб.
	дебиторская D	26713,8 млн руб.
По дополненной информации $C' = D$		52477,0 млн руб.
Из них встречных долгов пар предприятий, в сумме		20982,3 млн руб.
Остаток долгов после взаимозачёта	Фракционированием	10512,5 млн руб.
	Переадресацией	6115,0 млн руб.
Коэффициент взаимозачёта	Фракционированием	80,0 %
	Переадресацией	88,3 %

с нашим дополнением списка: они уже содержались в исходном списке из 13584 строк). Эти долги их бухгалтерии могли бы обнаружить и взаимно погасить сами, не обращаясь к региональному центру. После такого погашения список нетривиальных долгов сократился до 382 кредиторских и столько же дебиторских записей, в среднем по 3,5 записи на одно предприятие. Значительная часть этих записей содержала небольшие денежные суммы. В результате возможные нетривиальные цепочки быстро обрывались и гасили лишь небольшую часть оставшегося долга.

Причина этого понятна. У мелких предприятий много связей замыкается внутри региона, и для них региональный взаимозачёт эффективен в процентном отношении; но денежные суммы при этом невелики. У крупных предприятий и суммы крупные, но основная часть связей лежит вне региона; региональный центр их не может обрабатывать. Поэтому чем больше охваченный регион, тем эффективней взаимозачёт. Наиболее эффективен он в масштабе всей страны.

2.2. Программа и расчёты.

2.2.1. Проблема. Традиционным методом взаимозачёта долгов в системе большого числа юридических лиц (далее - *предприятий*) является отыскание замкнутых цепочек долгов и списание наименьшего из долгов со всех предприятий цепочки. Это юридически безукоризненный способ, но он трудоёмок, а эффективность его мала: удаётся погасить лишь малую часть долгов.

Наибольшую эффективность, соответствующую теоретическому пределу, даёт предложенный нами ранее метод переадресации долгов. Он во многом имитирует обращение переводных векселей. Но он не вполне этому эквивалентен, и поэтому не укладывается в существующее законодательство: хотя предприятие не вправе

возразить против перепродажи своей кредиторской задолженности, но только оно вправе распоряжаться своей дебиторской задолженностью.

Поэтому очень полезным был бы метод промежуточного типа: юридически правомерный, но достаточно эффективный и быстрый. Здесь разработан такой метод, названный методом фракционирования долгов. Он позволяет погасить значительную часть взаимных долгов, а по быстрдействию пригоден даже для системы из миллиона предприятий, то есть в масштабе всей страны.

2.2.2. Обозначения. Пусть для системы N предприятий собрана информация о взаимных долгах и, возможно, о других аспектах их финансового положения. Введём следующие обозначения исходной информации:

n, m — индексы предприятий, $1 \leq (n, m) \leq N$;

c_{nm} — кредиторская задолженность n -го предприятия m -му предприятию;

d_{nm} — аналогичная дебиторская задолженность (очевидно, $d_{mn} = c_{nm}$);

b_n — остаток оборотных средств на счёте n -го предприятия;

B_n — те средства, которые предприятие может быстро мобилизовать при угрозе банкротства (сюда входят его оборотные средства, срочная продажа неосновных фондов и т.п.; но в них нельзя включать продажу его дебиторской задолженности предприятиям данной системы).

2.2.3. Балансы. Введём основные балансные величины для каждого предприятия и всей системы:

$c_n = \sum_m c_{nm}$ — полная кредиторская задолженность предприятия;

$d_n = \sum_m d_{nm}$ — полная дебиторская задолженность предприятия;

$s_n = b_n + d_n - c_n$ — сальдо предприятия (с учётом остатка оборотных средств);

$C = \sum_n c_n$ — кредиторская задолженность всей системы;

$D = \sum_n d_n$ — дебиторская задолженность всей системы (очевидно, $D = C$);

$S = \sum_n s_n = \sum_n b_n$ — сальдо всей системы;

$S_a = \sum_{s_n < 0} s_n$ — суммарное сальдо предприятий «чистых» должников ($S_a \leq 0$);

$k_a = 1 + S_a/C$ — коэффициент идеального взаимозачёта с перерасацией долгов ($0 \leq k_a \leq 1$); это наибольшее значение, которое можно получить при взаимозачёте без привлечения дополнительных средств (займов или продаж).

Если хотя бы у одного предприятия сальдо $s_n < 0$, оно не сможет полностью расплатиться со своими кредиторами, даже если все его

дебиторы вернут ему долги. Это может породить каскад неуплат. Даже часть предприятий с положительным сальдо, не получив долги со своих дебиторов, окажется не в состоянии расплатиться с кредиторами.

2.2.4. Фракционирование долгов. В нашем предыдущем расчёте искались вероятности p_n того, что n -ые предприятия оплатят свои долги. Поставим вопрос иначе. Будем считать, что p_n — та доля своего долга, которую n -ое предприятие в состоянии выплатить ($0 \leq p_n \leq 1$); дополнительно предположим, что каждому своему кредитору оно выплачивает одну и ту же долю, не делая предпочтений. Тогда после этого взаимозачёта получим такое конечное состояние системы:

$\tilde{c}_{nm} = (1 - p_n)c_{nm}$ — кредиторская задолженность n -го предприятия m -му;

$\tilde{d}_{nm} = (1 - p_m)d_{nm}$ — аналогичная дебиторская задолженность;

$\tilde{c}_n = (1 - p_n)c_n$ — полная кредиторская задолженность предприятия;

$\tilde{d}_n = \sum_m (1 - p_m)d_{nm}$ — полная дебиторская задолженность предприятия;

$\tilde{b}_n = s_n - \tilde{d}_n + \tilde{c}_n$ — остаток оборотных средств предприятия (поскольку сальдо каждого предприятия сохраняется при взаимозачёте);

$\tilde{C} = \sum_n \tilde{c}_n$, $\tilde{D} = \sum_n \tilde{d}_n$ — окончательная кредиторская и дебиторская задолженности в системе; теперь $\tilde{C} \neq \tilde{D}$, поскольку при взаимозачёте могли расходоваться оборотные средства b_n ;

$k_f = 1 - \tilde{C}/C$ — коэффициент фракционного взаимозачёта ($0 \leq k_f \leq k_a \leq 1$).

Взаимозачёт *допустим*, если после его выполнения остатки оборотных средств всех предприятий неотрицательны: $\tilde{b}_n \geq 0$. Взаимозачёт *оптимален*, если его коэффициент максимален $k_f = \max$. Преобразуя эти условия с помощью балансных соотношений, приведённых выше, сведём проблему оптимального фракционного взаимозачёта к задаче линейного программирования:

$$\sum_{n=1}^N p_n c_n = \max, \quad 0 \leq p_n \leq 1$$

и

$$b_n - p_n c_n + \sum_{m=1}^N p_m d_{nm} \geq 0, \quad 1 \leq n \leq N.$$

Она в точности совпадает с задачей определения надёжности долгов, сформулированной и решённой в п. 1.1.1. Разработанный там алгоритм является очень быстрым, и даёт искомые доли p_n за число арифметических действий $\sim N$.

2.2.5. Программа и выдача. Программа DEFRA (debts fractioning) была дополнена расчётом итоговой матрицы долгов и балансов.

Выдача модернизированной программы содержит все ранее выдававшиеся величины, а также итоговые балансы и итоговую матрицу долгов. В зависимости от целей расчёта предусмотрено три варианта выдачи.

1°. Короткая. Выдаются только две строки:

- Начальные данные расчёта — число предприятий N , условная доля оборотных средств b , точность итераций $\varepsilon \ll 1$, ограничение числа итераций Q , начальный номер J для выбора псевдослучайных чисел при формировании матрицы долгов в методических расчётах.

- Полные балансы всей системы — число итераций q , коэффициенты идеального k_a и фракционного k_f взаимозачётов, сальдо системы S , суммарное сальдо «чистых» должников S_a , начальная C и конечная C_f суммы долгов в системе, начальная D и конечная D_f суммы кредитов, средняя доля уплаты долгов P_q .

Данных о состоянии отдельных предприятий в этой выдаче нет. Когда делают методические расчёты с $N > 50 \div 100$, целесообразно оставлять только короткую выдачу.

2°. Средняя. Она включает короткую выдачу и таблицу со сводными данными об исходном и конечном состоянии отдельных предприятий. Такую выдачу целесообразно делать в расчётах с $N \approx 15 \div 50$.

3°. Длинная. Она включает среднюю выдачу и ещё две таблицы. Это матрицы начальных C_{nm} и конечных $C_{f, nm}$ долгов. В них отрицательные числа — кредиторские задолженности, положительные — дебиторские (эти матрицы хорошо ложатся на лист бумаги только при $N \leq 14$).

Выдачи чётко иллюстрируют юридически бесспорный характер фракционного взаимозачёта: если некоторой долговой связи не было ($c_{nm} = 0$), то она не появляется ($\tilde{c}_{nm} = 0$); любой долг c_{nm} сохраняет свой знак и убывает по модулю (причём может стать нулём). Следовательно, переадресации долгов нет. Заметим, что у предприятий с $p_n = 1$ все величины $\tilde{c}_{nm} \geq 0$, то есть нет конечных кредиторских задолженностей.

Быстродействие. Расчёты выполнялись на маломощном РС-486. Поскольку программа была методической, её детали не «вылизывались»: проводилось суммирование по всем долгам, включая нулевые, считалось до $10N^2$ псевдослучайных чисел для формирования матрицы долгов и т. п. Даже такая программа позволяла разместить в оперативной памяти задачу с $N = 640$ предприятий и решить её за 15 мин.

Если же экономней написать программу, то на РС-486 можно будет проводить расчёты для $N \approx 3000$ предприятий, причём за время ≈ 7 мин. При увеличении N трудоёмкость расчёта и требуемая память будут расти $\sim N^{3/2}$.

2.2.6. Методика расчётов. Матрица долгов и столбец оборотных средств предприятий формировались по описанной ранее методике с помощью псевдослучайных чисел. Было сделано только одно, но

Короткая выдача

N = 14	B = 0,2	eps = 10 ⁻⁵	Q = 16	J = 100	nvud = 3	Итоговая выдача			
Iq	Ka	Kf	S	Sa	C	Cf	D	Df	Pq
6	0,7305	0,5519	7,9792	8,4552	31,3731	14,0593	31,3731	14,0593	0,7574

Средняя выдача

n	Rn	bn	sn	bnf	cn	cnf	Dn	dnf	Pqn
1	1,0000	0,0971	-1,4045	0,0000	1,7525	1,4045	0,2509	0,0000	0,1986
2	1,0377	0,5421	0,7373	0,3347	1,3791	0,0000	1,5743	0,4026	1,0000
3	1,0801	0,5380	-0,7491	-0,0001	3,2405	1,7806	1,9534	1,0315	0,4505
4	1,1282	0,4657	1,7916	0,1589	2,2923	0,0000	3,6181	1,6327	1,0000
5	1,1832	0,4487	-0,2843	0,0000	2,7796	1,6276	2,0466	1,3433	0,4144
6	1,2472	0,7489	3,4871	2,3292	0,1417	0,0000	2,8799	1,1578	1,0000
7	1,3229	0,1625	2,1450	0,6604	1,8318	0,0000	3,8142	1,4846	1,0000
8	1,4142	0,3912	-0,4089	0,0000	2,4294	1,0183	1,6293	0,6094	0,5808
9	1,5275	0,9277	2,2144	1,5494	0,0713	0,0000	1,3580	0,6651	1,0000
10	1,6733	1,2339	-2,7710	0,0000	4,5137	2,8956	0,5088	0,1246	0,3585
11	1,8708	0,9060	0,6372	0,0000	2,3309	0,2493	2,0621	0,8865	0,8931
12	2,1602	0,7181	-1,2806	0,0000	4,4012	2,0984	2,4024	0,8178	0,5232
13	2,6458	0,6801	5,4219	2,9468	0,0540	0,0000	4,7958	2,4751	1,0000
14	3,7417	0,1191	-1,5568	0,0000	4,1552	2,9851	2,4793	1,4283	0,2816

существенное изменение: знак долга стал определяться по формуле

$$\text{sign } c_{nm} = \begin{cases} +1 & \text{при } \gamma > 1 - N^{-1/2}, \\ 0 & \text{при } N^{-1/2} < \gamma < 1 - N^{-1/2}, \\ -1 & \text{при } \gamma < N^{-1/2}. \end{cases}$$

Это означает, что $\sim N^{-1/2}$ доли долгов будут ненулевыми, а остальные — нулевыми. Тем самым, каждое предприятие связано не со всеми предприятиями, а лишь с их частью. Поскольку всего предприятий N , то среднее число связей каждого предприятия $\sim 2N^{1/2}$, что правдоподобно.

Это важно по следующей причине. Наличие нулевых долгов разрывает долговые цепочки, то есть затрудняет взаимозачёт. Это приближает модельную систему долгов к реальности.

В реальности есть ещё один фактор: наличие естественных монополий, например, энергетиков. У него два аспекта. Первый — взвинчивание цен, из-за чего долги не случайны, а «перекошены» в сторону этих монополистов. Второй — предприятия могут быть не связаны между

Матрица конечных долгов C_{fin}

1	0,0000	0,0000	0,0000	-0,7030	-0,7015	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0000	0,0000	0,4026	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
3	0,0000	-0,4026	0,0000	0,0000	0,3735	0,0000	-0,4527	0,1347	-0,2222	0,1922	-0,4071	0,3311	-0,2959	0,0000	0,0000	0,0000	0,0000
4	0,7030	0,0000	0,0000	0,0000	0,2342	0,0000	0,0000	0,3566	0,0000	0,0000	0,0000	0,3389	0,0000	0,0000	0,0000	0,0000	0,0000
5	0,7015	0,0000	-0,3735	-0,2342	0,0000	0,0000	-0,5050	0,0000	0,0000	-0,1188	-0,0593	0,0000	-0,3368	0,6418	0,0000	0,0000	0,0000
6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,3423	0,0687	0,0000	0,0000	0,0000	0,0000	0,7469	0,0000
7	0,0000	0,0000	0,4527	0,0000	0,5050	0,0000	0,0000	0,5270	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
8	0,0000	0,0000	-0,1347	-0,3566	0,0000	0,0000	-0,5270	0,0000	0,0000	0,0000	0,0311	0,5783	0,0000	0,0000	0,0000	0,0000	0,0000
9	0,0000	0,0000	0,2222	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,4428	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
10	0,0000	0,0000	-0,1922	0,0000	0,1188	-0,3423	0,0000	0,0000	-0,4428	0,0000	0,0058	-0,6741	-0,2855	-0,9587	0,0000	0,0000	0,0000
11	0,0000	0,0000	0,4071	0,0000	0,0593	-0,0687	0,0000	-0,0311	0,0000	-0,0058	0,0000	-0,1436	0,0000	0,4200	0,0000	0,0000	0,0000
12	0,0000	0,0000	-0,3311	-0,3389	0,0000	0,0000	0,0000	-0,5783	0,0000	0,6741	0,1436	0,0000	-0,3805	-0,4696	0,0000	0,0000	0,0000
13	0,0000	0,0000	0,2959	0,0000	0,3368	0,0000	0,0000	0,0000	0,0000	0,2855	0,0000	0,3805	0,0000	1,1764	0,0000	0,0000	0,0000
14	0,0000	0,0000	0,0000	0,0000	-0,6418	-0,7469	0,0000	0,0000	0,0000	0,9587	-0,4200	0,4696	-1,1764	0,0000	0,0000	0,0000	0,0000

собой, но все они связаны с такими монополистами. Этот фактор пока не моделируется, считая экономику «нормальной».

Для улучшения статистики каждый методический расчёт повторялся 10 раз с разными псевдослучайными числами (ранее мы ограничивались 5 расчётами). Далее использовалось усреднение этих расчётов с вычислением стандартного отклонения.

7. Результаты. Были проведены две серии расчётов. Первая серия — при $N = 20$ с различными значениями b ; вторая — при $b = 0.1$ с различными значениями N . Величина b для второй серии выбрана меньше, чем ранее: среди долгов теперь много нулевых, что снижает требования к оборотным средствам.

Результаты расчётов показаны на рис. 9–18. Там же приведены подобранные для них аппроксимации. Точность этих аппроксимаций лучше, чем в предыдущем отчёте, что свидетельствует о более удачном выборе зависимостей. Приведём эти аппроксимации и границы их хорошей точности:

$$\left. \begin{aligned} q(b) &\approx (0.9 + b)/(0.018 + b), & 0 \leq b \leq 1.6 \\ k_a(b)\% &\approx 100 - 27.8(1 + 0.986b)^{-3}, & 0 \leq b \leq 1.6 \\ (k_a - k_f)\% &\approx 31 \cdot 2^{-10b}, & 0.05 \leq b \leq 1.6 \\ (k_a - p)\% &\approx -1.4, & 0.1 \leq b \leq 1.6 \end{aligned} \right\} N = 20,$$

$$\left. \begin{aligned} q(N) &\approx 14.63 - 2.25 \lg N, & 10 \leq N \leq 640 \\ k_a(N)\% &\approx 100 - 64.9N^{-1/3}, & 5 \leq N \leq 640 \\ (k_a - k_f)\% &\approx 62.3N^{-0.4}, & 10 \leq N \leq 640 \\ (k_a - p)\% &\approx 1.5, & 10 \leq N \leq 640 \end{aligned} \right\} b = 0.1.$$

Обсудим полученные результаты.

Из рис. 9 видно, что число q велико при отсутствии оборотных средств ($b = 0$) и $q \rightarrow 1$ при увеличении b . Аппроксимация $q(b)$ оказалась хорошей в широком диапазоне b .

Из рис. 10 видно, что коэффициент идеального взаимозачёта k_a быстро стремится к 100% при возрастании b , причём он велик даже при $b = 0$! Разумеется, это верно лишь для нормальной экономики, без монопольных цен. Отметим очень широкий диапазон аппроксимации $k_a(b)$.

Коэффициент фракционного взаимозачёта k_f всегда меньше k_a , но стремится к нему при $b \rightarrow \infty$ (рис. 11); стремление экспоненциальное, то есть быстрое. Лишь при $b = 0$ отличие k_f от k_a значительное. Уже при умеренных оборотных средствах фракционный взаимозачёт лишь немногим уступает идеальному.

Наконец, средняя надёжность долгов p почти не отличается от k_a уже при умеренных оборотных средствах (рис. 12). Различие видно лишь при очень малых $b \leq 0.05$, причём $p < k_a$.

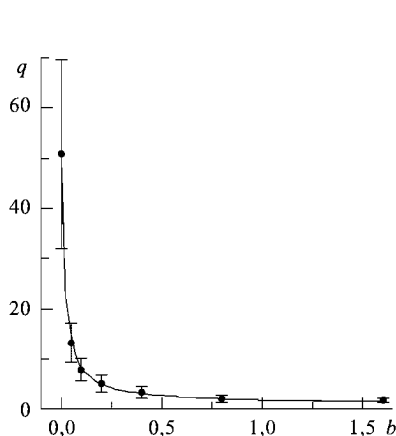


Рис. 9. Число итераций при $N = 20$; $\bullet$ — расчёты с указанием разброса, кривая — аппроксимация

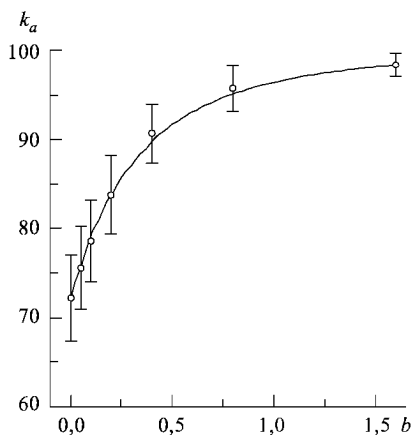


Рис. 10. Идеальный взаимозачёт при $N = 20$; $\circ$ — расчётные k_a , кривая — аппроксимация

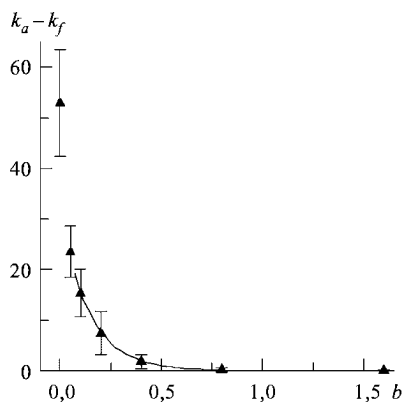


Рис. 11. Фракционный взаимозачёт при $N = 20$; $\blacktriangle$ — расчётные значения $k_a - k_f$, кривая — аппроксимация

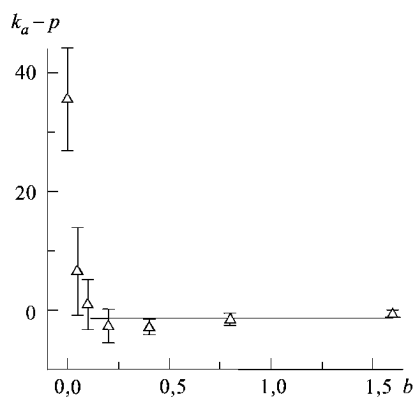


Рис. 12. Средняя надёжность долгов при $N = 20$; $\triangle$ — расчётные значения $k_a - p$, кривая — аппроксимация

Для удобства сопоставления разных величин все они показаны на одном рис. 13 в зависимости от b . Аналогичные кривые для зависимостей от N изображены на рис. 14.

На рис. 15 показана зависимость $q(N)$. При возрастании N число итераций не выходит на константу, а даже слегка убывает. Вероятно, это обусловлено тем, что все возможные долги были ненулевыми. Сейчас есть много нулевых долгов, причём их доля тем больше, чем больше N . Это обстоятельство обещает экономичность итерационного

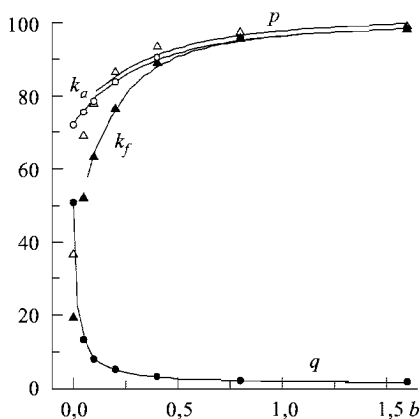


Рис. 13. Сводный график при $N = 20$; кривые — аппроксимации, расчёты $\bullet$ — q , $\circ$ — k_a , $\blacktriangle$ — k_f , Δ — p

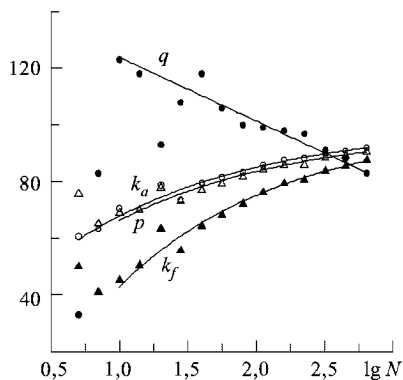


Рис. 14. Сводный график при $b = 0,1$; обозначения см. рис. 13 (но вместо числа итераций показано $10q$)

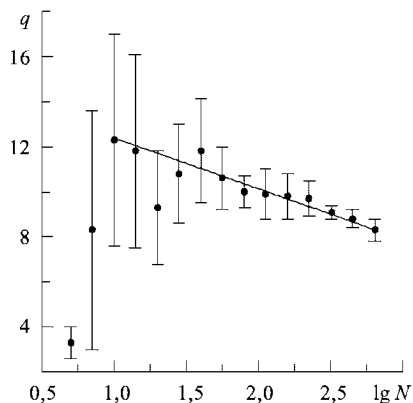


Рис. 15. Число итераций при $b = 0,1$; обозначения см. рис. 9

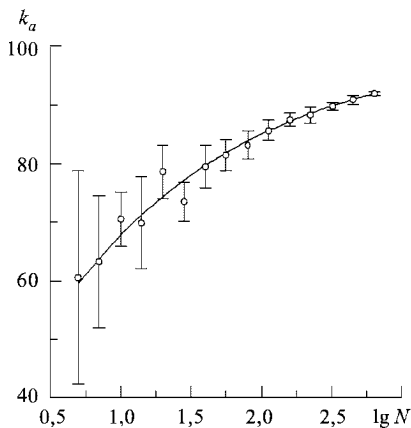


Рис. 16. Идеальный взаимозачёт при $b = 0,1$; обозначения см. рис. 10

процесса. Заметим, что выход на аппроксимацию происходит уже при $N \geq 10$, то есть для малых систем предприятий.

Зависимость $k_a(N)$ хорошо описывается аппроксимацией в очень широком диапазоне $N \geq 5$ (рис. 16), то есть для очень малых систем. Интересно, что скорость выхода на 100%-й предел соответствует не закону больших чисел $\sim N^{-1/2}$, а закону $N^{-1/3}$. Вероятная причина — наличие нулевых долгов, когда число ненулевых долгов само мало $\sim N^{-1/2}$.

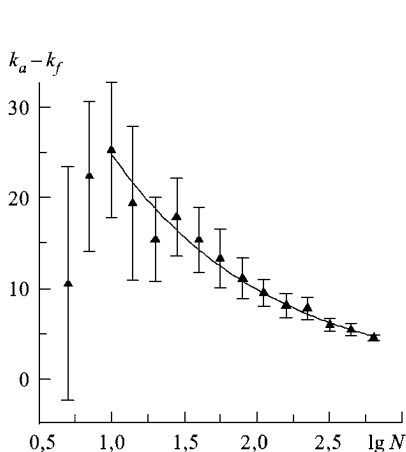


Рис. 17. Фракционный взаимозачёт при $b = 0,1$; обозначения см. рис. 11

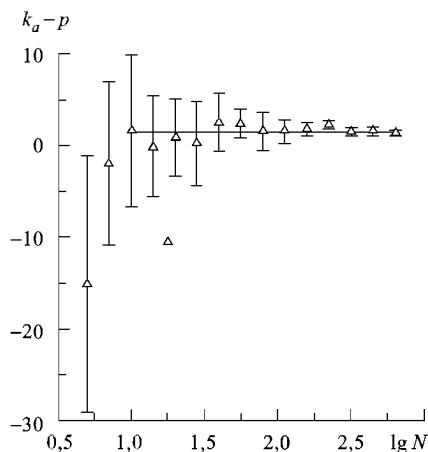


Рис. 18. Средняя надёжность долгов при $b = 0,1$; обозначения см. рис. 12

Наконец, отличие p от k_a почти несущественно при всех N , кроме самых маленьких $N \leq 7$ (рис. 18).

8. Выводы. Предложенный здесь фракционный взаимозачёт является очень перспективным: 1) он не требует переадресации долгов; 2) для «нормальной» экономики его эффективность мало уступает идеальному взаимозачёту, если число предприятий $N > 10$, а их оборотные средства не слишком малы; 3) алгоритм нахождения решения очень экономичен и позволяет решать задачи с $N \approx 10^6$.

2.3. Экспериментальная проверка метода фракционирования долгов для региона (Башкирия).

2.3.1. Методы взаимозачёта. Возможности метода поиска замкнутых цепочек долгов невелики. Обычно он позволяет погасить лишь небольшую часть взаимных долгов системы предприятий. Вдобавок для большого числа предприятий поиск цепочек становится очень трудоёмким. Поэтому сначала предложен метод переадресации долгов [1–4], имитировавший обращение векселей. Позднее, усовершенствуя этот метод, найден метод фракционного взаимозачёта [5, 7], эквивалентный нахождению всех возможных цепочек долгов, но имеющий малую трудоёмкость.

Проверка метода фракционирования проводилась на модельных примерах, имитирующих замкнутую систему предприятий, среди которых нет монополистов и которые имеют довольно густую сеть взаимных связей. Метод легко позволял гасить большую часть долгов и почти не уступал по эффективности методу переадресации.

Очень желательно было проверить метод не только на модельных примерах, но и на реальных. Реальная экономическая информа-

ция обычно труднодоступна. Однако удалось найти описанные ниже данные.

2.3.2. Сравнения в 1999 г.

По-видимому, здесь учтены долги предприятий налоговому ведомству и различным фондам, но «забыты» долги госбюджета предприятиям за оборонные заказы и т. п.

В Башкирии было взято 120 предприятий при 13573 долговых обязательствах по состоянию на 1997 г. для тестирования метода фракционирования.

Однако при начале этой работы, обнаружились неполнота исходной информации и наличие технических ошибок в ней. По-видимому, то же относится ко всем примерам, но в 1999 г. анализ исходной информации на полноту и непротиворечивость не проводился. Ниже описано, как была сделана выверка.

2.3.3. Выверка информации. Списки своих долговых обязательств подали 120 предприятий. Сводный список занимал 13573 строки, заполненных следующим образом соответственно для кредиторской или дебиторской задолженностей:

- ИНН предприятия • ИНН кредитора • сумма долга • ноль •
- ИНН предприятия • ноль • ИНН дебитора • сумма долга •

Все ИНН были десятизначными целыми числами, все денежные суммы приводились в неденоминированных рублях как нецелые числа с двумя знаками после точки (копейки).

Сначала мы провели проверку формальной правильности заполнения каждой строки. Было обнаружено 2 строки, где отсутствовал ноль в нужной позиции. Эти нули были восстановлены. Затем было обнаружено 11 двоек строк, содержащих пять ненулевых позиций. По структуре чисел (целые или нецелые) и порядку их расположения было видно, что в каждой такой строке содержались две записи о кредиторских задолженностях. Эти записи были вручную приведены к правильной форме, после чего число строк-обязательств увеличилось до 13584. Эти строки были перенумерованы (в исходных данных строки не нумеровались).

В графах дебиторов и кредиторов означены не только исходные 120 предприятий. Всего там имеется 7754 предприятия из различных регионов, включая налоговую инспекцию и различные фонды (некоторые расшифрованы в табл. 3). Для удобства выдач всем предприятиям были присвоены порядковые номера и составлена табл. 4. В ней помимо ИНН предприятий приведены их условные порядковые номера. Номера 1–120 соответствуют исходным предприятиям; 121–4997 тем предприятиям, которые дополнительно появляются в списке кредиторов; 4998–7754 тем, которые есть только в списке дебиторов.

Содержание столбцов табл. 4:

INN — ИНН предприятия, подавшего список;

INN_{cred} — ИНН кредитора для кредиторской задолженности, ноль для дебиторской;

Таблица 3. Расшифровка некоторых ИНН

Федеральный бюджет	7710 152 113	Гос. фонды
Фонд гос. соц. страхования	9999 113 090	
Амортизац. Фонд	9999 112 969	
Респ. фонд, амортиз.	9999 112 850	
Налоги во внебюдж. фонд	9999 112 705	
Фонд занятости	9999 112 301	
Пенсионный фонд	9999 111 489	
Фонд соц. страх.	0 275 003 528	
Фонд медицинский	0 274 026 307	Местн. фонды
Фонд мед. страх.	0 268 003 453	

S or INN_{deb} — сумма долга для кредиторской задолженности, ИНН дебитора для дебиторской;

0 or S — нуль для кредиторской задолженности, сумма долга для дебиторской;

N — номер строки по порядку (от 1 до 13584);

$N1$ — условный номер предприятия, подавшего список;

$N2$ — условный номер кредитора для кредиторской зависимости, нуль для дебиторской;

$N3$ — нуль для кредиторской зависимости, условный номер предприятия для дебиторской;

INN — ИНН предприятия;

$N0$ — условный номер этого предприятия (от 1 до 7754).

Все денежные суммы (в неденоминированных рублях) записаны как нецелые числа, но без знаков после точки (копейки отброшены). Цифры ИНН заменены звёздочками ввиду конфиденциальности этой информации.

Таблица 4. Выверенный список всех задолженностей

INN	INN_{cred}	S or INN_{deb}	0 or S	N	$N1$	$N2$	$N3$	INN	$N0$
*****	*****	1575.	0.	1	1	121	0	0.	0
*****	*****	1235.	0.	2	1	122	0	*****	1
*****	*****	78590.	0.	3	1	67	0	*****	2
*****	*****	1716.	0.	4	1	123	0	*****	3
*****	*****	1200.	0.	5	1	124	0	*****	4
*****	0.	*****	84364.	6	1	0	4998	*****	5
*****	0.	*****	32450.	7	1	0	4999	*****	6

Продолжение табл. 4

INN	INNcred	S or INNdeb	0 or S	N	N1	N2	N3	INN	N0
*****	0.	*****	1174.	8	1	0	5000	*****	7
*****	0.	*****	20460.	9	1	0	5001	*****	8
*****	0.	*****	10816.	10	1	0	5002	*****	9
*****	0.	*****	40661.	11	1	0	5003	*****	10
*****	0.	*****	7492.	12	1	0	5004	*****	11
*****	0.	*****	5086.	13	1	0	5005	*****	12
*****	0.	*****	1091.	14	1	0	5006	*****	13
*****	0.	*****	2650.	15	1	0	5007	*****	14
*****	0.	*****	1337.	16	1	0	5008	*****	15
*****	0.	*****	93570.	17	1	0	1264	*****	16
*****	0.	*****	6171.	18	1	0	5009	*****	17
*****	0.	*****	7698.	19	1	0	5010	*****	18
*****	0.	*****	11740.	20	1	0	5011	*****	19
*****	0.	*****	4609.	21	1	0	5012	*****	20
*****	0.	*****	16097.	22	1	0	5013	*****	21
*****	0.	*****	2987.	23	1	0	5014	*****	22
*****	0.	*****	175 777.	24	1	0	5015	*****	23
*****	0.	*****	571 088.	25	1	0	5016	*****	24
*****	0.	*****	1061.	26	1	0	5017	*****	25
*****	0.	*****	33379.	27	1	0	5018	*****	26
*****	0.	*****	7329.	28	1	0	5019	*****	27
*****	0.	*****	1304.	29	1	0	5020	*****	28
*****	0.	*****	8500.	30	1	0	2299	*****	29
*****	*****	38115.	0.	31	2	125	0	*****	30
*****	*****	482 160.	0.	32	2	126	0	*****	31
*****	*****	135 365.	0.	33	2	80	0	*****	32
*****	*****	164 461.	0.	34	2	80	0	*****	33
*****	0.	*****	47810.	35	2	0	5021	*****	34
*****	0.	*****	13210.	36	2	0	5022	*****	35

Ввиду большого объёма информации, приведена только первая страница выдачи в качестве образца. Она рекомендуется как пример с формально проверенной информацией, хотя и неполной.

2.3.4. Внутренние долги. Далее исходные 120 предприятий будем именовать *внутренними*, а все остальные — *внешними*. Внешние предприятия не подавали списков своих долгов, так что для них классический взаимозачёт невозможен (применим лишь метод переадресации). Поэтому из общего списка был выделен список *внутренних* долгов, когда и кредитор и дебитор были внутренними предприятиями. Этот список содержал 1040 кредиторских и 580 дебиторских задолженностей. Он приведён в табл. 5.

Содержание столбцов табл. 5:

K1 — порядковый номер кредиторской записи (от 1 до 1040)

S1 — сумма кредиторской задолженности в неденоминированных рублях;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер кредитора (от 1 до 120);

K2 — порядковый номер дебиторской записи (от 1 до 580);

S2 — сумма дебиторской задолженности в неденоминированных рублях;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер дебитора (от 1 до 120).

Таблица 5. Выверенный список внутренних задолженностей с отдельными обязательствами

k1	S1	N1	N2		k2	S2	N1	N2
1	78590.	1	67		1	1149.	3	49
2	135 365.	2	80		2	1 028 547.	4	22
3	164 461.	2	80		3	2 038 500.	4	72
4	460 482.	4	8		4	33 130 080.	4	80
5	287 771 937.	4	64		5	7500.	19	6
6	37876.	4	72		6	63462.	20	85
7	1 669 800.	4	72		7	21489.	20	109
8	155 688 588.	4	72		8	12344.	21	20
9	5 160 178.	4	72		9	5927.	22	21
10	17 355 784.	4	72		10	5776.	22	36
11	21 678 245.	4	72		11	34832.	22	44
12	18 514 002.	4	104		12	91568.	22	80
13	58670.	5	67		13	12400.	22	80
14	100 000.	6	67		14	37922.	22	80
15	281 962.	7	54		15	25179.	22	104

Продолжение табл. 5

k1	S1	N1	N2		k2	S2	N1	N2
16	1285.	7	60		16	70421.	23	22
17	118000.	9	36		17	23504.	23	64
18	254651.	10	54		18	24000.	24	39
19	897615.	12	54		19	3569.	24	109
20	13463.	17	44		20	2517.	31	16
21	947.	17	66		21	11206.	31	43
22	6328.	17	80		22	1228.	31	90
23	2248.	17	80		23	2517.	32	16
24	5149.	17	80		24	11206.	32	43
25	59982280.	17	85		25	1228.	32	90
26	1329.	17	92		26	14558.	34	41
27	4924690.	17	92		27	33059.	35	32
28	1675.	17	104		28	2880.	35	64
29	23699.	17	109		29	657.	35	72
30	4400.	19	23		30	149.	35	90
31	12058062.	20	21		31	33759.	35	104
32	15718.	20	80		32	294395600.	36	24
33	21051.	20	85		33	438700937.	36	64
34	14960.	21	80		34	176500.	38	22
35	63908.	22	20		35	11414.	38	66
36	96282.	22	21		36	74212.	39	44
37	99096.	22	21		37	74212.	39	44
38	81599.	22	23		38	397135.	39	72
39	3432.	22	23		39	450735.	39	80
40	201606.	22	38		40	288833.	39	80
41	2310.	22	41		41	97308.	40	42
42	35384.	22	44		42	2000.	40	42
43	16240.	22	64		43	8835.	40	80
44	638813.	22	64		44	4800.	40	80
45	354554.	22	64		45	10307.	40	104
46	129766.	22	104		46	253228.	41	38

Но ведь внутренние предприятия и их долги образуют замкнутую систему. В ней число кредиторских и дебиторских записей должно совпадать. Сами денежные величины должны удовлетворять соотношениям $c_{ит} = d_{ит}$, а суммарное сальдо всех предприятий должно

обращаться в нуль. Всё это не выполняется. Следовательно, исходный материал неполон. Первоначально опасались недостоверности, но 12 перекрёстных обязательств точно совпадали. Это указывало, что предприятия дали достоверные записи, но не полные, а частичные. Кроме того, между одной и той же парой предприятий в ряде случаев было несколько обязательств.

Поэтому были, во-первых, просуммированы все обязательства каждой пары. Тогда список сократился до 269 кредиторских и 259 дебиторских задолженностей. Он приведён в табл. 6. Из них только 84 относились к одинаковым парам, но лишь у 16 выполнялось равенство $c_{ит} = d_{ит}$. Тем самым, стопроцентно надёжными были лишь 16 суммарных долговых обязательств для 120 предприятий — в среднем 1/7 обязательства на одно предприятие! Такого материала совершенно недостаточно для взаимозачёта долгов. Так что сравнения методов в 1999 г. и не могли ничего выявить. Надо было сначала привести в порядок исходный материал.

Обозначения столбцов табл. 6:

K1 — порядковый номер кредиторской записи (от 1 до 269)

S1 — сумма кредиторской задолженности в неденоминированных рублях;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер кредитора (от 1 до 120);

K2 — порядковый номер дебиторской записи (от 1 до 259);

S2 — сумма дебиторской задолженности в неденоминированных рублях;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер дебитора (от 1 до 120).

1040 line → 269 k1 580 line → 259 k2

Таблица 6. Выверенный список внутренних задолженностей с суммированными обязательствами (начало)

k1	S1	N1	N2	k2	S2	N1	N2
1	78590.	1	67	1	1149.	3	49
2	299 826.	2	80	2	1 028 547.	4	22
3	460 482.	4	8	3	2 038 500.	4	72
4	287 771 937.	4	64	4	33 130 080.	4	80
5	201 590 471.	4	72	5	7500.	19	6
6	18 514 002.	4	104	6	63462.	20	85
7	58670.	5	67	7	21489.	20	109
8	100 000.	6	67	8	12344.	21	20
9	281 962.	7	54	9	5927.	22	21

Продолжение табл. 6

k1	S1	N1	N2	k2	S2	N1	N2
10	1285.	7	60	10	5776.	22	36
11	118 000.	9	36	11	34832.	22	44
12	254 651.	10	54	12	141 890.	22	80
13	897 615.	12	54	13	25179.	22	104
14	13463.	17	44	14	70421.	23	22
15	947.	17	66	15	23504.	23	64
16	13725.	17	80	16	24000.	24	39
17	59 982 280.	17	85	17	3569.	24	109
18	4 926 019.	17	92	18	2517.	31	16
19	1675.	17	104	19	11206.	31	43
20	23699.	17	109	20	1228.	31	90
21	4400.	19	23	21	2517.	32	16
22	12 058 062.	20	21	22	11206.	32	43
23	15718.	20	80	23	1228.	32	90
24	21051.	20	85	24	14558.	34	41
25	14960.	21	80	25	33059.	35	32
26	63908.	22	20	26	2880.	35	64
27	195 378.	22	21	27	657.	35	72
28	85031.	22	23	28	149.	35	90
29	201 606.	22	38	29	33759.	35	104
30	2310.	22	41	30	294 395 600.	36	24
31	35384.	22	44	31	438 700 937.	36	64
32	1 009 607.	22	64	32	176 500.	38	22
33	129 766.	22	104	33	11414.	38	66
34	7838.	22	109	34	148 424.	39	44
35	18313.	24	22	35	397 135.	39	72
36	15025.	24	44	36	739 568.	39	80
37	1 300 597.	24	64	37	99308.	40	42
38	63782.	24	80	38	13635.	40	80

Продолжение табл. 6

k1	S1	N1	N2	k2	S2	N1	N2
39	65018.	26	74	39	10307.	40	104
40	222 800.	27	67	40	253 228.	41	38
41	19811.	30	31	41	4445.	42	22
42	644.	30	65	42	91138.	42	32

Во-вторых, для получения разумных данных был дополнен список, считая бухгалтерии предприятий добросовестными, но неаккуратными. Если о долговом обязательстве сообщала только одна сторона, оно включалось в список. Если сообщали обе стороны, но суммы не совпадали, бралась большая из сумм. После этого список составил по 444 строки как кредиторской, так и дебиторской задолженности при совпадении $c_{nm} = d_{mn}$. При этом у 9 предприятий не оказалось никаких внутренних обязательств, так что этот список фактически включал 111 предприятий — в среднем по 4 связи у каждого предприятия.

Этот список в форме кредиторской задолженности приведён в табл. 7. Список дебиторской задолженности получается из данного по симметрии перестановке номеров $N1 \leftrightarrow N2$. Табл. 7 является выверенным материалом для тестирования методов. Ввиду её особой важности, её текст приведён здесь полностью.

Обозначения столбцов табл. 7:

K — порядковый номер кредиторской записи (от 1 до 222)

S1 — сумма кредиторской задолженности в неденоминированных рублях;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер кредитора (от 1 до 120);

K — порядковый номер кредиторской записи (от 223 до 444)

S1 — сумма кредиторской задолженности в неденоминированных рублях;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер кредитора (от 1 до 120);

Таблица 7. Дополненный по симметрии список внутренних кредиторских задолженностей

k	S1	N1	N2		k	S1	N1	N2
1	78590.	1	67		223	37410.	94	79
2	299 826.	2	80		224	13781.	94	80
3	460 482.	4	8		225	64043.	94	84
4	287 771 937.	4	64		226	169 115.	94	91

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
5	201 590 471.	4	72		227	62246.	94	114
6	18 514 002.	4	104		228	300 000.	95	89
7	58670.	5	67		229	250 304.	95	109
8	100 000.	6	67		230	300 000.	95	118
9	281 962.	7	54		231	99208.	97	96
10	1285.	7	60		232	104 735.	99	80
11	118 000.	9	36		233	4524.	99	85
12	254 651.	10	54		234	187 257.	99	89
13	897 615.	12	54		235	226 856.	99	91
14	13463.	17	44		236	26817.	99	92
15	947.	17	66		237	3366.	99	96
16	13725.	17	80		238	5881.	100	86
17	59 982 280.	17	85		239	73032.	100	88
18	4 926 019.	17	92		240	89 706 283.	102	64
19	1675.	17	104		241	41 008 409.	102	72
20	23699.	17	109		242	1 135 986.	102	80
21	4400.	19	23		243	24595.	103	64
22	12 058 062.	20	21		244	730.	103	72
23	15718.	20	80		245	11 407 535.	103	79
24	21051.	20	85		246	52 408 298.	103	80
25	14960.	21	80		247	128 923.	103	81
26	63908.	22	20		248	20471.	103	108
27	195 378.	22	21		249	920.	103	109
28	85031.	22	23		250	44337.	108	84
29	201 606.	22	38		251	24850.	108	103
30	2310.	22	41		252	28150.	109	54
31	35384.	22	44		253	5441.	109	77
32	1 009 607.	22	64		254	5 193 600.	109	80
33	129 766.	22	104		255	54342.	109	90

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
34	7838.	22	109		256	3648.	109	103
35	18313.	24	22		257	28150.	110	54
36	15025.	24	44		258	15400.	111	77
37	1 300 597.	24	64		259	62798.	112	107
38	63782.	24	80		260	4260.	114	27
39	65018.	26	74		261	59345.	114	79
40	222 800.	27	67		262	35115.	114	80
41	19811.	30	31		263	6659.	115	80
42	644.	30	65		264	1 497 289.	116	80
43	20733.	30	67		265	11081.	119	33
44	2000.	30	71		266	15330.	119	38
45	4 926 000.	32	44		267	107 025.	119	54
46	5556.	33	84		268	87204.	119	65
47	594.	35	24		***269	3640.	119	103
48	531 936.	35	32		270	1149.	49	3
49	660 357.	35	34		271	1 028 547.	22	4
50	230 406.	35	44		272	2 038 500.	72	4
51	961 696.	35	72		273	7500.	6	19
52	16796.	35	80		274	63462.	85	20
53	53 494 056.	35	85		275	21489.	109	20
54	178 928.	35	91		276	5927.	21	22
55	7661.	35	109		277	5776.	36	22
56	12 657 265.	36	37		278	34832.	44	22
57	596 901 864.	36	64		279	25179.	104	22
58	5 352 968.	36	104		280	23504.	64	23
59	5350.	38	77		281	24000.	39	24
60	13667.	39	38		282	3569.	109	24
61	687 811.	39	44		283	2517.	16	31
62	96857.	39	64		284	11206.	43	31

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
63	190038.	39	80		285	1228.	90	31
64	74322612.	39	85		286	2517.	16	32
65	793782.	39	91		287	11206.	43	32
66	30934.	39	92		288	1228.	90	32
67	27739.	40	17		289	14558.	41	34
68	517772.	40	38		290	33059.	32	35
69	9611.	40	64		291	2880.	64	35
70	10946.	40	77		292	657.	72	35
71	4577.	40	85		293	149.	90	35
72	9436.	40	90		294	33759.	104	35
73	4159.	40	104		295	294395600.	24	36
74	61881.	41	67		296	438700937.	64	36
75	4534752.	42	22		297	148424.	44	39
76	118714.	42	38		298	10307.	104	40
77	99308.	42	40		299	253228.	38	41
78	991.	42	43		300	4445.	22	42
79	777513.	42	44		301	91138.	32	42
80	50326.	42	79		302	188497.	38	42
81	464467366.	42	80		303	2.	40	42
82	25006359.	42	85		304	4697.	48	42
83	2251314.	42	90		305	444238.	57	42
84	144520.	42	91		306	17966.	77	42
85	3079568.	42	92		307	6882.	79	42
86	321095.	42	109		308	30643137.	85	42
87	29100.	43	67		309	179617.	90	42
88	4000.	43	84		310	74662.	103	42
89	3264.	44	35		311	559298634.	104	42
90	3400000.	44	42		312	5281240.	109	42
91	156054.	44	57		313	15120.	112	42

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
92	8549.	44	76		314	15100.	44	43
93	26023.	44	77		315	180473.	106	44
94	7980.	44	78		316	10556677.	109	44
95	567052.	44	90		317	134202.	119	44
96	2059.	44	111		318	62855.	44	45
97	412.	44	112		319	1446.	38	48
98	26486.	45	43		320	22746.	54	49
99	3700.	45	89		321	26100.	64	52
100	30976.	46	41		322	498435.	42	57
101	158269.	46	80		323	97675.	77	57
102	612176.	49	54		324	596252.	86	59
103	16021.	50	67		325	2604.	54	62
104	27442.	52	67		326	20509449339.	80	64
105	1757.	53	60		327	58410.	90	66
106	2453743.	53	109		328	45576.	109	66
107	5076.	55	54		329	35320.	11	67
108	1527.	56	80		330	102710.	13	67
109	6245.	57	38		331	40625.	18	67
110	586000.	57	54		332	55363.	25	67
111	4444.	57	85		333	101210.	31	67
112	332373.	59	54		334	102943.	34	67
113	21000.	61	54		335	92403.	37	67
114	5441.	61	77		336	65809.	51	67
115	385622.	62	79		337	146253.	65	67
116	19819.	63	20		338	781986.	114	67
117	75000.	63	54		339	3688.	92	71
118	23786054667.	64	80		340	1100.	17	72
119	150000.	65	80		341	800905.	39	72
120	11414.	66	38		342	597699.	40	72

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
121	173 658.	66	44		343	14135.	64	72
122	110 263.	66	64		344	899.	74	72
123	312 190.	66	72		345	584 247.	80	72
124	3 714 064.	66	80		346	265.	104	72
125	202 270 791.	66	85		347	119 363.	96	74
126	4 630 973.	66	91		348	536 516.	97	74
127	1440.	66	108		349	129 340.	98	74
128	8459.	67	71		350	181 500 000.	80	75
129	1 864 473.	68	79		351	6 780 893.	79	76
130	3 278 851.	68	80		352	9887.	17	77
131	10276.	69	73		353	195 524.	89	77
132	3 819 880.	70	79		354	114 017.	91	77
133	2 611 856.	70	80		355	97355.	79	78
134	142 177.	71	67		356	24073.	64	79
135	37410.	71	79		357	36.	67	79
136	68791.	71	84		358	13798.	72	79
137	210 168.	71	91		359	17874.	74	79
138	24826.	71	108		360	987.	77	79
139	401.	72	17		361	594.	88	79
140	135 372.	72	21		362	729 321.	89	79
141	397 135.	72	39		363	3127.	91	79
142	2253.	72	64		364	3559.	95	79
143	751.	72	66		365	25942.	96	79
144	126.	72	80		366	8767.	98	79
145	2 543 725.	73	79		367	681.	106	79
146	1 251 442.	73	80		368	16916.	108	79
147	5109.	76	44		369	13761.	111	79
148	7971.	76	71		370	65 348 009.	44	80
149	26489.	76	79		371	43562.	41	84

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
150	418 369.	76	80		372	36712.	65	84
151	326 861.	76	96		373	1639.	113	84
152	633 835.	76	97		374	21717.	80	86
153	150 297.	76	107		375	61204.	83	86
154	130 000.	76	108		376	176 135.	97	86
155	409.	77	71		377	1 079 477.	42	87
156	900 783 664.	77	75		378	780 977.	89	88
157	457.	77	81		379	121 280.	90	88
158	12435.	77	85		380	443 713.	91	88
159	45 237 025.	77	89		381	146 170.	92	88
160	1228.	77	90		382	13122.	38	90
161	26 651 479.	77	91		383	50000.	59	90
162	17 061 142.	77	92		384	62425.	79	90
163	79715.	77	112		385	16.	81	90
164	1 472 037.	78	21		386	1697.	98	90
165	2 637 549.	78	80		387	1 339 518.	79	92
166	114 723.	78	111		388	19763.	86	92
167	88 626 000.	79	80		389	10.	72	93
168	94602.	79	91		390	22207.	113	94
169	67994.	79	97		391	380 000.	85	95
170	33 130 080.	80	4		392	48633.	98	96
171	141 890.	80	22		393	2 600 535.	89	98
172	739 568.	80	39		394	32435.	92	98
173	13635.	80	40		395	296 387.	77	99
174	5531.	80	44		396	119 315.	95	99
175	144 339.	80	59		397	31699.	90	100
176	43307.	80	62		398	8748.	97	100
177	257.	80	66		399	174 228.	80	102
178	2 350 913 672.	80	85		400	37018.	64	103

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
179	614 722.	80	90		401	22126.	81	103
180	6878.	80	93		402	8140.	104	103
181	307 647.	80	103		403	155 196.	114	103
182	874.	80	119		404	118 999 000.	38	104
183	14 415 273.	81	79		405	2 852 000.	73	104
184	9 349 779.	81	80		406	85 026 000.	79	104
185	2 316 467.	81	85		407	8 764 000.	81	104
186	8277.	83	67		408	7 214 000.	86	104
187	2173.	83	103		409	227 817 000.	90	104
188	149 034.	85	79		410	19 334 000.	103	104
189	41451.	86	21		411	253 000.	44	106
190	100 083.	86	44		412	346 000.	89	106
191	305.	86	65		413	99000.	91	106
192	2548.	86	67		414	2 509 000.	92	106
193	147.	86	77		415	43000.	94	106
194	109 531.	86	79		416	176 103.	91	107
195	7 765 371.	86	80		417	155 230.	103	107
196	1354.	86	83		418	6640.	108	107
197	82597.	86	95		419	14152.	3	108
198	46509.	86	97		420	1984.	6	108
199	141 545.	86	100		421	4382.	33	108
200	819 061.	86	109		422	4803.	65	108
201	561 677.	90	44		423	26328.	94	108
202	1657.	90	70		424	1654.	97	108
203	36.	90	73		425	335 863.	44	109
204	4 492 053.	90	79		426	322 977.	77	109
205	215 047 248.	90	80		427	5897.	79	109
206	28 173 181.	90	85		428	84979.	92	109
207	6 831 101.	90	89		429	11 511 200.	75	111

Продолжение табл. 7

k	S1	N1	N2		k	S1	N1	N2
208	22 870 522.	90	92		430	262 000.	89	111
209	32.	90	94		431	434 400.	91	111
210	98.	90	96		432	238 300.	92	111
211	21.	90	98		433	10000.	107	113
212	6 024 788.	90	109		434	2315.	8	114
213	15726.	92	42		435	78657.	80	114
214	231 494.	92	44		436	18253.	54	116
215	254 738.	92	57		437	11909.	29	119
216	46980.	92	77		438	3312.	32	119
217	1 921 000.	92	79		439	68692.	44	119
218	1 771 021.	92	86		440	6931.	85	119
219	1208.	93	64		441	102 506.	42	120
220	558 971.	93	72		442	52814.	89	120
221	2006.	93	91		443	163 520.	91	120
222	960.	94	65		444	20760.	92	120

Более тщательный анализ показал ещё одну трудность. Среди погашенных долгов основную часть составили встречные долги 62 пар предприятий друг другу (заметим, что эти встречные долги не связаны с нашим дополнением списка: они уже содержались в исходном списке из 13573 строк). Эти долги их бухгалтерии могли бы обнаружить и взаимно погасить сами, не обращаясь к региональному центру. После такого погашения список нетривиальных долгов сократился до 382 кредиторских и столько же дебиторских записей, в среднем по 3,5 записи на одно предприятие. Это список приведён в табл. 8 (также полностью ввиду его важности). Значительная часть записей в табл. 8 содержит небольшие денежные суммы (до 2 неденоминированных рублей). В результате возможные нетривиальные цепочки быстро обрывались и гасили лишь небольшую часть оставшегося долга.

Содержание столбцов табл. 8:

K1 — порядковый номер остаточного долга (от 1 до 382)

S1 — сумма остаточного долга;

N1 — условный номер исходного предприятия (от 1 до 120);

N2 — условный номер кредитора (от 1 до 120);

K2 — порядковый номер исключаемого встречного долга (от 1 до 62)

S2 — величина исключаемого встречного долга;
 N1 — условный ном.ер исходного предприятия (от 1 до 120);
 N2 — условный номер кредитора (от 1 до 120);
 SUM — сумма исключённых долгов.

Таблица 8. Список внутренней кредиторской задолженности
с исключёнными встречными долгами

K1	S1	N1	N2	k2	S2	N1	N2
1	78590.	1	67	1	2 038 500.	72	4
2	299 826.	2	80	2	401.	72	17
3	14152.	3	108	3	21051.	20	85
4	460 482.	4	8	4	5927.	21	22
5	287 771 937.	4	64	5	4445.	22	42
6	199 551 971.	4	72	6	34832.	44	22
7	18 514 002.	4	104	7	25179.	104	22
8	58670.	5	67	8	33059.	32	35
9	7500.	6	19	9	3264.	44	35
10	100 000.	6	67	10	657.	72	35
11	1984.	6	108	11	438 700 937.	64	36
12	281 962.	7	54	12	118 714.	42	38
13	1285.	7	60	13	148 424.	44	39
14	2315.	8	114	14	397 135.	72	39
15	118 000.	9	36	15	190 038.	39	80
16	254 651.	10	54	16	2.	40	42
17	35320.	11	67	17	4159.	40	104
18	897 615.	12	54	18	777 513.	42	44
19	102 710.	13	67	19	444 238.	57	42
20	2517.	16	31	20	6882.	79	42
21	2517.	16	32	21	25 006 359.	42	85
22	13463.	17	44	22	179 617.	90	42
23	947.	17	66	23	15726.	92	42
24	699.	17	72	24	321 095.	42	109
25	9887.	17	77	25	5109.	76	44

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
26	13725.	17	80		26	5531.	80	44
27	59 982 280.	17	85		27	561 677.	90	44
28	4 926 019.	17	92		28	180 473.	106	44
29	1675.	17	104		29	335 863.	44	109
30	23699.	17	109		30	68692.	44	119
31	40625.	18	67		31	22746.	54	49
32	4400.	19	23		32	2253.	72	64
33	12 058 062.	20	21		33	20 509 449 339.	80	64
34	15718.	20	80		34	24595.	103	64
35	14960.	21	80		35	751.	72	66
36	1 028 547.	22	4		36	257.	80	66
37	63908.	22	20		37	8459.	67	71
38	189451.	22	21		38	126.	72	80
39	85031.	22	23		39	10.	72	93
40	201 606.	22	38		40	26489.	76	79
41	2310.	22	41		41	195 524.	89	77
42	552.	22	44		42	114 017.	91	77
43	1 009 607.	22	64		43	46980.	92	77
44	104 587.	22	104		44	5441.	109	77
45	7838.	22	109		45	62425.	79	90
46	18313.	24	22		46	3127.	91	79
47	294 395 600.	24	36		47	1 339 518.	79	92
48	15025.	24	44		48	21717.	80	86
49	1 300 597.	24	64		49	614 722.	80	90
50	63782.	24	80		50	174 228.	80	102
51	55363.	25	67		51	307 647.	80	103
52	65018.	26	74		52	35115.	114	80
53	222 800.	27	67		53	22126.	81	103
54	11909.	29	119		54	1354.	86	83

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
55	19811.	30	31		55	19763.	86	92
56	644.	30	65		56	46509.	86	97
57	20733.	30	67		57	5881.	100	86
58	2000.	30	71		58	21.	90	98
59	101 210.	31	67		59	54342.	109	90
60	91138.	32	42		60	8140.	104	103
61	4 926 000.	32	44		61	20471.	103	108
62	3312.	32	119		62	920.	103	109
63	5556.	33	84					
64	4382.	33	108	SUM=20 982 270 512				
65	102 943.	34	67					
66	594.	35	24					
67	498 877.	35	32					
68	660 357.	35	34					
69	227 142.	35	44					
70	961 039.	35	72					
71	16796.	35	80					
72	53 494 056.	35	85					
73	178 928.	35	91					
74	7661.	35	109					
75	5776.	36	22					
76	12 657 265.	36	37					
77	158 200 927.	36	64					
78	5 352 968.	36	104					
79	92403.	37	67					
80	253 228.	38	41					
81	69783.	38	42					
82	1446.	38	48					
83	5350.	38	77					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
84	13122.	38	90					
85	118 999 000.	38	104					
86	24000.	39	24					
87	13667.	39	38					
88	539 387.	39	44					
89	96857.	39	64					
90	403 770.	39	72					
91	74 322 612.	39	85					
92	793 782.	39	91					
93	30934.	39	92					
94	27739.	40	17					
95	517 772.	40	38					
96	9611.	40	64					
97	597 699.	40	72					
98	10946.	40	77					
99	4577.	40	85					
100	9436.	40	90					
101	14558.	41	34					
102	61881.	41	67					
103	43562.	41	84					
104	4 530 307.	42	22					
105	99306.	42	40					
106	991.	42	43					
107	54197.	42	57					
108	43444.	42	79					
109	464 467 366.	42	80					
110	1 079 477.	42	87					
111	2 071 697.	42	90					
112	144 520.	42	91					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
113	3 063 842.	42	92					
114	102 506.	42	120					
115	11206.	43	31					
116	11206.	43	32					
117	29100.	43	67					
118	4000.	43	84					
119	2 622 487.	44	42					
120	15100.	44	43					
121	62855.	44	45					
122	156 054.	44	57					
123	3440.	44	76					
124	26023.	44	77					
125	7980.	44	78					
126	65 342 478.	44	80					
127	5375.	44	90					
128	72527.	44	106					
129	2059.	44	111					
130	412.	44	112					
131	26486.	45	43					
132	3700.	45	89					
133	30976.	46	41					
134	158 269.	46	80					
135	4697.	48	42					
136	1149.	49	3					
137	589 430.	49	54					
138	16021.	50	67					
139	65809.	51	67					
140	27442.	52	67					
141	1757.	53	60					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
142	2453 743.	53	109					
143	2604.	54	62					
144	18253.	54	116					
145	5076.	55	54					
146	1527.	56	80					
147	6245.	57	38					
148	586 000.	57	54					
149	4444.	57	85					
150	332 373.	59	54					
151	50000.	59	90					
152	21000.	61	54					
153	5441.	61	77					
154	385 622.	62	79					
155	19819.	63	20					
156	75000.	63	54					
157	23504.	64	23					
158	2880.	64	35					
159	26100.	64	52					
160	11882.	64	72					
161	24073.	64	79					
162	3 276 605 328.	64	80					
163	12423.	64	103					
164	146 253.	65	67					
165	150 000.	65	80					
166	36712.	65	84					
167	4803.	65	108					
168	11414.	66	38					
169	173 658.	66	44					
170	110 263.	66	64					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
171	311 439.	66	72					
172	3 713 807.	66	80					
173	202 270 791.	66	85					
174	4 630 973.	66	91					
175	1440.	66	108					
176	36.	67	79					
177	1 864 473.	68	79					
178	3 278 851.	68	80					
179	10276.	69	73					
180	3 819 880.	70	79					
181	2 611 856.	70	80					
182	133 718.	71	67					
183	37410.	71	79					
184	68791.	71	84					
185	210 168.	71	91					
186	24826.	71	108					
187	135 372.	72	21					
188	13798.	72	79					
189	2 543 725.	73	79					
190	1 251 442.	73	80					
191	2 852 000.	73	104					
192	899.	74	72					
193	17874.	74	79					
194	11 511 200.	75	111					
195	7971.	76	71					
196	418 369.	76	80					
197	326 861.	76	96					
198	633 835.	76	97					
199	150 297.	76	107					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
200	130 000.	76	108					
201	17966.	77	42					
202	97675.	77	57					
203	409.	77	71					
204	900 783 664.	77	75					
205	987.	77	79					
206	457.	77	81					
207	12435.	77	85					
208	45 041 501.	77	89					
209	1228.	77	90					
210	26 537 462.	77	91					
211	17 014 162.	77	92					
212	296 387.	77	99					
213	317 536.	77	109					
214	79715.	77	112					
215	1 472 037.	78	21					
216	2 637 549.	78	80					
217	114 723.	78	111					
218	6 754 404.	79	76					
219	97355.	79	78					
220	88 626 000.	79	80					
221	91475.	79	91					
222	67994.	79	97					
223	85 026 000.	79	104					
224	5897.	79	109					
225	33 130 080.	80	4					
226	141 890.	80	22					
227	549 530.	80	39					
228	13635.	80	40					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
229	144 339.	80	59					
230	43307.	80	62					
231	584 121.	80	72					
232	181 500 000.	80	75					
233	2 350 913 672.	80	85					
234	6878.	80	93					
235	43542.	80	114					
236	874.	80	119					
237	14 415 273.	81	79					
238	9 349 779.	81	80					
239	2 316 467.	81	85					
240	16.	81	90					
241	8 764 000.	81	104					
242	8277.	83	67					
243	59850.	83	86					
244	2173.	83	103					
245	42411.	85	20					
246	5 636 778.	85	42					
247	149 034.	85	79					
248	380 000.	85	95					
249	6931.	85	119					
250	41451.	86	21					
251	100 083.	86	44					
252	596 252.	86	59					
253	305.	86	65					
254	2548.	86	67					
255	147.	86	77					
256	109 531.	86	79					
257	7 743 654.	86	80					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
258	82597.	86	95					
259	135 664.	86	100					
260	7 214 000.	86	104					
261	819061.	86	109					
262	594.	88	79					
263	729 321.	89	79					
264	780 977.	89	88					
265	2 600 535.	89	98					
266	346 000.	89	106					
267	262 000.	89	111					
268	52814.	89	120					
269	1228.	90	31					
270	1228.	90	32					
271	149.	90	35					
272	58410.	90	66					
273	1657.	90	70					
274	36.	90	73					
275	4 429 628.	90	79					
276	214 432 526.	90	80					
277	28 173 181.	90	85					
278	121 280.	90	88					
279	6 831 101.	90	89					
280	22 870 522.	90	92					
281	32.	90	94					
282	98.	90	96					
283	31699.	90	100					
284	227 817 000.	90	104					
285	5 970 446.	90	109					
286	443 713.	91	88					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
287	99000.	91	106					
288	176 103.	91	107					
289	434 400.	91	111					
290	163 520.	91	120					
291	231 494.	92	44					
292	254 738.	92	57					
293	3688.	92	71					
294	581 482.	92	79					
295	1 751 258.	92	86					
296	146 170.	92	88					
297	32435.	92	98					
298	2 509 000.	92	106					
299	84979.	92	109					
300	238 300.	92	111					
301	20760.	92	120					
302	1208.	93	64					
303	558 961.	93	72					
304	2006.	93	91					
305	960.	94	65					
306	37410.	94	79					
307	13781.	94	80					
308	64043.	94	84					
309	169 115.	94	91					
310	43000.	94	106					
311	26328.	94	108					
312	62246.	94	114					
313	3559.	95	79					
314	300 000.	95	89					
315	119 315.	95	99					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
316	250 304.	95	109					
317	300 000.	95	118					
318	119 363.	96	74					
319	25942.	96	79					
320	536 516.	97	74					
321	129 626.	97	86					
322	99208.	97	96					
323	8748.	97	100					
324	1654.	97	108					
325	129 340.	98	74					
326	8767.	98	79					
327	1676.	98	90					
328	48633.	98	96					
329	104 735.	99	80					
330	4524.	99	85					
331	187 257.	99	89					
332	226 856.	99	91					
333	26817.	99	92					
334	3366.	99	96					
335	73032.	100	88					
336	89 706 283.	102	64					
337	41 008 409.	102	72					
338	961 758.	102	80					
339	74662.	103	42					
340	730.	103	72					
341	11 407 535.	103	79					
342	52 100 651.	103	80					
343	106 797.	103	81					
344	19 325 860.	103	104					

Продолжение табл. 8

K1	S1	N1	N2		k2	S2	N1	N2
345	155 230.	103	107					
346	33759.	104	35					
347	6148.	104	40					
348	559 298 634.	104	42					
349	265.	104	72					
350	681.	106	79					
351	10000.	107	113					
352	16916.	108	79					
353	44337.	108	84					
354	4379.	108	103					
355	6640.	108	107					
356	21489.	109	20					
357	3569.	109	24					
358	4 960 145.	109	42					
359	10 220 814.	109	44					
360	28150.	109	54					
361	45576.	109	66					
362	5 193 600.	109	80					
363	2728.	109	103					
364	28150.	110	54					
365	15400.	111	77					
366	13761.	111	79					
367	15120.	112	42					
368	62798.	112	107					
369	1639.	113	84					
370	22207.	113	94					
371	4260.	114	27					
372	781 986.	114	67					
373	59345.	114	79					

Продолжение табл. 8

K1	S1	N1	N2	k2	S2	N1	N2
374	155 196.	114	103				
375	6659.	115	80				
376	1 497 289.	116	80				
377	11081.	119	33				
378	15330.	119	38				
379	65510.	119	44				
380	107 025.	119	54				
381	87204.	119	65				
382	3640.	119	103				

Причина этого понятна. У мелких предприятий много связей замыкается внутри региона. Поэтому для них региональный взаимозачёт эффективен в процентном отношении, но денежные суммы при этом невелики. У крупных предприятий и суммы крупные, но основная часть связей лежит вне региона; региональный центр их не может обрабатывать. Поэтому чем больше охваченный регион, тем эффективней взаимозачёт. Наиболее эффективен он в масштабе всей страны.

2.3.5. Расчёт погашения. Взаимозачёт погашения долгов методом фракционирования был произведён в два этапа. Предлагаемая программа DEFRA не предусматривала погашения встречных долгов. Поэтому применять её непосредственно к табл. 7 было нельзя. Чтобы не переделывать программу, её применили к табл. 8. Результат такого фракционного взаимозачёта представлен табл. 9.

Параметры расчёта: $N = 120$; $B = 0,00000$; $\text{eps} = 0,70\text{E-}06$; $Q = 1000$; $J = 3\,686\,500$; $\text{nvud} = 2$; REZULTAT.

Интегральные результаты: $I_q = 454$; $K_a = 0,418\,314$; $K_i = 0,0$; $S = 0$; $S_a = 6\,114\,952\,679$; $C = 10\,512\,462\,066$; $C_i = 10\,512\,460\,934$; $D = 10\,512\,462\,066$; $D_i = 10\,512\,460\,934$; $P_q = 0,2041$.

Таблица 9. Взаимозачёт по программе DEFRA для данных из табл. 8

N	Rn	Bn	sn	bnf	cn	cnf	Dn	dnf	Pqn
1	0.	0	-78590.	0.	78590	78590.	0.	0.	0,0000
2	0.	0.	-299 826.	0.	299 826.	299 826.	0.	0.	0,0000
3	0.	0.	-13003.	0.	14152.	14152.	1149.	1149.	0,0000
4	0.	0.	-472 139 765.	0.	506 298 392.	506 298 387.	34 158 627.	34 158 622.	0,0000
5	0.	0	-58670.	0.	58670.	58670.	0.	0.	0,0000
6	0.	0.	-109 484.	0.	109 484	109 484.	0.	0.	0,0000

Продолжение табл. 9

N	Rn	Bn	sn	bni	cn	cni	Dn	dni	Pqn
7	0	0	-283 247.	0.	283 247	283 247.	0.	0.	0.0000
8	0.	0.	458 167.	0.	2315.	2315.	460 482.	460 482.	0.0000
9	0.	0	-118 000	0.	118 000.	118 000	0.	0.	0.0000
10	0.	0.	-254 651.	0.	254 651.	254 651.	0.	0.	0.0000
11	0.	0.	-35320.	0.	35320.	35320.	0.	0.	0.0000
12	0.	0	-897 615	0.	897 615	897 615	0.	0.	0.0000
13	0.	0.	-102 710.	0.	102 710.	102 710.	0.	0.	0.0000
14	0.	0.	0.	0.	0.	0.	0.	0.	0.0000
15	0.	0.	0.	0.	0.	0.	0.	0.	0.0000
16	0.	0	-5034	0.	5034.	5034	0.	0.	0.0000
17	0.	0.	-64 944 655.	0.	64 972 394.	64 972 394.	27739.	27739.	0.0000
18	0.	0.	-40625.	0.	40625.	40625.	0.	0.	0.0000
19	0.	0.	3100.	0.	4400.	4400.	7500.	7500.	0.0000
20	0.	0.	-11 926 153.	0.	12 073 780.	12 073 778.	147 627.	147 625.	0.0000
21	0.	0	13 881 413.	0.	14960	14956.	13 896 373	13 896 369	0.0003
22	0.	0.	2 002 849.	0.	2 693 437.	2 693 434.	4 696 286.	4 696 283.	0.0000
23	0.	0.	112 935.	0.	0.	0.	112 935.	112 935.	1.0000
24	0.	0.	-295 765 154	0.	295 793 317	295 793 317.	28163.	28163.	0.0000
25	0.	0	-55363.	0.	55363.	55363.	0.	0.	0.0000
26	0.	0.	-65018.	0.	65018.	65018.	0.	0.	0.0000
27	0.	0.	-218 540.	0.	222 800.	222 800.	4260.	4260.	0.0000
28	0.	0	0	0.	0	0	0.	0.	0.0000
29	0.	0.	-11909.	0.	11909.	11909.	0.	0.	0.0000
30	0.	0.	-43188.	0.	43188.	43188.	0.	0.	0.0000
31	0.	0.	-66448.	0.	101 210.	101 210.	34762.	34762.	0.0000
32	0.	0	-4 506 622.	0.	5 020 450.	5 020 450	513 828.	513 828.	0.0000
33	0.	0.	1143	0.	9938.	9938.	11081.	11081.	0.0000
34	0.	0.	571 972.	0.	102 943.	102 943.	674 915.	674 915.	0.0000
35	0.	0.	-56 008 662.	0.	56 045 450.	56 045 450.	36788.	36788.	0.0000
36	0.	0.	118 296 664.	0.	176 216 936.	176 216 936.	294 513 600.	294 513 600.	0.0000
37	0.	0	12 564 862.	0.	92403	92403.	12 657 265	12 657 265	0.0000
38	0.	0.	-118 575 895.	0.	119 341 929.	119 341 929.	766 034.	766 034.	0.0000
39	0.	0.	-75 675 479.	0.	76 225 009.	76 225 009.	549 530.	549 530.	0.0000
40	0.	0.	-1 058 691.	0.	1 177 780	1 177 780.	119 089.	119 089.	0.0000
41	0.	0	166 513.	0.	120 001.	120 001	286 514.	286 514.	0.0000
42	0.	0.	97 133 757.	-4.	475 657 653.	475 657 370.	572 791 410.	572 791 131.	0.0000
43	0.	0.	-12935.	0.	55512.	55512.	42577.	42577.	0.0000
44	0.	0	-51 803 662.	0.	68 316 790.	68 316 788.	16 513 128	16 513 126	0.0000
45	0.	0.	32669.	0.	30186.	30186.	62855.	62855.	0.0000
46	0.	0.	-189 245.	0.	189 245.	189 245.	0.	0.	0.0000

Продолжение табл. 9

N	Rn	Bn	sn	bnf	cn	cnf	Dn	dnf	Pqn
47	0.	0	0.	0.	0.	0	0.	0.	0,0000
48	0.	0.	-3251.	0.	4697.	4697	1446	1446.	0,0000
49	0.	0	-590 579	0.	590 579.	590 579	0.	0.	0,0000
50	0.	0.	-16021.	0.	16021.	16021.	0.	0.	0,0000
51	0.	0.	-65809.	0.	65809.	65809.	0.	0.	0,0000
52	0.	0	-1342	0.	27442	27442.	26100.	26100.	0,0000
53	0.	0.	-2 455 500.	0.	2 455 500.	2 455 500.	0.	0.	0,0000
54	0.	0.	3 185 575.	0.	20857.	20857.	3 206 432.	3 206 432.	0,0000
55	0.	0.	-5076.	0.	5076.	5076.	0.	0.	0,0000
56	0.	0	-1527.	0.	1527.	1527.	0.	0.	0,0000
57	0.	0.	-34025.	0.	596 689	596 689.	562 664.	562 664.	0,0000
58	0.	0.	0.	0.	0.	0.	0.	0.	0,0000
59	0.	0.	358 218.	0.	382 373.	382 373.	740 591.	740 591.	0,0000
60	0.	0.	3042.	0.	0.	0.	3042.	3042.	0,0000
61	0.	0	-26441.	0.	26441	26441.	0.	0.	0,0000
62	0.	0.	-339 711.	0.	385 622.	385 622.	45911.	45911.	0,0000
63	0.	0.	-94819.	0.	94819.	94819.	0.	0.	0,0000
64	0.	0.	-2 738 498 900	0.	3 276 706 190.	3 276 706 186	538 207 290.	538 207 286.	0,0000
65	0.	0	-248 655	0.	337 768.	337 768	89113.	89113.	0,0000
66	0.	0.	-211 118 852.	0.	211 223 785.	211 223 785.	104 933.	104 933.	0,0000
67	0.	0.	2 284 366.	0.	36.	36.	2 284 402.	2 284 402.	0,0028
68	0.	0	-5 143 324.	0.	5 143 324	5 143 324	0.	0.	0,0000
69	0.	0.	-10276.	0.	10276.	10276.	0.	0.	0,0000
70	0.	0.	-6 430 079.	0.	6 431 736.	6 431 736.	1657.	1657.	0,0000
71	0.	0.	-460 845.	0.	474 913.	474 913.	14068.	14068.	0,0000
72	0.	0	243 842 714	0.	149 170.	149 168	243 991 884	243 991 882.	0,0000
73	0.	0.	-6 636 855.	0.	6 647 167.	6 647 167	10312.	10312.	0,0000
74	0.	0.	831 464.	0.	18773.	18769.	850 237.	850 233.	0,0002
75	0.	0.	1 070 772 464.	-1.	11 511 200.	11 511 155.	1 082 283 664.	1 082 283 619.	0,0000
76	0.	0.	5 090 511.	0.	1 667 333.	1 667 332.	6 757 844.	6 757 843.	0,0000
77	0.	0	-990 128 390	0.	990 201 584	990 201 559	73194.	73170.	0,0000
78	0.	0.	-4 118 974.	0.	4 224 309.	4 224 309.	105 335.	105 335.	0,0000
79	0.	0.	-139 929 024.	-1.	180 669 125.	180 669 086.	40 740 101.	40 740 062.	0,0000
80	0.	0.	1 634 020 213.	0.	2 567 071 868.	2 567 071 558	4 201 092 081.	4 201 091 771.	0,0000
81	0.	0	-34 738 281.	0.	34 845 535.	34 845 535.	107 254.	107 254.	0,0000
82	0.	0.	0.	0.	0.	0.	0.	0.	0,0000
83	0.	0.	-70300.	0.	70300.	70300.	0.	0.	0,0000
84	0.	0	268 640	0.	0	0	268 640.	268 640.	1,0000
85	0.	0.	2 765 283 885.	0.	6 215 154.	6 214 870.	2 771 499 039.	2 771 498 755.	0,0000
86	0.	0.	-14 904 559.	0.	16 845 293.	16 845 292.	1 940 734.	1 940 733.	0,0000

Продолжение табл. 9

N	Rn	Bn	sn	bnf	cn	cnf	Dn	dnf	Pqn
87	0.	0	1 079 477.	1.	0.	0	1 079 477.	1 079 476.	1.0000
88	0.	0.	1 564 578	0.	594.	592	1 565 172.	1 565 170.	0.0032
89	0.	0	47 591 912.	0.	4 771 647	4 771 640	52 363 559	52 363 552.	0.0000
90	0.	0.	-508 587 671.	0.	510 740 221.	510 740 220	2 152 550.	2 152 549.	0.0000
91	0.	0.	31 668 549.	0.	1 316 736.	1 316 734.	32 985 285.	32 985 283.	0.0000
92	0.	0	42 077 992.	0.	5 854 304	5 854 302	47 932 296	47 932 294	0.0000
93	0.	0.	-555 297.	0.	562 175.	562 175.	6878.	6878.	0.0000
94	0.	0.	-394 644.	0.	416 883.	416 883.	22239.	22239.	0.0000
95	0.	0.	-510 581.	0.	973 178.	973 161.	462 597.	462 580.	0.0000
96	0.	0	332 861.	0.	145 305.	145 304	478 166.	478 165.	0.0000
97	0.	0.	-73923.	0.	775 752	775 751.	701 829.	701 828.	0.0000
98	0.	0.	2 444 554.	0.	188 416.	188 412.	2 632 970.	2 632 966.	0.0000
99	0.	0.	-137 853.	0.	553 553.	553 553.	415 702.	415 700.	0.0000
100	0.	0.	103 079.	0.	73032.	73032.	176 111.	176 111.	0.0000
101	0.	0	0	0.	0	0	0.	0.	0.0000
102	0.	0.	-131 676 450.	0.	131 676 450.	131 676 450.	0.	0.	0.0000
103	0.	0.	-82 990 926.	0.	83 171 465.	83 171 465.	180 539.	180 539.	0.0000
104	0.	0.	-65 367 714.	0.	559 338 806.	559 338 786	493 971 092.	493 971 072.	0.0000
105	0.	0	0.	0.	0.	0	0.	0.	0.0000
106	0.	0.	3 068 846.	0.	681.	679.	3 069 527.	3 069 525.	0.0026
107	0.	0.	541 068.	0.	10000.	10000.	551 068.	551 068.	0.0000
108	0.	0	137 297.	0.	72272	72272.	209 569.	209 569.	0.0000
109	0.	0.	-10 534 907.	0.	20 476 071.	20 476 066.	9 941 164.	9 941 159.	0.0000
110	0.	0.	-28150.	0.	28150.	28150.	0.	0.	0.0000
111	0.	0.	12 533 521.	0.	29161.	29115.	12 562 682.	12 562 636.	0.0016
112	0.	0	2209.	0.	77918.	77918.	80127.	80127	0.0000
113	0.	0.	-13846.	0.	23846	23846.	10000.	10000.	0.0000
114	0.	0.	-892 684.	0.	1 000 787.	1 000 787.	108 103.	108 103.	0.0000
115	0.	0.	-6659.	0.	6659.	6659.	0.	0.	0.0000
116	0.	0.	-1 479 036.	0.	1 497 289.	1 497 289.	18253.	18253.	0.0000
117	0.	0	0	0.	0	0	0.	0.	0.0000
118	0.	0.	300 000.	5.	0.	0.	300 000.	299 995.	1.0000
119	0.	0.	-266 764.	0.	289 790.	289 790.	23026.	23026.	0.0000
120	0.	0.	339 600	0.	0	0.	339 600.	339 600.	1.0000

В этом расчёте сходимость итераций была гораздо более медленной, чем для модельных примеров. Причина в том, что в медленных примерах сеть долгов предприятий была гораздо более густой и разветвлённой, а здесь долговых связей очень мало. Это приближает знаменатель сходимости к 1. Поэтому сейчас 450 итераций дали погрешность

$\approx 10^{-10}$ (1 рубль на 10 млрд), что можно видеть по ненулевым значениям b'_n в табл. 9. Однако скорость расчёта оставалась более высокой — менее 1 сек. на PC Pentium IV.

Результат табл. 9 надо объединить с погашением встречных долгов, сумма которых составляла около 21 млрд руб. Окончательные результаты представлены в табл. 2. Там же остаток долгов, не погашенный фракционированием, частично погашен методом переадресации. Видно следующее:

- начальная информация очень неполна — она содержала сведения лишь примерно о половине действительных долгов — так что разумные результаты можно получить лишь по дополнительной информации;
- метод фракционирования в данном примере погасил 80 % внутренних долгов системы, что демонстрирует его эффективность;
- метод переадресации гасит 42 % от остатка, что ещё более 8 % от начальной задолженности.

Таким образом, наиболее эффективна следующая комбинация. Сначала погасить всё, что возможно, методом фракционирования, так как он является юридически безукоризненным, и его результаты обязательны для предприятий-участников. К оставшейся части долга применить метод переадресации, необязательный для предприятий, и выдать его как рекомендацию.

2.3.6. Внешняя задолженность. Предприятия с № 121 по № 7754 сами не давали списков долгов, но сведения об их отношениях с исходными предприятиями содержатся в табл. 4. Если какое-то из внешних предприятий является только дебитором или только кредитором, то метод фракционирования не уменьшает его долгов. Однако фракционное погашение возможно, если внешнее предприятие есть одновременно в кредиторском и дебиторском списках.

Общее число записей по внешним долгам равно $13\,584 - (1040 + 580) = 11\,964$. Беглый просмотр показал, что не менее 22 % (возможно, до 36 %) этих записей таковы, что внешнее предприятие имеет одновременно кредиторскую и дебиторскую задолженности.

Значит методом фракционирования можно ещё существенно уменьшить остаточные долги (конечно, не внутренние, а внешние). Однако количественный расчёт не проводился.

§ 3. Переадресация

Найдено решение задачи погашения взаимных долгов системы большого числа предприятий, при котором сумма остаточных долгов минимальна. Оно не единственно, и среди множества этих решений можно выбрать такие, что число остаточных долгов менее числа самих предприятий, причём, выполнены условия, существенно облегчающие последующие банковские переводы денег. Алгоритм является сверхбыстрым.

3.1. Уменьшение остаточного долга. В § 1 показано, что для замкнутой системы с хорошо развитыми связями метод фракционирования высокоэффективен и даёт малый остаточный долг. Однако даже в этом случае желательно уменьшить насколько можно величину этого остатка. Кроме того, в определённых ситуациях эффективность метода фракционирования заметно понижается. К таким ситуациям относятся: 1) незамкнутая система, в которой значительная часть связей осуществляется с внешними предприятиями; 2) система с «перекосами» экономики, включающая монополистов, искусственно завышающая цену на свою продукцию; 3) система с малоразветвлёнными связями, даже при большом числе предприятий. В этих случаях остаточный долг после фракционирования может быть значительным, так что его уменьшение становится насущным.

В этих случаях надо сначала максимально уменьшить долг методом фракционирования насколько возможно. Затем к полученному состоянию системы целесообразно применять метод переадресации подробно приводимый ниже.

3.2. Метод переадресации. Нетрудно показать на примере, что способ замкнутых цепочек имеет принципиальный недостаток. Рассмотрим цепочку, в которой каждое предприятие с 1-го до $(N-1)$ -е должно следующему одинаковую сумму x (рис. 19а), но N -е не должно 1-му. Эта цепочка разомкнута, и указанный способ неприменим. В то же время очевидно следующее решение этой задачи: 1-е должно N -му сумму x , то есть долг переадресовывается, а остальные долги гасятся (рис. 19б).

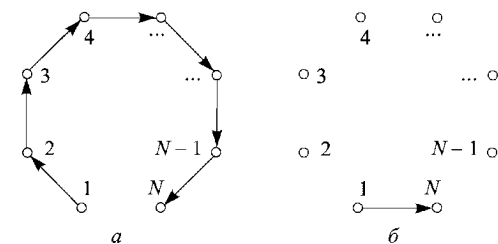


Рис. 19. Переадресация долга обращением цепочки

Это решение приемлемо с экономической точки зрения, даже если 1-е и N -е предприятия никогда ранее не знали друг о друге. Оно соответствует обычному вексельному обращению: 1-е предприятие, не имея свободных средств, заплатило векселем на предъявителя свой долг второму, второе этим же векселем 3-му и так далее, а N -е предъявило этот вексель 1-му.

Общая математическая формулировка переадресации долгов дана в [1-2]. Пусть x_{nm} есть величина долга m -го предприятия n -му ($1 \leq n$,

$m \leq N$); тем самым мы условились, что $x_{nm} > 0$, если m -е предприятие должно n -му, и $x_{nm} < 0$ в обратном случае. Очевидно,

$$x_{nm} = -x_{mn} \quad \text{при } n \neq m \quad \text{и} \quad x_{nn} = 0, \quad (14)$$

то есть матрица долгов X кососимметрична. Введём балансы кредитов и долгов каждого предприятия (сальдо)

$$s_n = \sum_{m=1}^n x_{nm}, \quad 1 \leq n \leq N. \quad (15)$$

В силу (14) достаточно задавать только поддиагональную часть матрицы X ($1 \leq m < n \leq N$), содержащую $N(N-1)/2$ элементов, а (15) переписать в следующем виде:

$$s_n = \sum_{m=1}^{n-1} x_{nm} - \sum_{m=n+1}^N x_{mn}. \quad (16)$$

Заметим, что

$$\sum_{n=1}^N s_{nm} = 0; \quad (17)$$

последнее означает замкнутость системы (поскольку рассматриваются долги только между предприятиями данной системы).

Взаимозачёт долгов формально является заменой исходной матрицы X на новую матрицу Z . В [1–2] сформулировано и обосновано следующее

Утверждение. Любая кососимметричная матрица Z , сохраняющая вектор сальдо $s(Z) = s(X)$, соответствует экономически допустимому зачёту взаимных долгов.

Существует бесконечное множество допустимых зачётов. Среди них имеются как погашение долгов в цепочках, так и переадресации. Для оценки качества получаемых решений введём векторы кредитов $c(Z)$ и долгов $d(Z)$ предприятий:

$$c_n(Z) = \frac{1}{2} \sum_{m=1}^N (|z_{nm}| + z_{nm}), \quad (18)$$

$$d_n(Z) = \frac{1}{2} \sum_{m=1}^N (|z_{nm}| - z_{nm}); \quad (19)$$

долги здесь определены как положительные величины. Выражения (18), (19) также можно преобразовать к виду, аналогичному (16). Очевидно, сальдо связаны с ними соотношениями

$$s_n(Z) = c_n(Z) - d_n(Z), \quad (20)$$

Критерием качества выберем сумму долгов всех предприятий

$$D(Z) = \sum_{n=1}^N d_n(Z) = \sum_{n=1}^N c_n(Z) = \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^{n-1} |z_{nm}|; \quad (21)$$

последние равенства следуют из (14)–(17). Очевидно, чем меньше $D(Z)$, тем лучше проведён зачёт долгов с экономической точки зрения. Однако, сделать $D(Z)$ сколь угодно малым нельзя. В самом деле, поскольку $c_n \geq 0$ и $d_n \geq 0$, то из (7) следует, что $|s_n(Z)| \leq c_n(Z) + d_n(Z)$. Суммируя это неравенство от 1 до N и сравнивая с (21), получим

$$D(Z) \geq D_{\min} \equiv \frac{1}{2} \sum_{n=1}^N |s_n(Z)| \geq 0. \quad (22)$$

Величина s_n инвариантна относительно процедуры зачёта долгов. Значит, $D_{\min}$ тоже есть инвариант системы. Если хотя бы одно сальдо ненулевое, то $D_{\min} > 0$

3.3. Идеальное решение. В [1–3] из допустимых зачётов выбирались те, которые минимизируют эвклидову норму матрицы Z (при определении этой нормы вводились веса). На примерах было видно, что величина $D(Z)$ при этом близка к $D_{\min}$, особенно для вариантов метода, описанных в [3]. Алгоритм нахождения таких решений оказался быстрым. Однако, можно построить лучшие решения.

Определение. Назовём допустимое решение идеальным, если $D(Z) = D_{\min}$.

Идеальное решение имеет одно важное свойство. Равенство $|s_n(Z)| = c_n(Z) + d_n(Z)$ возможно только если $c_n = 0$ и/или $d_n = 0$. Из (20) следует, что на идеальном решении такие равенства должны выполняться при всех n . В этом случае $c_n = s_n$ и $d_n = 0$ при $s_n > 0$, $c_n = 0$ и $d_n = -s_n$ при $s_n < 0$, $c_n = d_n = 0$ при $s_n = 0$; то есть у предприятий с положительным сальдо не остаётся никаких долгов, при отрицательном сальдо — только долги, при нулевом — ни долгов, ни кредитов. Это действительно наиболее целесообразно экономически.

Идеальное решение не единственно. Такие решения составляют многогранник в некотором подпространстве $[N(N-1)/2]$ -мерного пространства, образованного поддиагональными переменными z_{nm} . Некоторые из них могут выражаться простыми явными формулами, например

$$z_{nm} = \frac{s_n(s_m + |s_m|)}{D(X)}, \quad n \neq m, \quad z_{nn} = 0. \quad (23)$$

Решение (23) соответствует тому, что долг любого предприятия с отрицательным сальдо распределяется между всеми кредиторами пропорционально величинам их сальдо. Недостатком решения (23) является то, что матрица z_{nm} содержит много ненулевых элементов, то есть между предприятиями оставляется много связей.

Наибольший практический интерес представляет подмножество решений, построение которого описано ниже.

Обозначим через P число строго положительных сальдо s_n , а через Q — число строго отрицательных; очевидно, $0 \leq P + Q \leq N$, а $N - P - Q$ есть число нулевых сальдо. Возьмём все положительные сальдо, произведём произвольную их перестановку и после этого обозначим их через α_p , $1 \leq p \leq P$. Аналогично возьмём модули отрицательных сальдо, произведём некоторую перестановку и обозначим через β_q , $1 \leq q \leq Q$.

Дальнейшая процедура легко поясняется геометрически с помощью рис. 20. Из величин α_p строится линейка итоговых кредитов, а из β_q — итоговых долгов предприятий. Длины этих линеек одинаковы в силу (17). Проведём границы, соответствующие узлам обеих линеек (пунктир). Они разобьют эти линейки на R отрезков. Если среди внутренних узлов нет совпадающих, то $R = P + Q - 1$, в противном случае $R < P + Q - 1$. Длины этих отрезков примем за величины кредитов долгов $|z_{nm}|$ между соответствующими предприятиями. Остальные долги z_{nm} , в том числе все, относящиеся к предприятиям с нулевыми сальдо, положим равными нулю.

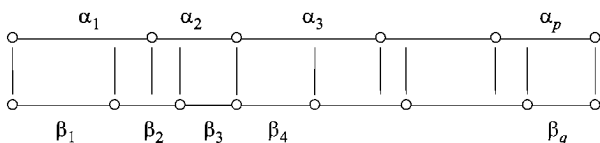


Рис. 20. Вариант идеальной переадресации

Таким образом, построено идеальное решение, содержащее не более $N - 1$ не нулевых поддиагональных элементов z_{nm} . Описанный алгоритм настолько быстр и прост, что не предъявляет серьёзных требований к компьютеру.

Проиллюстрируем его примером системы с $N = 10$ и матрицей X с 45 ненулевыми поддиагональными элементами (табл. 10); начальная сумма долгов равна $D(X) = 3729$. После упорядочения по убыванию $|s_n|$ получается матрица Z всего с 7 ненулевыми элементами (табл. 11) и суммой долгов $D(Z) = 62$, многократно меньшей начальной.

3.4. Обсуждение.

3.4.1. Эффективность. Под «предприятием» можно понимать любое юридическое (или физическое) лицо, в том числе банки и госбюджет. Это позволяет рассмотреть и зачесть все виды долгов, имеющиеся в стране.

Предложенный алгоритм позволяет произвести погашение всех взаимных долгов системы предприятий, которые в принципе могут быть погашены. Это делается практически мгновенным расчётом, минуя

Таблица 10

$n \backslash m$	1	2	3	4	5	6	7	8	9
2	-25								
3	-1	-20							
4	4	25	-2						
5	25	-450	25	30					
6	-15	150	-30	20	-928				
7	3	-40	3	3	5	25			
8	1	-22	-2	-2	4	-15	5		
9	10	332	-15	-25	498	-800	-10	20	
10	1	-25	-2	1	-20	15	-1	-3	30

Таблица 11

$n \backslash m$	1	2	3	4	5	6	7	8	9
2	2								
3	0	0							
4	0	0	0						
5	0	0	0	0					
6	0	0	0	0	-28				
7	1	0	0	0	0	0			
8	0	-7	0	0	0	0	0		
9	0	-18	0	0	-2	0	0	0	
10	0	0	0	0	0	4	0	0	0

длительную процедуру последовательных банковских платежей и не требуя наличия свободных оборотных средств у предприятий.

Непогашаемая доля остаточных долгов $\delta = D_{\min}/D(X)$ зависит от исходной системы. Она в принципе может колебаться от 0 (если все предприятия имеют нулевые сальдо) до 1 (если предприятия в долгу у одного монополиста, взвинтившего цену на свою продукцию). Значение $\delta \approx 1$ свидетельствует о катастрофической несбалансированности

цен. При нормальных соотношениях цен должно быть $\delta \ll 1$, то есть предложенный зачёт ликвидирует кризис неплатежей.

Последнее рассуждение явно подразумевает, что система замкнута. Однако, реально это не так: предприятия области имеют связи по всей стране, а предприятия страны — за рубежом. Чем больше взятая система, тем ближе она к замкнутой, тем меньше (в среднем) величина δ и более эффективен зачёт долгов. Его наиболее выгодно делать для всего мира или, хотя бы, для группы стран с тесно связанной экономикой.

Рассмотренный алгоритм производит переадресацию остаточных долгов на конкретных юридических лиц, а не на клиринговую палату или другой центральный орган. Тем самым, этот метод не требует выдачи кредитов Центрального Банка и не вызывает инфляцию. Всё это обуславливает высокую экономическую эффективность метода.

3.4.2. Перестановки. Различные перестановки предприятий кредиторов и предприятий-должников на линейках рис. 20 могут принести дополнительные преимущества.

Среди предприятий есть крупные и малые. Если сохранить на рис. 20 исходный порядок предприятий, то вероятны случаи, когда большое предприятие останется должником или кредитором многих мелких предприятий. Если же упорядочить каждую линейку по убыванию $|s_n|$, то такой случай мало вероятен, и в среднем каждое предприятие останется связанным с небольшим числом предприятий.

Целесообразно сгруппировать предприятия так, чтобы кредиторы и должники, лежащие в линейках друг против друга, имели корреспондентские счета в одном и том же банке. Это существенно ускоряет последующие расчёты между ними: перевод денег внутри банка выполняется много быстрее межбанковского платежа. Из этих групп можно составить укрупнённые группы по банкам, охваченным той или иной автоматизированной системой межбанковских расчётов.

Это должно облегчить предприятиям выплату значительной части остаточных долгов за счёт собственных оборотных средств.

3.4.3. Быстродействие. Предложенный алгоритм очень прост. Основное число операций приходится на вычисление сальдо ($\approx N^2$) и на перестановки в линейках ($< N^2$). Если вычислять сальдо по формуле (16), вводя матрицу X по одной строке и столбцу, то достаточно оперативной памяти $\approx 2N$ чисел. Таким образом, для всех предприятий России можно провести расчёт на персональном компьютере.

Расчёт целесообразно производить в целых числах. Надо иметь в виду, что 32-разрядный компьютер при расчётах в целых числах позволяет без специальных ухищрений оперировать с числами не более $2^{31} \approx 2 \cdot 10^9$.

Л. В. Кузьмина составила демонстрационную программу такого расчёта DECAN-2 (Debts Canceling) на языке FORTRAN.

3.4.4. Мировая экономика. На Западе звенья экономики хорошо согласованы. Однако, и там во время биржевых паник возника-

ют кратковременные, но серьёзные кризисы неплатежей, приводящие к значительным потерям. Если располагать введённой в строй программой на мощном компьютере, такой кризис можно ликвидировать в считанные минуты. В обычной же обстановке такая программа позволит уменьшить размер оборотных средств, что принесёт ощутимую прибыль.

§ 4. Банковские платежи

4.1. Введение. Рассмотрена задача межбанковского клирингового расчёта при отрицательных сальдо части банков. Предложен переход от дискретных переменных к непрерывным. После этого задачу можно решить методом линейного программирования. Найдены такие непрерывные переменные, в которых решение всегда существует, если суммарное сальдо всех банков неотрицательно. Это решение находится двойственным симплекс-методом.

Разработан эффективный алгоритм перехода от непрерывных переменных к формированию пакетов из дискретного числа платёжных поручений. Расчёт на примерах показал, что этот алгоритм позволяет зачесть значительную часть платежей даже при существенных нехватках денежных средств у части банков.

Сначала был разработан быстрый алгоритм, эффективный в ситуации, когда сальдо почти всех банков много меньше оборота средств, т. е. сумма производимых каждым банком платежей почти равна сумме встречных платежей из других банков.

Однако, после этого нашлись три примера из практики с резко несбалансированными платежами (например, один из банков посылает платежей на вчетверо большую сумму, чем адресовано ему). Эта ситуация потребовала существенной доработки метода.

Ниже изложен модернизированный метод.

4.2. Исходная постановка. Имеется N банков — участников клиринга. В каждый банк платёжные поручения от клиентов (имеющие разный приоритет) могут поступать непрерывно в течение рабочего дня. Они немедленно обрабатываются и (если они приняты банком) заносятся в базу данных этого банка. В определённые часы, несколько раз в течение рабочего дня, состояние баз данных всех банков записывается в базу данных клирингового центра. После положенной перепроверки данных производится клиринговый расчёт, и его результаты немедленно сообщаются банкам.

Сформулируем разумные требования к алгоритму клирингового расчёта. Будем ориентироваться на некоторые цифры по г. Москве: число банков $N \approx 20-100$, полное число платёжных поручений в день $\approx 300\,000$; банки достаточно сильно различаются по мощности, так что суточное количество платёжных поручений из n -го банка в m -й может колебаться от ≈ 5000 до ≈ 30 . Предполагается, что в целом со-

стояние системы нормальное (подробнее об этом будет сказано далее). Но некоторые банки могут испытывать нехватку свободных средств и, тем самым, полного взаимозачёта всех платёжных поручений не произойдёт.

Желательно, чтобы расчёт проводился раз в час. Для этого полный комплекс программ межбанковского расчёта должен требовать всего 20–30 мин. расчёта. Тогда на алгоритм клирингового расчёта можно отвести 1–3 мин. расчёта, т. е. $\approx 10^8$ арифметических операций на недорогом компьютере. Видно, что алгоритм должен быть очень быстрым.

Будем требовать, чтобы платежи с высоким приоритетом обязательно выполнялись в первый же расчёт. Допустим, что небольшой процент платёжных поручений с низким приоритетом может остаться неоплаченным в первый расчёт даже для банков с достаточными свободными средствами; но в следующем расчёте он должен быть зачтён (разумеется, для «благополучных» банков). Очевидно, при расчёте раз в час такая задержка небольшой части платежей не вызовет серьёзных последствий.

4.3. Обозначения. Информация об исходном состоянии системы содержит следующие величины:

n ($1 \leq n \leq N$) или m — индекс, обозначающий номер банка;

$K_{nm} \geq 0$ — число платёжных поручений из n -го банка в m -й;

a_{nmk} ($k \leq K_{nm}$) — величина k -го платёжного поручения из n -го банка в m -й;

A_n — свободные средства n -го банка до расчёта;

C_n — резерв, который оставляется на счёте n -го банка на первом этапе расчётов;

ДНР — допустимое непрерывное решение;

ОНР — оптимальное непрерывное решение.

Случай, когда у всех банков достаточно свободных средств для взаимного расчёта по всем платежам, маловероятен. Поэтому считаем, что в каждом пакете платежей n -го банка в m -й проходит не полное число поручений K_{nm} , а уменьшенное.

Тогда исходная сумма платежей из n -го банка в m -й равна

$$A_{nm} = \sum_{k=1}^{K_{nm}} a_{nmk}, \quad (24)$$

а уменьшенная на некоторую долю x_{nm} составляет

$$A_{nm}(1 - x_{nm}).$$

Тогда сальдо n -го банка после расчёта будет равно:

$$B_n(X) = A_n - \sum_{m=1}^{N'} [A_{nm}(1 - x_{nm}) - A_{mn}(1 - x_{mn})]. \quad (25)$$

Оно складывается из его начальных средств, поступлений от других банков и его платежей. Здесь и далее штрих у суммы означает, что

она берётся по всем индексам, кроме $m = n$. Символом X обозначается матрица (x_{nm}) . Очевидно, диагональные элементы x_{nn} в ней отсутствуют.

4.4. Непрерывные переменные. На первом этапе клирингового расчёта решение проводится в *непрерывных* переменных. Требуется найти матрицу X такую, что все банки после расчёта не затронут свои резервы:

$$B_n(X) \geq C_n \quad \text{при } 1 \leq n \leq N, \quad (26)$$

а сами элементы матрицы удовлетворяют очевидным ограничениям

$$0 \leq x_{nm} \leq 1 \quad \text{при } 1 \leq n \leq N, \quad 1 \leq m \leq N \quad (n \neq m). \quad (27)$$

Любая матрица, удовлетворяющая неравенствам (26), (27), является решением в непрерывных переменных. Будем называть его *допустимым непрерывным решением* (ДНР).

При суммировании (25) по n все платежи банков взаимно сокращаются, так что

$$\sum_{n=1}^N B_n(X) = \sum_{n=1}^N A_n. \quad (28)$$

Сравнивая с (3), получаем необходимое условие существования ДНР:

$$\sum_{n=1}^N A_n \geq \sum_{n=1}^N C_n. \quad (29a)$$

Достаточным условием существования ДНР является

$$A_n \geq C_n \quad \text{для всех } n. \quad (29б)$$

Если оно выполнено, то существует по крайней мере тривиальное ДНР $x_{nm} \equiv 1$ (его смысл — ни один платёж не произведён). Однако использование (29б) неудобно, так как конечные лимиты C_n ограничены начальными активами A_n , которые для «хороших» банков с положительным сальдо невелики.

С другой стороны, можно подобрать такие тестовые примеры, в которых условие (29а) не является достаточным для существования ДНР. Вопрос о необходимом и достаточном условии существования ДНР труден для решения. Однако на практике при большом N и большом суммарном обороте системы решение при выполнении (29а) удаётся найти.

Введём величины сальдо банков при условии полного прохождения всех платежей ($x_{nm} \equiv 0$):

$$B_n \equiv B_n(X=0) = A_n + \sum_{m=1}^{N'} (A_{mn} - A_{nm}). \quad (30)$$

Очевидно, если для всех n выполняется условие

$$B_n \geq C_n \quad \text{при } 1 \leq n \leq N, \quad (31)$$

то $x_{nm} \equiv 0$ является ДНР. Такое решение можно считать *идеальным* (все платежи прошли). Нетрудно видеть, что (31) является необходимым и достаточным условием существования идеального решения.

Если условие (31) не выполнено, а (29) выполнено, то существует бесконечно много ДНР. Тогда можно ставить вопрос о нахождении оптимального (в каком-то смысле) решения.

4.5. Опимальное решение. ДНР, минимизирующее линейную целевую функцию

$$L(x) \equiv \sum_{n=1}^N \sum_{m=1}^{N'} c_{nm} x_{nm} = \min, \quad c_{nm} > 0 \quad (32)$$

будем считать *оптимальным* непрерывным решением (ОНР). Очевидно, задача на минимум (32) при линейных ограничениях (26), (27) есть задача линейного программирования.

В общем случае ОНР может быть не единственным — достигаться на ребре или грани выпуклого многогранника условий. Даже если оно единственно, т.е. достигается в вершине, возможна ситуация, когда выходящие из этой вершины ребро или грань окажутся почти параллельными плоскостям $L(x) = \text{const}$. Тогда ОНР будет плохо обусловленным, и небольшие вариации величин c_{nm} , A_{nm} , A_n или даже ошибки округления при вычислениях могут заметно изменить решение (в частности, вызвать переброс решения на соседнюю вершину).

Заметим, что если условие (29) является точным равенством, то ошибки округления могут превратить его в неравенство обратного знака, когда ДНР отсутствует. Поэтому в алгоритм всегда нужно вводить небольшую страховку от ошибок округления.

Коэффициенты c_{nm} можно задать из разных соображений. Рассмотрим два случая, которые можно считать противоположными крайностями. В первом полагаем

$$c_{nm} = A_{nm}. \quad (33)$$

Такой выбор соответствует минимизации денежной суммы непрошедших платежей. На первый взгляд это требование разумно. Однако, при этом фактически предпочтение отдаётся крупным платежам. Может быть целиком отброшен пакет из многих мелких платёжных поручений ради некоторого увеличения процента прохождения большого пакета. Целевую функцию, соответствующую (33), обозначим $L_A(x)$, а соответствующее ОНР через X_A .

Противоположная крайность — выбор одинаковых весов. Без ограничения общности можно положить

$$c_{nm} \equiv 1, \quad n \neq m. \quad (34)$$

Это означает равноправие малых и больших пакетов, минимизируется суммарный процент непрохождения платежей во всех пакетах. Такой выбор должен больше устраивать многочисленных некрупных клиентов

банков. Соответствующую целевую функцию и ОНР обозначим через $L_C(x)$ и X_C .

Сравнение весов (33) и (34) проводилось на численных примерах, описанных далее в п. 10. Оно показало, что алгоритм нахождения ОНР двойственным симплекс-методом очень устойчив при весах (34), а при весах (33) могут возникать осложнения. Более детальный анализ показал, что веса (33) приводят к параллельности некоторых рёбер двойственного многогранника условий и плоскости целевой функции. Если процесс приведёт на такое ребро, то алгоритм иногда сбивается (на практике это встречается чрезвычайно редко).

В принципе, небольшие изменения весов по сравнению с (33) должны ликвидировать такое вырождения. Однако, это делать нецелесообразно. Оказалось, что хотя $L_C(X_C) > L_A(X_A)$, но это превышение незначительно. Поэтому лучше использовать веса (11) и находить решение X_C ; при этом алгоритм устойчив, целевая функция $L_C(x)$ точно минимизируется, а целевая функция $L_A(x)$ близка к своему минимуму.

Рекомендация: вести расчёты только с $c_{nm} \equiv 1$.

4.6. Задача линейного программирования (ЗЛП). Очевидно, именно такова задача минимизации линейной функции (32) при линейных ограничениях (26), (27). Эта задача содержит $N(N-1)$ неизвестных x_{nm} , $1 \leq n, m \leq N$, $n \neq m$. Переменные x_{nn} вообще не вводятся.

Если какой-то пакет платежей отсутствует, т. е. его $A_{nm} = 0$, то соответствующую переменную x_{nm} можно, в принципе, также исключить из расчёта. Однако это заметно усложняет логику алгоритма. Выгоднее оказалось полагать в этом случае $x_{nm} = 0$, записывая это условие в виде неравенства $0 \leq x_{nm} \leq 0$.

После этого исходная задача переписывается следующим образом:

$$L(x) \equiv \sum_{n=1}^N \sum_{m=1}^{N'} c_{nm} x_{nm} = \min, \quad c_{nm} \geq 0 \quad (\text{обычно } c_{nm} \equiv 1), \quad (35a)$$

$$\sum_{m=1}^{N'} (A_{nm} x_{nm} - A_{mn} x_{mn}) + (B_n - C_n) \geq 0, \quad 1 \leq n \leq N, \quad (35б)$$

$$-x_{nm} + 1 \geq 0 \quad \text{при } A_{nm} \neq 0, \quad 1 \leq n, \quad m \leq N, \quad n \neq m, \quad (35в)$$

$$-x_{nm} \geq 0 \quad \text{при } A_{nm} = 0, \quad 1 \leq n, \quad m \leq N, \quad n \neq m, \quad (35г)$$

$$-x_{nm} \geq 0, \quad 1 \leq n, \quad m \leq N, \quad n \neq m, \quad (35д)$$

где

$$B_n \equiv B_n(X=0) = A_n + \sum_{m=1}^{N'} (A_{mn} - A_{nm}), \quad 1 \leq n \leq N. \quad (35е)$$

Это классическая неканоническая форма записи задачи линейного программирования. Неравенства (35д) выделяют первый координатный угол, а неравенства (35б)–(35г) описывают многогранник ограничений.

Если в разрешающей строке нет положительных элементов, то задача *не имеет решения*.

Затем из положительных элементов разрешающей строки выбираем элемент по наименьшему двойственному отношению, т.е. отношению соответствующих элементов строки C и найденной разрешающей строки ($c_j/a_{kn,j}$). Обозначим его индекс как km (в терминологии симплекс-метода найден разрешающий столбец с номером km).

Элемент $a_{kn,km}$, стоящий на пересечении *разрешающих* строки и столбца, называется *разрешающим* элементом. С этим *разрешающим* элементом делается один шаг обыкновенных жордановых исключений:

— меняем местами kn -й элемент вспомогательного столбца и km -й элемент вспомогательной строки.

— выполняем преобразование элементов симплексной таблицы по следующим формулам:

$$\begin{aligned}
 & \text{для } i \neq kn, \quad j \neq km \\
 & a_{i,j} = a_{i,j} - \frac{a_{i,km}a_{kn,j}}{a_{kn,km}}, \quad b_i = b_i - \frac{a_{i,km}b_{kn}}{a_{kn,km}}, \quad c_j = c_j - \frac{c_{km}a_{kn,j}}{a_{kn,km}}, \\
 & \text{для } i = kn, \quad j \neq km \\
 & a_{i,j} = -\frac{a_{i,j}}{a_{kn,km}}, \quad b_{kn} = -\frac{b_{kn}}{a_{kn,km}}; \\
 & \text{для } i \neq kn, \quad j = km \\
 & a_{i,j} = a_{i,j}a_{kn,km}, \quad c_{km} = \frac{c_{km}}{a_{kn,km}}; \\
 & \text{для } i = kn, \quad j = km \\
 & a_{kn,km} = \frac{1}{a_{kn,km}}.
 \end{aligned} \tag{37}$$

При этих преобразованиях для защиты от неустойчивости и от ошибок округления величины $\{a_{ij}\}$ и $\{c_j\}$, меньшие по модулю 10^{-15} , заменяются нулями. На этом процедура однократного применения «жорданового исключения» закончена.

4.8. Резерв. Рассмотрим, зачем и как вводятся C_n . По банковским правилам решение X допустимо, если сальдо всех банков неотрицательны $B_n(X) \geq 0$. Однако, расчёт в непрерывных переменных даёт доли x_{nm} , которые (если они не равны 0 или 1) не могут, вообще говоря, быть точно представлены в виде суммы каких-то платёжных поручений из данного пакета.

Поэтому на втором этапе алгоритма приходится формировать пакет, то есть находить какие-то наборы a_{nmk} с суммами, возможно более близкими к $A_{nm}x_{nm}$. Эти суммы могут быть больше или меньше искомым величин; соответственно изменяются $B_n(X)$, причём в любую сторону. Если первоначально было удовлетворено ограничение $B_n(X) \geq 0$,

то после формирования пакета некоторые величины $B_n(X)$ окажутся отрицательными, и полученное решение станет недопустимым.

Поэтому на стадии расчёта в непрерывных переменных надо требовать $B_n(X) \geq C_n$, где C_n — некоторые положительные величины, вообще говоря, не малые. Их выбор целесообразно сделать исходя из некоторой характерной величины единичного платежа в каждом пакете a_{nm} .

Считаем, что при формировании одного пакета происходит отклонение от искомой суммы на $\pm a_{nm}/2$. Все пакеты, посылаемые и получаемые n -м банком, формируются независимо. Поэтому за резерв естественно взять удвоенный (по правилу 2σ) статистический разброс сальдо $B_n(X)$, имеющий вид

$$C_n^0 = \left[\sum_{m=1}^{N'} (a_{mn}^2 + a_{nm}^2) \right]^{1/2}, \quad 1 \leq n \leq N. \quad (38)$$

Определение величин a_{nm} тоже неоднозначно. Рассмотрим три естественных определения:

$$a_{nm} = \max_k a_{nmk}, \quad (39a)$$

$$a_{nm} = \left(\sum_k \frac{a_{nmk}^2}{N} \right)^{1/2}, \quad (39б)$$

$$a_{nm} = \sum_k \frac{a_{nmk}}{N}. \quad (39в)$$

Выбор (39a) является очень надёжным, но может приводить к нереально большим значениям C_n . Выбор (39в) не даёт достаточного запаса надёжности алгоритма. Поэтому мы рекомендуем (39б).

Однако, если формально ввести C_n^0 по формулам (38), (39б), то может получиться

$$\sum_n C_n^0 > \sum_n A_n,$$

и решения в непрерывных переменных не будет существовать. Поэтому была предложена процедура введения *коэффициента запаса* q :

$$C_n = \frac{qC_n^0}{Q}, \quad 0 < q \leq 1, \quad (40a)$$

$$Q = (1 + \varepsilon) \frac{\sum_{n=1}^N C_n^0}{\sum_{n=1}^N A_n}, \quad 0 < \varepsilon \ll 1. \quad (40б)$$

(здесь $\varepsilon \approx 10^{-6}$ страхует от ошибок округления). Расчёты проводят с различными значениями q , начиная от $q = 1$ и постепенно уменьшая. Наиболее приемлемый результат удобно оценивать визуально.

Очевидно, при любом q в указанных пределах существует ОНР. Чем меньше q , тем больше средняя доля прохождения платежей. Но при этом уменьшается резерв на формирование пакета и повышается вероятность того, что после формирования пакета решение станет недопустимым. Поскольку этот процесс вероятностный, зависимость результатов от q будет не вполне регулярной. Именно поэтому нужна визуальная оценка.

Замечание. Формировать требуется только те пакеты, для которых строго получилось $0 < x_{nm} < 1$. Это небольшая часть всех возможных пакетов. Только их необходимо включать в сумму (38). Однако, заранее неизвестно, какие пакеты будут формироваться, так что приходится брать полную сумму.

4.9. Формирование пакета. Рассмотрим, как целесообразнее набрать из K_{nm} величин a_{nmk} сумму, достаточно близкую к заданной величине $A_{nm}x_{nm}$ ($0 < x_{nm} < 1$). Очевидно, перебор всех возможных комбинаций даёт наилучшее приближение. Но число таких комбинаций $2^{K_{nm}}$ при часто встречающихся $K_{nm} \approx 40\text{--}50$ настолько велико, что объём вычислений невыполним.

Кроме того, нетрудно привести пример, когда даже наилучшая комбинация даёт плохой результат. На рис. 21 изображён случай, когда пакет состоит из одного большого платёжного поручения и одного или нескольких очень маленьких, а $x_{nm} \approx 1/2$. Любая комбинация даст здесь отличие суммы от требуемого значения почти на половину величины пакета!

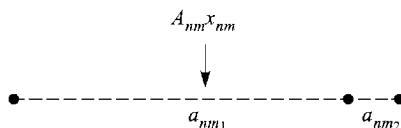


Рис. 21. Случай плохой укладки пакета

Поэтому целесообразно ограничиться каким-либо несложным и быстрым алгоритмом. Хорошо зарекомендовала себя следующая процедура.

1. *Упорядочение.* Расположим платежи a_{nmk} в порядке убывания их величин. Если необходимо учесть приоритеты, то сначала ставят платежи с высшим приоритетом в порядке убывания величины, а после них — платежи с низким приоритетом в аналогичном порядке. Требуемое число операций $\sim K_{nm}^2$ невелико.

2. *Формирование.* Опишем получение сумм с недобором s_{nmk}^- и с перебором s_{nmk}^+ (соответствующие пакеты назовём *нижним* и *верхним*). Начинаем со значений $s_{nm0}^- = 0$ и $s_{nm0}^+ = A_{nm}$ и «пустых» пакетов. Далее поочерёдно прибавляем платежи a_{nmk} :

а) $s_{nmk} = s_{nm,k-1}^- + a_{nmk}$, $k = 1, \dots, K_{nm}$;

б) если $s_{nmk} \leq A_{nm}x_{nm}$, то зачисляем платёж a_{nmk} в нижний пакет, верхний пакет не меняем, за новые суммы берём $s_{nmk}^- = s_{nmk}$, $s_{nmk}^+ = s_{nm,k-1}^+$;

в) если $s_{nmk} > A_{nm}x_{nm}$, то платёж a_{nmk} не зачисляем в нижний пакет, а его сумму сохраняем: $s_{nmk}^- = s_{nm,k-1}^-$;

г) если выполнено в) и $s_{nmk} < s_{nm,k-1}^+$, то новый состав верхнего пакета есть текущий нижний пакет плюс платёж a_{nmk} , а его новая сумма равна $s_{nmk}^+ = s_{nmk}$;

д) если в) выполнено, но $s_{nmk} \geq s_{nm,k-1}^+$, то состав и сумма верхнего пакета не изменяются;

е) выполнив шаг $k = K_{nm}$, производим последнюю проверку: если $s_{nmK}^- + s_{nmK}^+ \geq 2A_{nm}x_{nm}$, то в качестве окончательного берётся нижний пакет, в обратном случае — верхний.

Требуемое число операций $\sim K_{nm}$ ничтожно.

З а м е ч а н и е. Для сформированного пакета целевая функция $L(X)$ и процент прохождения платежей могут быть как хуже, так и лучше, чем для ОНР. В последнем нет ничего недопустимого: при нахождении ОНР резервы C_n считаются неприкосновенными, а при формировании пакета они могут расходоваться.

4.10. Тестирование. Сначала данный алгоритм проверялся на пакетах, искусственно построенных с помощью датчика случайных чисел по некоторому закону распределения. Однако эти тесты оказались чересчур лёгкими. В таких распределениях есть платежи самых разных размеров, больших средних и малых; ими нетрудно хорошо заполнить требуемую сумму. Сформированные пакеты совпадали с исходными значениями $A_{nm}x_{nm}$ на 3–4 знака.

Настоящую трудность представляют три реальные подборки из практики. Среди них было немало пакетов, похожих по структуре на рис. 21. Основные параметры этих подборок представлены в табл. 12. Каждая из них содержит 10 банков (на большее число не рассчитана предложенная методическая программа) и, соответственно, 90 пакетов платёжных поручений. Видны следующие особенности.

Во-первых, платёжные балансы отдельных банков практически не скомпенсированы. Например, имеются банки, которые должны выплатить всем другим банкам в совокупности в 4–6 раз большую сумму, чем получают от них. Даже в среднем по системе платежи несбалансированы от 38% для примера 3 до 45% для примера 2. Очевидно, если у соответствующих банков нет значительных свободных средств A_n , то средний процент непрохождения платежей в системе

$$\delta = \frac{\sum'_{n,m} A_{nm}x_{nm}}{\sum'_{n,m} A_{nm}} \quad (41)$$

будет большим, несколько десятков процентов.

Таблица 12

Номер подборки	Пример 1	Пример 2	Пример 3
Сумма всех заявленных платежей (млн руб.)	237,618,0	95,382,6	45,157,4
Сумма отрицательных балансов (млн руб.)	99,820,0	43,083,2	17,336,9
Среднеарифметический пакет (млн руб.)	2,640,2	1,059,8	501,7
Полное число платёжных поручений	3970	1529	3259
Среднеарифметический платёж (млн руб.)	59,9	62,4	13,9
Максимальный платёж (млн руб.)	9,000,0	7,000,0	3,766,5
Число платежей крупнее среднего пакета	370	97	454

Во-вторых, имеется много больших платёжных поручений, превышающих величину среднеарифметического пакета. Наибольший же платёж в несколько (до 6) раз превышает величину среднего пакета. Есть пакеты, содержащие ≈ 20 поручений, причём $\approx 95\%$ денежной суммы пакета составляет один из платежей.

Очевидно, если при формировании окончательного пакета значение $A_{nm}x_{nm}$ «режет пополам» такое крупное платёжное поручение, то отличие набранной суммы пакета от требуемой будет большим. Такая ситуация благополучно разрешается только при наличии достаточно больших резервов C_n (а их можно ввести лишь при наличии свободных средств A_n).

В примерах 1–3 свободные средства A_n не указывались. Поэтому они были заданы в данных методических расчётах следующим образом: для банков с положительным платёжным балансом $A_n = 0$, для банков с отрицательным балансом: $A_n = r \times$ (величину дефицита). Очевидно, при $r = 1$ все платежи проходят, а при $r = 0$ решение может отсутствовать, так что интересен случай $0 < r < 1$, особенно небольшие r .

Было проведено более 50 различных методических расчётов. Рассмотрим важнейшие результаты.

Целевая функция. Расчёт первой стадии (т.е. с непрерывными переменными) был проведён для примера 1 с тремя различными весами целевой функции: (33), (34) и ещё одним промежуточным между ними. Были взяты маленький коэффициент свободных средств $r = 0,1$ и умеренный коэффициент резерва $q = 0,25$.

Эти расчёты привели к трём различным ОНР, отвечающим разным вершинам многогранника условий. Однако, коэффициент прохождения платежей (41) оказался во всех случаях одинаковым $\delta = 41,18\%$. Правда, на других тестах величины δ были не одинаковыми, но достаточно

близкими (например, в пределах от 55 % до 59 % для примера 2 и от 40 % до 42 % для примера 3).

Поэтому выше была дана рекомендация использовать веса для целевой функции $c_{nm} \equiv 1$.

Ненулевые пакеты. В этих расчётах ОНР содержало только 7–10 отличных от 0 и 1 пакетов из 90 возможных. Это не случайно. Можно строго доказать, что в ОНР будет не более N (число условий на окончательные позиции банков) пакетов с $0 < x_{nm} < 1$. Формировать надо только пакеты с $0 < x_{nm} < 1$. Видно, что их число невелико.

Заметим также, что если некоторое $x_{nm} \neq 0$, то симметричное $x_{mn} = 0$. Если бы это было не так, то во встречных пакетах $n \rightarrow m$ и $m \rightarrow n$ можно было бы провести клиринговый зачёт, то есть рассмотренное решение не было бы оптимальным.

Число пакетов с $x_{nm} = 1$ может быть значительным при плохой сбалансированности системы.

Влияние резерва. Были проведены расчёты с разными коэффициентами свободных средств r и резерва q . Очевидна следующая зависимость ОНР от них: чем больше r , тем меньше величина непрохождения платежей δ ; при $r = 0$ она максимальна, при $r = 1$ обращается в нуль. Наоборот, чем больше q , тем больше δ (поэтому «неулучшаемое» ОНР при данном r получается при $q = 0$). Такая зависимость иллюстрируется левой частью табл. 13.

Таблица 13

$\begin{smallmatrix} q \\ r \end{smallmatrix}$	Оптимальное непрерывное решение				Сформированные пакеты				% непрошедших поручений			
	1,00	0,75	0,50	0,25	1,00	0,75	0,50	0,25	1,00	0,75	0,50	0,25
0,5	31,4	27,3	25,6	24,4	31,4	25,7	25,3	24,1		7,9	13,5	13,7
0,3	39,0	36,1	36,1	32,1	37,8	38,2	37,1	30,1	19,1	16,8	16,4	12,8
0,1	48,6	46,5	46,0	45,2	48,4	44,9	44,2	47,3	26,2	24,2	17,6	19,0

В средней части табл. 13 приведена зависимость для сформированных (согласно п. 4.9) пакетов. Она не вполне монотонна, поскольку процедура формирования оперирует с дискретными величинами. Сами значения δ иногда больше, а иногда меньше, чем соответствующие значения для ОНР.

После формирования возникали ситуации, когда выделенного резерва не хватало, и сальдо одного или нескольких банков оказывались отрицательными. Такое решение неприемлемо. Эти случаи лежат ниже пунктирной линии. Видно, что наиболее сильно влияет в этом смысле коэффициент свободных средств: при $r = 0,5$ решение существует при всех рассмотренных q , при $r = 0,1$ его не оказалось ни при каком q .

Однако, влияние q также существенно: чем больше q , тем выше вероятность того, что формирование пакета оставит решение приемле-

мым (не приведёт к отрицательному сальдо). Расчёты других тестов показывают, что в целом поведение близко не к горизонтальной прямой, а к наклонной слева-вверх.

Интересен также средний процент прохождения числа платёжных поручений. Для ОНР он совпадает с процентом прохождения денежных сумм (левая часть табл. 13). Но наш алгоритм формирования пакета использует значительное количество мелких платежей для заполнения «пустот». Поэтому процент прохождения числа платежей (правая часть табл. 13) при этом сильно уменьшается по сравнению с ОНР.

4.11. Тактика расчётов. В реальных расчётах нет коэффициента r , а заданы действительные свободные средства A_n . Поэтому единственным свободным коэффициентом остаётся q . Можно рекомендовать следующую тактику.

Производим серию расчётов с несколькими значениями q , включая $q = 1$ и $q = 0$, например, $q = 1,0; 0,8; 0,6; 0,4; 0,2$ и 0 . Выберем из данной серии те варианты, когда решение со сформированными пакетами приемлемо (т. е. все сальдо неотрицательны). Из них оставим вариант с наименьшим значением δ .

ОНР при $q = 0$ показывает верхнюю грань для процента прохождения платежей ($\delta_m = \min$). Если итоговое значение δ близко к δ_m , то алгоритм формирования пакета сработал хорошо. Если часто будут встречаться ситуации с заметно отличающимися δ и δ_m , то алгоритм желательно усовершенствовать.

Если при каком-то q не найдено приемлемого решения, это означает, что свободные средства A_n у каких-то банков недостаточны, то есть им требуются кредиты.

Кредиты. Эффективность кредитов может сильно зависеть от того, какому банку они предоставлены. В некоторых случаях кредит обслужит единственный платёж. Тогда сумма неплатежей в системе уменьшится на величину кредита. Но если кредит инициирует длинную цепочку платежей, он особенно эффективен: уменьшение суммы неплатежей станет во много раз больше.

Данная программа даёт возможность найти места для наиболее целесообразного предоставления кредитов. Увеличим A_n некоторого банка на величину ΔA_n и вычислим соответствующее уменьшение суммы неплатежей. Совершая это для всех банков поочерёдно, найдём тот банк, для которого это уменьшение максимально. Ему и следует в первую очередь предоставлять кредит.

Можно разработать алгоритм нахождения наиболее целесообразной величины такого кредита, найти следующий банк для кредитования и т. д. Данный метод решения задачи клирингового расчёта настолько быстр, что позволяет проделывать все эти эксперименты на компьютере за небольшое время.

Применение двойственного симплекс-метода к задаче линейного программирования может натолкнуться на сложности, если задача вырождена. Геометрически это означает, что одинаковые значения целевой функции достигаются более чем в одной вершине двойственного многогранника условий. В этом случае одно или несколько c_j на каком-либо шаге оказываются равными нулю для $a_{kn,j} > 0$.

З а м е ч а н и е. Так как c_j изменяется на каждом шаге (37), то $c_j = 0$ возможно даже при $c_j^{\text{нач}} \neq 0$ для всех j . Например, для $c_{nm} = A_{nm}$ получаем некоторые $c_j = 0$ уже после первого шага.

Тогда в качестве разрешающего столбца берём как раз столбец с $c_j = 0$, $a_{kn,j} > 0$, так как именно на нём достигается наименьшее двойственное отношение.

Легко заметить, что в этом случае значение целевой функции f не изменяется. При этом гипотетически возможна ситуация, когда после ряда преобразований можно вернуться к уже встречавшемуся ранее набору базисных переменных IBZ (значению элементов вспомогательного столбца). Таким образом, мы сталкиваемся с заикливанием двойственного симплекс-метода.

Оказывается, заикливание можно предотвратить, если внести в симплекс-алгоритм определённые уточнения для выбора вводимых и выводимых базисных переменных IBZ.

Доказано, что заикливания не возникает (то есть после конечного числа шагов либо получим ответ, что данное значение целевой функции оптимально, либо перейдём к другому значению целевой функции f , либо окажется, что задача не имеет решения), если при выборе разрешающих строки и столбца пользоваться следующим алгоритмом, реализованным в программе SIMPLD1. Для упрощения программы используется предположение, справедливое для решаемой здесь прикладной задачи, что $\{c_j\} \geq 0 \quad j = 1, \dots, m$.

4.12. SUBROUTINE SIMPLD1 (n, m, A, Nmx, Mmx, B, F, IBZ, ISV, idd). Входные данные:

n, m — размерность решаемой задачи;

Nmx, Mmx — максимально возможные размеры задачи;

A(Nmx, Mmx) — матрица коэффициентов $\{a_{i,j}\}$ неравенств;

B(Nmx) — вектор свободных членов $\{b_i\}$ неравенств;

F(Mmx) — вектор коэффициентов $\{c_j\}$ целевой функции.

Рабочие массивы:

IBZ(Nmx) — номера расширенных переменных, являющиеся базисными;

ISV(Mmx) — номера расширенных переменных, являющиеся свободными.

Выходные данные:

F(Mmx) — значения решения $\{x_{i,j}\}$.

1. Подготовительная часть:

— проверяется, что все элементы массива F больше 0;

— подготавливаются массивы IBZ и ISV%

$$\text{ISV}(j) = j, \quad j = 1, m,$$

$$\text{IBZ}(i) = m + i, \quad i = 1, n.$$

2. Ищем разрешающую строку:

Если все элементы массива В положительны, то решение найдено и идём к п. 5.

Если нет, то среди всех отрицательных элементов В ищем минимальный, который определяет разрешающую строку. Если минимальных несколько, то в качестве разрешающей строки берём строку с наименьшим индексом. Обозначим этот индекс как kn (в терминологии симплекс-метода — найдена разрешающая строка с номером kn).

3. Ищем разрешающий столбец по наименьшему двойственному отношению, т.е. для всех положительных элементов разрешающей строки $\{A(kn, j), j = 1, m\}$ находим отношение элементов F -строки к элементам данной разрешающей строки: $\{F(j)/A(kn, j), j = 1, m\}$.

Рассмотрим множество G_0 , состоящее из тех j , на которых достигается минимум соотношения $\{F(j)/A(kn, j), A(kn, j) > 0, j = 1, m\}$.

Замечание. Если в разрешающей строке нет положительных элементов, то задача не имеет решения.

Если множество G_0 содержит более одного элемента, то составляем множество G_1 . В G_1 входят такие $j \in G_0$, что на них достигается минимум отношения $\{A(1, j)/A(kn, j), A(kn, j) > 0, j \in G_0\}$. Если в G_1 входит более одного элемента, то составляем множество G_2 . В G_2 входят такие $j \in G_1$, что на них достигается минимум отношения $\{A(2, j)/A(kn, j), A(kn, j) > 0, j \in G_1\}$, и так далее, пока не будет получено множество G_t , состоящее из одного элемента.

Строго доказано, что такое t ($1 \leq t \leq n$) существует. Единственный элемент множества G_t и определяет разрешающий столбец $km = j, j \in G_t$.

4. Элемент, стоящий на пересечении разрешающих строки и столбца, называется разрешающим элементом — $A(kn, km)$. С этим разрешающим элементом производим один шаг обыкновенных жордановых исключений:

— Обмен содержимого ячеек $\text{ISV}(km)$ и $\text{IBZ}(kn)$

$$\text{для } i \neq kn, \quad j \neq km$$

$$a_{i,j} = a_{i,j} - \frac{a_{i,km}a_{kn,j}}{a_{kn,km}}, \quad b_i = b_i - \frac{a_{i,km}b_{kn}}{a_{kn,km}}, \quad f_j = f_j - \frac{f_{km}a_{kn,i}}{a_{kn,km}};$$

$$\text{для } i = kn, \quad j \neq km$$

$$a_{i,j} = -\frac{a_{i,j}}{a_{kn,km}}, \quad b_{kn} = -\frac{b_{kn}}{a_{kn,km}};$$

$$\text{для } i \neq kn, \quad j = km$$

$$a_{i,j} = \frac{a_{i,j}}{a_{kn,km}}, \quad f_{km} = \frac{f_{km}}{a_{kn,km}};$$

для $i = kn, j = km$

$$a_{kn,km} = \frac{1}{a_{kn,km}}$$

— для защиты от неустойчивости и от ошибок округления величины $\{a_{i,j}\}$ и $\{f_j\}$, меньшие по модулю 10^{-15} , заменяем нулями. Возвращаемся к п. 2 данной программы.

5. Решение найдено: Массив $\{B(i), i = 1, n\}$ содержит решение в расширенном базисе. Массив $\{IBZ(i), i = 1, n\}$ содержит номера расширенных переменных, являющиеся базисными. Надо выбрать только реальные переменные $\{x_{i,j}\}$ с номерами, не превосходящими m .

если $j = IBZ(i) < m - F(j) = B(i)$

для остальных $F(j) - F(j) = 0$

Включение в алгоритм дополнительных действий по формированию множеств $G_0, G_1, \dots, G_t$ — достаточно трудоёмкая процедура. Она усложняет программу и увеличивает время расчёта.

Однако практические расчёты показывают, что явление заикливания настолько редкое, что использование описанных алгоритмов вряд ли оправдано. Какую же тактику расчётов выбирать?

Рекомендации. При решении задачи следует из G_0 выбирать минимальное значение, то есть $km = \min_{j \in G_0}(j)$. Этот алгоритм был использован в основной программе двойственного симплекс-метода.

Если же заикливание всё-таки возникает, то для расчётов придётся воспользоваться описанным выше правилом преодоления заикливания.

§ 5. Оценка долгов

Ниже рассмотрена проблема продажи долгов, требующая объективной оценки того, какая часть долга может быть возвращена. Сформулированы и исследованы математические постановки возникающих при этом задач. Эти подходы могут быть применены для задачи взаимозачёта долгов.

Построен эффективный алгоритм расчёта вероятности возврата долгов, составлена программа для персонального компьютера и произведены методические расчёты с числом предприятий до 640. Проанализированы результаты этих расчётов. Показано, что скорость алгоритма достаточна для расчёта системы порядка миллиона предприятий, так что алгоритм пригоден для включения в федеральные системы.

5.1. Проблема. Продажа долгов (разумеется, со скидкой) по частным соглашениям между юридическими лицами происходит не первый год. О подобных сделках кое-что сообщала пресса. Так, во время финансового кризиса в августе 1998 г некоторые предприятия скупали за полцены долговые обязательства тех банков, которым сами были

должны, и затем возвращали долг банкам этими обязательствами по номиналу.

Появлялись также официальные извещения о принудительной продаже долгов с аукциона (было ли это связано с процедурой банкротства, не сообщалось). Обычно стартовые цены составляли от 10 до 50 суммы долга. Из этих примеров видно, что подобные сделки носят спекулятивный характер, и объективной оценки реальной стоимости долга при них не устанавливается.

Объективную оценку получить нелегко. Рассмотрим факторы, определяющие её. В первую очередь рассматривают кредиторскую задолженность данного предприятия. Во вторую — наличие имущества, на которое можно наложить взыскание и легко реализовать. В третью — дебиторскую задолженность. Её учесть особенно сложно, ибо надо оценить, насколько вероятен возврат этих долгов, т. е. проделать для всех должников данного предприятия ту оценку, которую надо сделать для него самого. Такую задачу покупатель долга (или участник аукциона) не сможет решить.

Однако, если собраны те сведения, которые необходимы для взаимозачёта долгов по всей стране или, хотя бы, по достаточно большой группе тесно связанных предприятий, то можно предложить неплохие способы решения задачи оценки долгов. Правда, для этого потребуются некоторые дополнительные сведения, а также отдельные экспертно подобранные константы.

5.2. Обозначения. Пусть собрана информация для взаимозачёта долгов. Предприятия, участвующие во взаимозачёте, назовём *внутренними*; не участвующие, но связанные с внутренними, назовём *внешними*. Внешние предприятия сами не сообщают о себе сведений. О них узнаётся только то, что сообщают их контрагенты внутренние предприятия. Введём следующие обозначения:

n, m — индексы только внутренних предприятий;

r — индексы только внешних предприятий;

B_n — рыночная (не балансовая) стоимость основных фондов внутреннего предприятия;

b_n — остаток оборотных средств на счёте внутреннего предприятия;

c_{nm}^i — кредиторская задолженность внутреннего предприятия внутреннему;

c_{nr}^0 — кредиторская задолженность внутреннего предприятия внешнему;

d_{nm}^i — дебиторская задолженность внутреннего предприятия внутреннему;

d_{nr}^0 — дебиторская задолженность внутреннего предприятия внешнему;

p_n — доля долга $0 \leq p_n \leq 1$, которую реально может выплатить данное внутреннее предприятие (для всех его долгов она прини-

мается одинаковой, т.е. приоритеты пока не учтены). Эту долю можно рассматривать так же, как вероятность выплаты долгов;
 q_r — доля долга, которую реально может выплатить внешнее предприятие (вероятность выплаты его долгов);
 ξ_n — доля основных фондов ($0 \leq \xi_n \leq 1$), которую данное предприятие вынуждено срочно продать для расплаты с долгами. Реально даже при банкротстве и полной продаже величина ξ_n будет существенно меньше 1, ибо полную стоимость при этом не выручишь.

5.3. Балансы. Предполагаем, что все предприятия частично выплачивают свои долги. Тогда каждое внутреннее предприятие выплачивает своим внутренним кредиторам сумму

$$C_n^i = p_n \sum_m c_{nm}^i, \quad (42a)$$

своим внешним кредиторам

$$C_n^0 = p_n \sum_r c_{nr}^0, \quad (42б)$$

то есть всего выплачивает

$$C_n^t = C_n^i + C_n^0. \quad (42в)$$

Оно получает от своих внутренних дебиторов

$$D_n^i = \sum_m p_m d_{nm}^i, \quad (42г)$$

от внешних дебиторов

$$D_n^0 = \sum_r q_r c_{nr}^0, \quad (42д)$$

то есть всего

$$D_n^t = D_n^i + D_n^0. \quad (42е)$$

Аналогичные величины при 100%-х выплатах долгов ($p_n 1$, $q_r 1$) назовём идеальными выплатами и обозначим соответственно

$$\bar{C}_n^i, \bar{C}_n^0, \bar{C}_n^t, \bar{D}_n^i, \bar{D}_n^0, \bar{D}_n^t. \quad (43)$$

Нетрудно заметить, что

$$C_n^i = p_n \bar{C}_n^i, C_n^0 = p_n \bar{C}_n^0, C_n^t = p_n \bar{C}_n^t. \quad (44)$$

При расплате с учётом продажи части основных фондов в итоге у предприятия на счёте должна оставаться неотрицательная величина. Следовательно, для всех n должны выполняться условия

$$\xi_n B_n + b_n + \sum_m p_m d_{nm}^i + \sum_r q_r d_{nr}^0 \geq p_n \left(\sum_m \bar{c}_{nm}^i + \sum_r \bar{c}_{nr}^0 \right), \quad (45)$$

причём $0 \leq p_n, q_r, \xi_n \leq 1$. Это основные уравнения балансов, из которых следуют все дальнейшие рассуждения.

5.4. Удовлетворение кредиторов. Рассмотрим следующую постановку задачи. Потребуем, чтобы все внутренние предприятия (для которых проводится взаимозачёт долгов), полностью оплатили претензии всех внутренних и внешних кредиторов. Для этого положим $p_n \equiv 1$ в (45), но q_n сохраним, поскольку на внешние предприятия наше влияние не распространяется.

Введём ту долю продажи основных фондов, которая в этих условиях превращает неравенства (45) в равенства:

$$\bar{\xi}_n = \frac{\bar{C}_n^t - b_n - \bar{D}_n^t - D_n^0}{B_n}. \quad (46)$$

Величина $\bar{\xi}_n$ даёт хорошее представление о состоянии предприятия и удобна для мониторинга. Если $\bar{\xi}_n < 0$, то данное предприятие может расплатиться без продажи основных фондов (разумеется при условии, что с ним расплатятся его дебиторы), т. е. его дела благополучны. Если же $\bar{\xi}_n > 0$, то для расплаты надо продавать основные фонды. Отсюда получаем долю, подлежащую продаже:

$$\xi_n = \max(0, \bar{\xi}_n). \quad (47)$$

Чем больше ξ_n , тем хуже дела предприятия. Если $\xi_n \ll 1$, то положение можно считать удовлетворительным. Но при отрицательном сальдо и малом B_n возможны случаи $\xi_n > 1$. Такая ситуация нередка для фирмпосредников с малыми основными фондами, а также для банков, ведущих рискованные операции. В этом случае даже полная продажа предприятия не покрывает его долгов.

На практике может потребоваться разновидность данной задачи, в которой не учитываются внешние задолженности, причём или все, или только дебиторские. Если не учитываются внешние дебиторские задолженности, то из формулы (46) надо исключить D_n^0 , а если и дебиторские, и кредиторские, то ещё дополнительно следует заменить $\bar{C}_n^t$ на $\bar{C}_n^i$.

Заметим, что такая постановка задачи несколько отличается от традиционной, ибо рассматривает не отдельное предприятие, а систему предприятий, и использует преимущества проводимого взаимозачёта. Поэтому значения ξ_n получаются обычно меньше, чем для отдельных предприятий, за счёт учёта их дебиторов.

5.5. Надёжности долгов. Вычислим $\bar{\xi}_n$ согласно (46), а затем произвольно выберем некоторые значения ξ_n так, чтобы выполнялось $0 \leq \xi_n \leq \max(0, \bar{\xi}_n)$. Эти значения должны выбирать эксперты, исходя из того, чтобы соответствующие продажи имущества были оперативными и не нарушали бы нормальной работы предприятий (автомобили дирекции, дома отдыха и т. п.). В частности, можно выбрать все $\xi_n = 0$, что означает отказ от продаж имущества.

Теперь уже невозможно полностью выплатить все долги, так что для части предприятий будет $p_n < 1$. Условия выплаты уценённых долгов (45) перепишем в виде

$$p_n \leq \frac{1}{C_n^t} \left(b_n + \xi_n B_n + D_n^0 + \sum_m p_m d_{nm}^i \right), \quad (48)$$

причём $0 \leq p_n \leq 1$.

Чтобы оплатить максимальную долю долгов, надо потребовать, чтобы p_n не только удовлетворяли (48), но и были бы как можно большими. Учтём, что для благополучных предприятий правые части неравенств (48) могут оказаться больше 1. Поэтому получим для всех n

$$p_n = \min(1, \bar{p}_n(\mathbf{p})), \quad (49a)$$

где

$$\bar{p}_n(\mathbf{p}) \leq \frac{1}{C_n^t} \left(b_n + \xi_n B_n + D_n^0 + \sum_m p_m d_{nm}^i \right). \quad (49б)$$

Это система уравнений для определения максимально возможных (при данных продажах имущества) долей погашения долгов p_n . Она напоминает уравнения задач линейного программирования. Исследуем её.

1. *Геометрическая интерпретация.* Рассмотрим многомерное пространство с координатами p_n , $n = 0, 1, 2, \dots$. Допустимые решения лежат в единичном кубе $0 \leq p_n \leq 1$ (см. рис. 6 п. 2.1 для трёхмерного случая) и дополнительно удовлетворяют неравенствам (48), а оптимальное решение определяется равенствами (49).

Если (48) превратить в равенство, то оно определяет гиперплоскость. Очевидно, сумма по m в (48) и (49б) не содержит слагаемого с $m = n$: предприятие само себе не должно. Коэффициенты d_{nm}^i неотрицательны. Это означает, что n -я гиперплоскость описывает монотонно возрастающую (точнее, неубывающую) зависимость координаты p_n от всех остальных координат. Эта гиперплоскость пересекает ось p_n в точке с координатой

$$\tilde{p}_n \leq \frac{b_n + \xi_n B_n + D_n^0}{C_n^t} \geq 0; \quad (50)$$

она получается из (49б) при $p_m = 0$. Поэтому все гиперплоскости выглядят так, как изображено на рис. 6 (они заштрихованы), с одной оговоркой: возможно $\tilde{p}_n > 1$, когда гиперплоскость целиком лежит вне единичного куба.

Каждое неравенство (48) описывает полупространство, лежащее по ту же сторону от соответствующей гиперплоскости, что и начало координат. Таким образом, множество допустимых решений есть выпуклый многогранник (назовём его G), ограниченный сторонами единичного куба и n описанными гиперплоскостями (чтобы этот многогранник лучше был виден на рис. 6, лежащие вне него куски гиперплоскостей

показаны более редкой штриховкой). Отсюда вытекают следующие утверждения.

2. *Существование решения.* Существует бесконечно много допустимых решений. В частности, каждая точка прямоугольного параллелепипеда $0 \leq p_n \leq \min(1, p_n)$ является допустимым решением, ибо этот параллелепипед лежит в единичном кубе и соответствующих полупространствах. Однако ясно, что эти решения неоптимальны, т.е. можно увеличить возврат долгов.

3. *Оптимальное решение.* Для его выделения из множества допустимых надо ввести определение оптимальности. Естественно такое определение:

Оптимально такое допустимое решение, где увеличение хотя бы одного p_n нарушает какие-то из неравенств (48).

Очевидно, оптимальное решение удовлетворяет системе (49). Решением этой системы является та из вершин многогранника G , которая наиболее близка к вершине $(1, 1, \dots)$ единичного куба (см. рис. 6). Таким образом, оптимальное решение существует и единственно.

5.6. Алгоритмы. 1. Задача продажи имущества для удовлетворения кредиторов решается по явным формулам (46), (47). Алгоритм очевиден, он экономичен и легко реализуется для любых масштабов задачи хоть совокупности всех предприятий страны.

2. Оптимальное решение задачи надёжности долгов (49) можно свести к задаче линейного программирования. Геометрическая интерпретация показывает, что оптимальная вершина многогранника G удовлетворяет условию

$$\sum_n \alpha_n p_n = \max, \quad \alpha_n > 0. \quad (51a)$$

Этот экстремум надо найти при ограничениях типа неравенств

$$p_n \leq 1, \quad p_n \leq \bar{p}_n(\mathbf{p}), \quad n = 1, 2, \dots; \quad (51б)$$

отсюда следует сделанное утверждение. Для решения задачи (51) пригодны любые алгоритмы линейного программирования. Однако все они сравнительно трудоёмки. Поэтому ими можно воспользоваться при региональном взаимозачёте долгов (тысячи предприятий), но не в масштабе страны (миллиона предприятий).

3. Из сказанного видно, что желательно разработать сверхбыстрый алгоритм решения задачи (49). Надежды здесь можно возлагать только на специфику данной задачи неотрицательность всех коэффициентов уравнений (49б) и связанные с этим наклоны гиперплоскостей.

Целесообразно опробовать следующий алгоритм. Выберем нулевое приближение $p_n^{(0)} \equiv 0$, а следующие приближения вычислим по фор-

мулам

$$p_n^{(q)} = \min(1, \bar{p}_n^{(q)}), \quad \bar{p}_n^{(q)} = \frac{1}{C_n^t} \left(b_n + \xi_n B_n + D_n^0 + \sum_m p_m^{(q-1)} d_{nm}^i \right). \quad (52)$$

В частности, $p_n^{(1)} = \min(1, p_n)$. Исследуем сходимость этого процесса. Она определяется производными

$$\frac{\partial p_n^{(q)}}{\partial p_m^{(q-1)}} = \begin{cases} \frac{\partial p_n^{(q)}}{\partial p_m^{(q-1)}} & \text{при } \bar{p}_n^{(q)} > 1, \\ 0 & \text{при } \bar{p}_n^{(q)} < 1. \end{cases} \quad (53)$$

Поскольку все производные неотрицательны, из (52)(53) вытекает следующее: если $p_n^{(q)} \geq p_n^{(q-1)}$ для всех n , то $p_n^{(q+1)} \geq p_n^{(q)}$. Но для первой итерации $p_n^{(1)} \geq p_n^{(0)}$, поскольку $p_n^{(1)} \geq 0$, а $p_n^{(0)} = 0$. Тем самым для всех итераций $p_n^{(q)} \geq p_n^{(q-1)}$. Значения $p_n^{(q)}$ ограничены сверху единицей. Следовательно, итерационный процесс (52) сходится к некоторому пределу p_n^* . Но если перейти к пределу в (52), то получим равенства (49). Следовательно, этот предел есть оптимальное решение.

Заметим также, что из (52) следует, что

$$p_n^{(q)} \leq \frac{1}{C_n^t} \left(b_n + \xi_n B_n + D_n^0 + \sum_m p_m^{(q-1)} d_{nm}^i \right).$$

Но выше установлено, что $p_m^{(q-1)} \leq p_m^{(q)}$. Тем самым,

$$p_n^{(q)} \leq \frac{1}{C_n^t} \left(b_n + \xi_n B_n + D_n^0 + \sum_m p_m^{(q)} d_{nm}^i \right),$$

то есть $p_n^{(q)}$ удовлетворяют неравенствам (48). Они лежат в единичном кубе. Следовательно, значения $p_n^{(q)}$ на всех итерациях принадлежат многограннику G , то есть являются допустимыми решениями.

Скорость сходимости алгоритма (52) трудно оценить. Далее была составлена программа и проверена на тестовых примерах (пока на методических). Для проверки на реальном примере нужны сведения об основных фондах и оборотных средствах предприятий. Таким реальным материалом мы пока не располагаем.

5.7. Программа. Для доработки алгоритма (52) и проведения методических расчётов была написана программа для персонального компьютера на языке FORTRAN. Исходная информация о состоянии системы предприятий выбиралась следующим образом:

1) Все продажи основных средств $\xi_n B_n$ и взыскания внешней дебиторской задолженности D_n^0 входят только в сумме с оборотными средствами b_n в формуле (52). Поэтому задавались только b_n , остальные члены опускались.

2) Предприятия распределялись по «экономической мощности» с помощью произвольно выбранной весовой функции $R(n)$, $1 \leq n \leq N$. Для определённости $R(n)$ бралась монотонно возрастающей, что не ограничивает общности. Конкретно в расчётах задавалась

$$R(n) = \sqrt{\frac{N}{N+1-n}}. \quad (54)$$

Это означало, что самое крупное предприятие в $\sqrt{N}$ раз мощнее самого мелкого, причём крупных предприятий немного, а малых очень много. Это достаточно правдоподобное распределение.

3) Матрица долгов (с учётом их знаков) кососимметрична. Достаточно сформировать её верхнюю половину $m > n$. Она формировалась построчно, начиная с верхней строки $n = 1$ (наименьшее предприятие). Рассчитывалась последовательность псевдослучайных чисел и брался её отрезок начиная с некоторого номера J . В очередной n -й строке бралась очередная клетка (n, m) с $m = n + 1, n + 2, \dots, N$ и пара очередных случайных чисел γ_j, γ_{j+1} . Долг и его знак определялись как

$$a_{nm} = \gamma_j R_n \cdot \text{sign}(\gamma_{j+1} - 0,5). \quad (55)$$

Если получился знак плюс, a_{nm} зачислялась в дебиторскую задолженность d_{nm} , если минус — в кредиторскую c_{nm} . Для следующей клетки бралась следующая пара случайных чисел и т. д. Таким образом, долги каждого предприятия получаются случайными, но их величины соответствуют мощности предприятия.

Заметим, что рассматривается замкнутая система. В ней все долги сбалансированы, так как все они внутренние:

$$\sum_{n,m} d_{nm} = \sum_{n,m} c_{nm}. \quad (56)$$

4) Оборотные средства предприятия считались неотрицательными. Их строили с помощью очередных случайных чисел также соответственно мощности предприятия:

$$b_n = b\sqrt{N} R_n \gamma_j, \quad b \geq 0. \quad (57)$$

Здесь b — свободный параметр, варьировавшийся в методических расчётах. Множитель $\sqrt{N}$ вводился потому, что каждое предприятие имеет $N - 1$ долг. Их знаки любые, так что они частично гасятся, и по статистике средняя величина сальдо должна быть $\sim N^{1/2}$ при больших N . Существенное отклонение от этого закона возможно лишь в случае монопольного завышения части цен.

Для удобства анализа результатов рассчитывались также следующие величины. Начальные сальдо каждого предприятия и всей систе-

мы (когда долги не уценены, т.е. $p_n = 1$):

$$s_n = b_n + \sum_m d_{nm} - \sum_m c_{nm}, \quad S = \sum_{n=1}^N s_n \equiv \sum_{n=1}^N b_n. \quad (58a)$$

Сумма исходных долгов в системе и коэффициент идеального взаимозачёта без уценки долгов:

$$C = \sum_{n,m} c_{nm}, \quad K = 1 - \frac{1}{C} \sum_{s_n < 0} |s_n|. \quad (58б)$$

Средняя вероятность возврата долгов в системе на каждой итерации (когда итерации сошлись, это окончательный результат):

$$p^{(q)} = \left[\frac{1}{N} \sum_{n=1}^N (p_n^{(q)})^2 \right]^{1/2}; \quad (58в)$$

её можно рассматривать как норму L_2 вектора $\mathbf{p}^{(q)}$. Сальдо каждого предприятия и всей системы на каждой итерации:

$$s_n^{(q)} = b_n + \sum_m p_m^{(q)} d_{nm} - p_n \sum_m c_{nm}, \quad S^{(q)} = \sum_{n=1}^N s_n^{(q)} = S. \quad (58г)$$

Если итерации сошлись, то все $s_n^{(q)}$ должны стать неотрицательными (сальдо всей системы выдаётся для контроля расчёта, и на всех итерациях должно равняться начальному).

5.8. Описание выдачи. Пример выдачи результатов расчёта приведён ниже (табл. 14). Он содержит следующие величины.

Строки начальных данных: N — число предприятий, b — множитель для оборотных средств, eps — точность сходимости итераций (всегда бралось $\epsilon = 10^{-5}$), Q — принудительное ограничение числа итераций на случай их плохой сходимости, J — начало отрезка последовательности случайных чисел.

Итоговая выдача с последней итерации содержит сначала следующую строку: K — коэффициент идеального взаимозачёта (58б), S — сальдо полной системы (58а), C — исходная сумма долгов системы (58б), q — число итераций до сходимости (если $q < Q$, то итерации сошлись), pq — окончательная средняя вероятность возврата долгов $p^{(q)}$ (58в).

Затем следует таблица итоговых величин для отдельных предприятий с номерами n : условная мощность предприятий Rn (54), их оборотные средства bn , их начальные сальдо sn , их новые сальдо после уценки долгов sqn (если среди них остались отрицательные, это указывает на незавершённость итераций), вероятности возврата долгов данными предприятиями pqn .

Далее следует матрица долгов Cnm ; в ней дебиторским задолженностям присвоен знак плюс, а кредиторским — минус. Если $N \leq 14$,

то матрица с таким числом значащих цифр удобно выдаётся на лист бумаги. При большем N её можно исключить из выдачи.

Далее следуют две таблицы, иллюстрирующие процесс сходимости итераций. В них по горизонтали отложен номер очередной итерации q , а по вертикали — номер предприятия n . В таблице «надёжность долгов» первая строка даёт знаменатель сходимости rq (59а), вторая — разности dq значений $p^{(q)}$ на соседних итерациях, третья — величину $p^{(q)}$. Далее следуют столбцы $p_n^{(q)}$. В столбцах таблицы «сальдо с уценкой долгов» приведены $S_n^{(q)}$ (58г), а отдельной строкой — $S^{(q)}$ (они оказываются равными S , как и должно быть). Эти две таблицы также можно не выдавать.

Т а б л и ц а 14. Пример выдачи.

$$N = 14, b = 0.2, eps = 10^{-5}, Q = 16, J = 100$$

Итоговая выдача с последней итерации

	K	S	C	q	pq
	0,7780	7,9792	54,4294	16	0,7002
n	Rn	bn	sn	sqn	pqn
1	1,0000	0,0971	-3,1357	0,0000	0,1707
2	1,0377	0,5421	-2,5144	0,0000	0,3615
3	1,0801	0,5380	0,7329	-0,0001	0,6931
4	1,1282	0,4657	0,5395	-0,0001	0,7414
5	1,1832	0,4487	-0,8648	0,0000	0,5367
6	1,2472	0,7489	5,9005	3,2874	1,0000
7	1,3229	0,1625	4,4622	1,7142	1,0000
8	1,4142	0,3912	0,8111	-0,0001	0,6218
9	1,5275	0,9277	2,3706	0,4560	1,0000
10	1,6733	1,2339	-4,4135	0,0000	0,2572
11	1,8708	0,9060	-0,7847	-0,0001	0,5386
12	2,1602	0,7181	-0,3725	-0,0001	0,5904
13	2,6458	0,6801	5,1406	2,5223	1,0000
14	3,7417	0,1191	0,1076	-0,0001	0,5424

5.9. Методические расчёты. Из самих формул (52) видно, что процесс построен по типу простых итераций, и должен иметь линейную сходимость. При этом погрешность убывает за каждую итерацию примерно в одно и то же число раз r . Поскольку выше доказана сходимость итераций, то $r > 1$. Вопрос только в величине r : если

Продолжение табл. 14

МАТРИЦА ДОЛГОВ C_{nm}

n/m	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0.0000	0.1101	-0.5003	-0.8772	-0.8753	-0.7482	0.2509	-0.3223	-0.4548	0.2655	-0.5508	0.1611	-0.0545	0.3630
2	-0.1101	0.0000	0.7327	0.8416	-0.2779	-0.2497	-0.8708	-0.9677	-0.1215	-0.2379	-0.1003	-0.0085	-0.7030	-0.9832
3	0.5003	-0.7327	0.0000	0.2407	0.6379	0.0286	-0.8238	0.3214	-0.4045	0.2996	-0.7409	0.6945	-0.5385	0.7123
4	0.8772	-0.8416	-0.2407	0.0000	0.4000	-0.6642	0.7794	0.8508	-1.1267	-0.2516	-0.5349	0.7107	0.2699	-0.1546
5	0.8753	0.2779	-0.6379	-0.4000	0.0000	0.4221	-0.8624	-1.1524	0.9163	-0.2028	-0.1013	-0.7666	-0.5752	0.8934
6	0.7482	0.2497	-0.0286	0.6642	-0.4221	0.0000	-0.3095	0.6037	-0.1417	0.5335	0.6425	0.6460	0.9260	1.0396
7	-0.2509	0.8708	0.8238	-0.7794	0.8624	0.3095	0.0000	1.2572	1.2395	1.2249	0.3343	-0.0633	-0.8014	-0.7277
8	0.3223	0.9677	-0.3214	-0.8508	1.1524	-0.6037	-1.2572	0.0000	0.0713	0.3900	0.2910	1.2130	0.0540	-1.0086
9	0.4548	0.1215	0.4045	1.1267	-0.9163	0.1417	-1.2395	-0.0713	0.0000	0.6903	1.0190	-0.8852	0.4622	0.1345
10	-0.2655	0.2379	-0.2996	0.2516	0.2028	-0.5335	-1.2249	-0.3900	-0.6903	0.0000	0.0544	-1.0509	-0.4450	-1.4944
11	0.5508	0.1003	0.7409	0.5349	0.1013	-0.6425	-0.3343	-0.2910	-1.0190	-0.0544	0.0000	-1.3431	-0.6194	0.5847
12	-0.1611	0.0085	-0.6945	-0.7107	0.7666	-0.6460	0.0633	-1.2130	0.8852	1.0509	1.3431	0.0000	-0.7981	-0.9849
13	0.0345	0.7030	0.5385	-0.2699	0.5752	-0.9260	0.8014	-0.0540	-0.4622	0.4450	0.6194	0.7981	0.0000	1.6376
14	-0.3630	0.9832	-0.7123	0.1546	-0.8934	-1.0396	0.7277	1.0086	-0.1345	1.4944	-0.5847	0.9849	-1.6376	0.0000

Продолжение табл. 14

НАДЕЖНОСТИ ДОЛГОВ $p_n(q)$

n/q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
rq	7.6575	1.8326	1.4538	2.0762	2.0327	1.8177	1.8680	1.8894	1.8513	1.8526	1.8621	1.8550	1.8529	1.8560	1.8552	1.8540
dq	0.1306	0.0713	0.0490	0.0236	0.0116	0.0064	0.0034	0.0018	0.0010	0.0005	0.0003	0.0002	0.0001	0.0000	0.0000	0.0000
pq	0.8694	0.7982	0.7491	0.7255	0.7139	0.7075	0.7041	0.7023	0.7013	0.7008	0.7005	0.7003	0.7003	0.7002	0.7002	0.7002
1	0.2846	0.2266	0.1934	0.1843	0.1784	0.1748	0.1728	0.1718	0.1713	0.1710	0.1708	0.1708	0.1707	0.1707	0.1707	0.1707
2	0.4570	0.4570	0.4469	0.4014	0.3810	0.3728	0.3676	0.3647	0.3632	0.3624	0.3620	0.3617	0.3616	0.3616	0.3615	0.3615
3	1.0000	1.0000	0.8297	0.7620	0.7326	0.7146	0.7044	0.6992	0.6964	0.6949	0.6941	0.6936	0.6934	0.6933	0.6932	0.6931
4	1.0000	0.9443	0.8424	0.7890	0.7696	0.7566	0.7494	0.7458	0.7438	0.7427	0.7421	0.7418	0.7416	0.7415	0.7415	0.7414
5	0.8159	0.6506	0.5856	0.5645	0.5523	0.5452	0.5412	0.5391	0.5380	0.5374	0.5370	0.5369	0.5368	0.5367	0.5367	0.5367
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	1.0000	0.8600	0.7334	0.6875	0.6573	0.6401	0.6318	0.6272	0.6247	0.6233	0.6226	0.6222	0.6220	0.6219	0.6218	0.6218
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0.3098	0.2822	0.2738	0.2669	0.2619	0.2598	0.2586	0.2580	0.2576	0.2574	0.2573	0.2573	0.2572	0.2572	0.2572	0.2572
11	0.8177	0.7091	0.6522	0.5940	0.5673	0.5549	0.5474	0.5432	0.5411	0.5400	0.5393	0.5390	0.5388	0.5387	0.5387	0.5386
12	0.9285	0.7142	0.6563	0.6303	0.6108	0.6011	0.5964	0.5936	0.5921	0.5913	0.5909	0.5906	0.5905	0.5904	0.5904	0.5904
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	1.0000	0.7151	0.6402	0.5986	0.5735	0.5585	0.5510	0.5471	0.5449	0.5437	0.5431	0.5428	0.5426	0.5425	0.5424	0.5424

САЛДО С УЩЕНКОЙ ДОЛГОВ $sn(q)$

N/q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
sq	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792	7,9792
1	-0,2546	-0,1452	-0,0399	-0,0261	-0,0159	-0,0085	-0,0043	-0,0024	-0,0013	-0,0007	-0,0004	-0,0002	-0,0001	-0,0001	0,0000	0,0000
2	0,0000	-0,0469	-0,2105	-0,0946	-0,0378	-0,0241	-0,0135	-0,0068	-0,0037	-0,0021	-0,0011	-0,0006	-0,0003	-0,0002	-0,0001	0,0000
3	0,0011	-0,5519	-0,2193	-0,0953	-0,0582	-0,0330	-0,0168	-0,0091	-0,0050	-0,0027	-0,0014	-0,0008	-0,0004	-0,0002	-0,0001	-0,0001
4	-0,2125	-0,3885	-0,2039	-0,0739	-0,0498	-0,0275	-0,0137	-0,0076	-0,0042	-0,0022	-0,0012	-0,0006	-0,0003	-0,0002	-0,0001	-0,0001
5	-0,7771	-0,3053	-0,0988	-0,0577	-0,0334	-0,0188	-0,0098	-0,0052	-0,0028	-0,0015	-0,0008	-0,0004	-0,0002	-0,0001	-0,0001	0,0000
6	4,6981	4,0140	3,6863	3,5039	3,4046	3,3499	3,3211	3,3056	3,2971	3,2926	3,2901	3,2888	3,2881	3,2877	3,2875	3,2874
7	2,9241	2,5354	2,1418	1,9427	1,8371	1,7807	1,7500	1,7334	1,7245	1,7197	1,7171	1,7157	1,7149	1,7145	1,7143	1,7142
8	-0,5660	-0,5116	-0,1854	-0,1223	-0,0693	-0,0335	-0,0186	-0,0104	-0,0055	-0,0029	-0,0016	-0,0009	-0,0005	-0,0002	-0,0001	-0,0001
9	1,3170	1,0599	0,7860	0,6191	0,5462	0,5055	0,4822	0,4701	0,4636	0,4600	0,4581	0,4571	0,4565	0,4562	0,4561	0,4560
10	-0,1764	-0,0535	-0,0443	-0,0317	-0,0137	-0,0073	-0,0043	-0,0022	-0,0012	-0,0007	-0,0004	-0,0002	-0,0001	-0,0001	0,0000	0,0000
11	-0,4671	-0,2451	-0,2504	-0,1147	-0,0534	-0,0326	-0,0178	-0,0091	-0,0050	-0,0027	-0,0015	-0,0008	-0,0004	-0,0002	-0,0001	-0,0001
12	-1,1160	-0,3015	-0,1352	-0,1019	-0,0506	-0,0244	-0,0145	-0,0079	-0,0041	-0,0022	-0,0012	-0,0006	-0,0003	-0,0002	-0,0001	-0,0001
13	4,1369	3,3216	2,9757	2,7668	2,6536	2,5930	2,5603	2,5427	2,5332	2,5281	2,5253	2,5239	2,5230	2,5226	2,5224	2,5223
14	-1,5283	-0,4021	-0,2229	-0,1352	-0,0803	-0,0401	-0,0209	-0,0118	-0,0064	-0,0034	-0,0018	-0,0010	-0,0005	-0,0003	-0,0002	-0,0001

$r \gg 1$, то сходимость быстрая; если $r \approx 1$, то потребуется очень много итераций.

Первые же расчёты показали, что важен выбор нулевого приближения. Нельзя брать $p_n^{(0)} = 0$: тогда при $b_n = 0$ на всех итерациях будет $p_n^{(q)} = 0$, то есть сходимости нет (хотя при $b_n > 0$ она не исключена). Поэтому было опробовано $p_n^{(0)} = 1$. При этом итерации всегда хорошо сходились.

Из неотрицательности производных $\partial p_n^{(q+1)} / \partial p_m^{(q)}$ (53) есть важное теоретическое следствие: если все $p_n^{(0)}$ больше окончательного решения, то итерации сходятся монотонно сверху. Все компоненты p от итерации к итерации не должны возрастать: $p_n^{(q+1)} \leq p_n^{(q)}$, $1 \leq n \leq N$. Численные расчёты полностью подтвердили это. Разумеется, средняя вероятность ведёт себя также: $p^{(q+1)} \leq p^{(q)}$.

Если сходимость линейная, то её знаменатель есть

$$r_q \approx \frac{p^{(q)} - p^{(q+1)}}{p^{(q-1)} - p^{(q)}}. \quad (59a)$$

Расчёт показал (см. выдачу), что отношение r_q уже после 3–4 итераций почти переставало изменяться. Это чётко подтверждает линейную сходимость, позволяет найти её знаменатель и дать надёжный критерий окончания итераций:

$$p^{(q)} - p^{(q+1)} < (r_q - 1)\varepsilon. \quad (59б)$$

Здесь ε — требуемая точность нахождения вероятности. Во всех расчётах принималось $\varepsilon = 10^{-5} = 0,001\%$, что заведомо превосходит требования практики.

Дополнительный контроль сходимости вёлся по значениям $s_n^{(q)}$. Поскольку $p_n^{(0)} = 1$ соответствуют неуплаченным долгам, то для части предприятий $s_n^{(0)} < 0$ при не слишком большом b . Все $s_n^{(q)}$ зависели от q монотонно, причём положительные $s_n^{(q)}$ стремились к неотрицательным пределам $s_n^{(\infty)}$, а отрицательные $s_n^{(q)}$ стремились к нулю. Когда выполнялся критерий сходимости (59), отрицательные $s_n^{(q)}$ не превосходили $\varepsilon s_n^{(0)}$, то есть были практически нулями (см. выдачу).

Помимо исследования чисто математических аспектов алгоритма, изучалось влияние N и b на результаты. Коэффициент оборотных средств варьировался в пределах $0 \leq b \leq 1,6$; число предприятий бралось от $N = 5$ до 640 (сколько позволяла память IBM PC-486, загруженного ещё и другими задачами). Время расчёта даже для $N = 640$ составляло ≈ 90 с. Причём, при небольшом усовершенствовании программы его можно было сократить вдвое.

Результаты расчёта зависели от тех случайных чисел, которые использовались в формировании исходных данных (55), (57). Поэтому при каждой паре значений N и b расчёты проводились с пятью разными значениями $J \approx N^2 \cdot (1, 2, 3, 4, 5)$. Тем самым, для каждого из этих рас-

чётов брались неперекрывающиеся отрезки последовательности случайных чисел. Затем для важнейших выходных данных проводилось усреднение по этим пяти расчётам и находились средние значения и среднеквадратичные отклонения.

5.10. Анализ результатов. Наиболее интересно поведение следующих величин: числа итераций q , необходимых для достижения точности $\varepsilon = 10^{-5}$; среднего процента платежей долгов p ; коэффициента идеального взаимозачёта K (последние две величины приводятся в процентах).

Зависимости этих величин от b (при $N = 14$) приведены на рис. 22–27 (эти и последующие рисунки похожи на рис. 9–18, но рассчитаны для других значений параметров).

Из рис. 22 видно, что число итераций q велико (25–55) при $b = 0$, и быстро уменьшается при увеличении b . Это понятно: когда нет оборотных средств, то конечное состояние далеко отстоит от нулевого приближения, и нужно много итераций. Уже при $b = 0,1$ достаточно $q \approx 10$, а при $b = 1,6$ всего $q \approx 2$, то есть сходимость становится быстрой. Зависимость во всём рассчитанном диапазоне хорошо аппроксимируется гиперболой

$$q(b) \approx \frac{1,72 + b}{0,043 + b}, \quad N = 14; \quad (60a)$$

при $b \rightarrow \infty$ она имеет правильную асимптотику $q \rightarrow 1$.

Из рис. 23 видно, что процент выплаты невелик ($p = 40$ –60 %) при $b = 0$, и стремится к 100 % при возрастании b . Расчёты также хорошо аппроксимируются гиперболической зависимостью с правильной асимптотикой

$$p(b) \approx \frac{0,055 + b}{0,110 + b} \cdot 100 \%, \quad N = 14. \quad (60б)$$

Из рис. 24 видно, что процент идеального взаимозачёта при $b = 0$ существенно выше, чем процент выплаты долгов. Это интересный результат, показывающий преимущество идеального взаимозачёта в условиях большой задолженности. При увеличении оборотных средств качественное поведение K аналогично p :

$$K(b) \approx \frac{0,255 + b}{0,351 + b} \cdot 100 \%, \quad N = 14, \quad (60в)$$

а количественное различие между ними уменьшается.

Зависимости исследуемых величин от N в диапазоне $5 \leq N \leq 640$ (при $b = 0,2$) приведены на рис. 25–27. Для удобства по оси N выбран логарифмический масштаб.

Исключительно интересной оказалась зависимость $q(N)$ на рис. 25. При $N \leq 20$ число итераций возрастало с увеличением N . Но при $N \geq 40$ рост прекратился, и среднее число итераций составило $q \approx 10$. Разброс этого числа при варьировании J уменьшается с ростом N и становится незначительным. Это очень важно для перспектив ме-

тогда: при такой малой трудоёмкости метод становится пригодным для федеральных банковских систем.

Зависимость $p(N)$ при $N \geq 100$ хорошо аппроксимируется формулой (рис. 26)

$$p(N) \approx \left(1 - \frac{1,27}{\sqrt{N}}\right) \cdot 100\%. \quad (61a)$$

Видно, что для больших систем со случайными долгами (без монопольных цен) при $N \rightarrow \infty$ средняя цена долга стремится к 100 %. Это значит, что большинство предприятий могут полностью расплатиться с долгами и лишь у немногих долги придётся уценивать. Стремление к этому пределу происходит по закону больших чисел $N^{-1,2}$. Однако для небольших систем $N \leq 50$ формула (61a) плохо применима. Самое удивительное, что среднее p больше, чем по этой формуле, а не меньше!

Зато зависимость $K(N)$ при всех N , от малых до больших, хорошо описывается формулой такого же вида (рис. 27)

$$K(N) \approx \left(1 - \frac{0,71}{\sqrt{N}}\right) \cdot 100\%. \quad (61б)$$

5.11. Детальная картина. Помимо общего состояния системы, интересен детальный анализ по предприятиям. Для него достаточно первой таблицы выдачи. В ней рассмотрена маленькая система $N = 14$ предприятий, не сильно различающихся по мощности (R_n от 1 до ≈ 4). Их оборотные средства малы: сумма всех оборотных средств $\sum b_n = S = 7,97$ примерно в 4 раза меньше чем $\sum |s_n| = 32,15$, а у некоторых предприятий отношение $|s_n|/b_n$ очень велико (при $n = 1$ оно $\approx 32!$).

Неудивительно, что среднее $p^{(q)}$ оказалось небольшим: 70 %. Из предприятий только у четырёх получилось $p_n^{(q)} = 100\%$, то есть их долги совершенно надёжны. У остальных эти величины заметно ниже, а у наиболее неблагополучного предприятия надёжность долга составила всего 17 %. Это предприятие $n = 1$ с большим отрицательным сальдо и малыми оборотными средствами.

Любопытно, что оценка состояния этого предприятия простейшими экономическими способами — лишь по его балансам — дала бы не 17 %, а гораздо худшее значение (возможно, даже $b_1/|s_1| \approx 3\%$). Видно, что учёт состояния всей системы предприятий позволяет улучшить оценки. Это важно, поскольку позволяет избежать продажи сравнительно «неплохих» долгов за бесценнок.

Для предприятий с положительным сальдо долги также не стопроцентно надёжны; так, у последнего предприятия это всего $p_{20} = 54\%$ (оно выдало много кредитов, а его оборотные средства малы).

Разумеется цифры в этих расчётах учитывают только оборотные средства, то есть ориентированы на краткосрочные финансовые расчёты. Такое важно при срочных платежах и особенно при биржевых паниках. Предлагаемый метод особенно полезен в таких ситуациях.

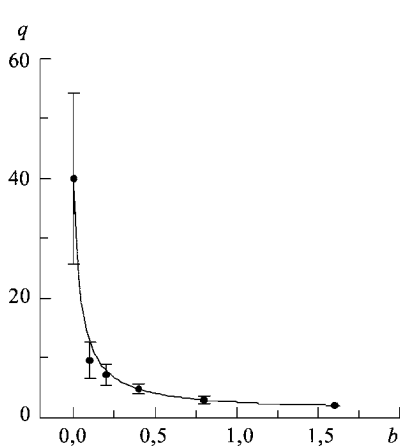


Рис. 22. Точки — численные расчёты при $N = 14$ с указанием разброса; линия — формула (60а)

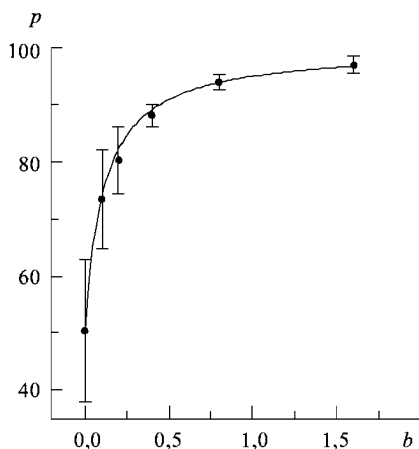


Рис. 23. Точки — численные расчёты при $N = 14$ с указанием разброса; линия — формула (60б)

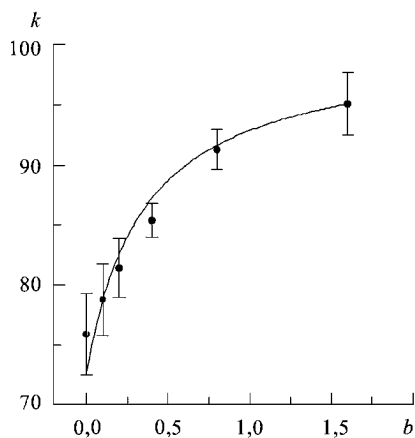


Рис. 24. Точки — численные расчёты при $N = 14$ с указанием разброса; линия — формула (60в)

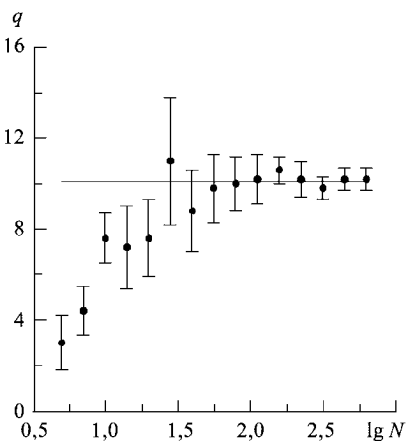


Рис. 25. Точки — численные расчёты при $b = 0,2$ с указанием разброса; линия — $q = 10,1$

Если рассматривать долгосрочные расчёты, то надо учесть возможность продажи части имущества должников (в первую очередь, непроизводственных фондов). Это моделируется увеличением коэффициента b в расчётах и приводит к возрастанию надёжности всех долгов $p_n^{(q)}$.

Выводы. Предложенный алгоритм оказался очень надёжным и быстрым. Итерации всегда сходились, причём даже при больших

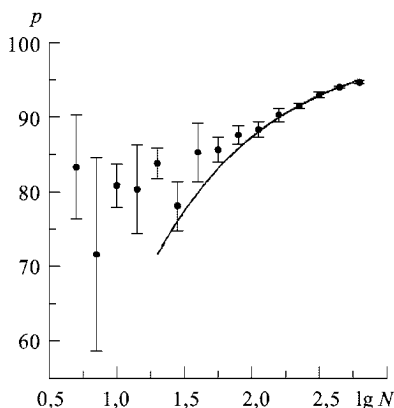


Рис. 26. Точки — численные расчёты при $b = 0,214$ с указанием разброса; линия — формула (61a)

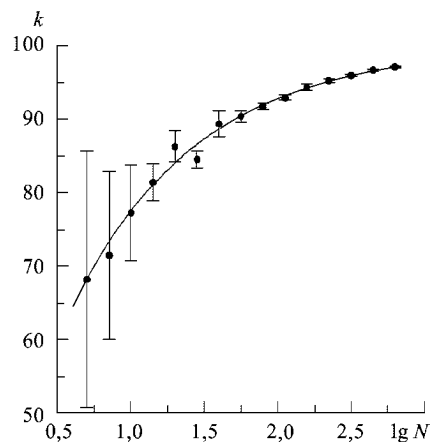


Рис. 27. Точки — численные расчёты при $b = 0.2$ с указанием разброса; линия — формула (61b)

$N = 640$ хватало $q = 10$ – 11 итераций, а при малых N обычно требовалось ещё меньше.

Оценим трудоёмкость алгоритма для больших систем $N = 10^3$ – 10^6 предприятий, считая, что матрица долгов уже задана и введена в компьютер. Предприятие никогда не связано со всеми предприятиями системы, так что матрица долгов слабо заполнена. Обычно малые предприятия имеют 10 – 100 связей, наиболее крупные до 10^3 – 10^4 связей. Разумно предположить, что среднее число связей-долгов одного предприятия не более $\sim N^{1/2}$ (скорее всего, среднее будет заметно меньше). В расчёте по формулам (52) будем суммировать только фактически имеющиеся долги, пропуская нулевые. Тогда одна итерация потребует $\leq 2N^{3/2}$ операций (умножений и сложений), и запоминания $\approx N^{3/2}$ чисел.

Поскольку число итераций оказалось $q \approx 10$ не зависящим от N , то общий объём расчётов составит $\approx 20N^{3/2}$ операций (дополнительная память не нужна). При расчёте в масштабе всей страны $N \approx 10^6$, так что потребуются $\approx 2 \cdot 10^{10}$ операций и $\approx 10^9$ чисел в памяти. Это реально на крупном компьютере.

Строго говоря, постоянное число итераций $q_0 = 10$ было получено для модельного примера с чисто случайной матрицей долгов. В этом примере нулевым приближением было $p^{(0)} = 1$, а окончательным результатом $p^{(q)} \approx 1 - a/\sqrt{N}$ согласно (61a). Разница между ними невелика и уменьшается при увеличении N :

$$\Delta_0 \equiv 1 - p^{(q)} \approx \frac{a}{\sqrt{N}} \rightarrow 0.$$

Можно получить величину знаменателя сходимости из условия $r^{q_0} = \Delta_0/\varepsilon$. Но для неслучайной матрицы долгов при наличии монопольных цен средняя вероятность может существенно отличаться от 1 даже при большом N : $\Delta = 1 - p^{(q)} \approx 1$. Число итераций q тогда определяется по условию $r^q = \Delta/\varepsilon$. Отсюда следуют формулы

$$r \approx \left(\frac{a}{\varepsilon \sqrt{N}} \right)^{1/q_0}, \quad q \approx q_0 \frac{\lg \left(\frac{\Delta}{\varepsilon} \right)}{\lg \left(\frac{a}{\varepsilon \sqrt{N}} \right)}. \quad (62)$$

Принимая для тестового примера $\varepsilon = 10^{-5}$, $a \approx 1,2$, $q_0 \approx 10$ и $\Delta \approx 1$, получим следующую табл. 15 зависимости r и q от N для наихудшей комбинации долгов:

Таблица 15

N	10^1	10^2	10^3	10^4	10^5	10^6
q	11	12	14	16	19	24
r	2,87	2,56	2,28	2,03	1,81	1,61

Видно, что возможный рост $q(N)$ крайне медленный, так что им можно пренебречь.

Таким образом, разработанный здесь алгоритм расчёта надёжности долгов пригоден для использования в федеральной системе банковских расчётов.

§ 6. Заключение

6.1. Фракционирование. Для системы предприятий с большим числом взаимных связей метод фракционного взаимозачёта долгов по эффективности мало уступает даже методу переадресации долгов, дающему теоретический предел для коэффициента погашения долгов. Он соответствует существующему законодательству и не требует привлечения кредитов. Его трудоёмкость настолько невелика, что расчёт для масштабов страны ($N \approx 10^6$) легко выполнить на современных компьютерах. Для масштабов района ($N \approx 10^3$) достаточно всего 1–10 мин. работы персонального компьютера.

Отметим ещё одну рекордную характеристику этого метода: ранее никогда не решались задачи линейного программирования с таким числом неизвестных.

Поэтому из всех существующих сейчас методов взаимозачёта долгов только метод фракционирования пригоден для создания государственной и региональных сетей для финансовых расчётов.

Если сеть взаимных связей предприятий редкая, то коэффициент взаимозачёта может стать небольшим. Поэтому региональный взаимозачёт целесообразен, во-первых, лишь для небольших местных

предприятий, основная часть связей которых лежит внутри региона. Во-вторых, он должен охватывать возможно большую долю таких предприятий, лучше всего — все. Наиболее же эффективным будет федеральный центр, охватывающий все предприятия страны.

6.2. Кредиты. Список Башкирии не содержал сведений об оборотных средствах предприятий, так что в расчётах принималось $b_n = 0$. В итоге около 20 % долгов не гасилось фракционированием, а 12 % даже переадресацией. В модельных примерах было видно, что увеличение b_n быстро приближает коэффициент взаимозачёта к 100 %.

Не представляет труда провести моделирование оптимальной выдачи кредитов. Для этого мысленно выдадим кредит b_n какому-то предприятию с отрицательным сальдо и рассчитаем, насколько увеличится погашение суммарного долга системы по сравнению с $b_n = 0$. Такую виртуальную выдачу кредита проведём для каждого предприятия-должника. Сравнивая результаты, выделим предприятия, где каждый рубль кредита обеспечивает наибольшее погашение долгов; им и выдаём кредиты.

Эта процедура целесообразна, если коэффициент «усиления» много больше 1. Такое возможно при густой сети долговых связей. Если же коэффициент усиления невелик, то централизованные кредиты невыгодны (именно такова ситуация со списком Башкирии).

Описанная процедура требует многократного расчёта фракционного взаимозачёта. Но быстрое действие предложенного метода таково, что это не вызывает затруднений. Для списка Башкирии однократный расчёт фракционирования на PC Pentium IV составлял менее 1 с.

6.3. Мониторинг. Собранная информация позволяет найти немало «болевых точек» в экономике региона или страны. Рассмотрим некоторые ситуации.

а) Одно из предприятий имеет большое положительное сальдо. Это похоже на монополиста, необоснованно завышающего цены. Если добавок ему должно много различных предприятий, это усиливает подозрения. Остаётся выяснить, что это за предприятие.

б) У предприятия большое отрицательное сальдо. Это может указывать на плохое управление предприятием и опасность банкротства.

в) Большие прямые встречные долги пар предприятий (как в списке Башкирии). Возможны такие причины. Одна — неаккуратность в работе бухгалтерий («забыли»). Вторая — недобросовестность одного из партнёров: он хочет получить с контрагента, но не платить ему. Третья — взаимный сговор: выдача взаимных кредитов, чтобы не показывать наличие оборотных средств.

Эта ситуация известна во многих регионах. На Дальнем Востоке энергетики не платят железной дороге долг за перевозку угля, а та не платит им за электроэнергию. Между РАО ЕЭС и Москвой также есть встречные неплатежи.

6.4. Переадресация. Этот метод [1–4] может оказаться полезным для частичного погашения тех долгов, которые остаются после применения метода фракционирования. Против него выдвигают возражение, что он не укладывается в действующее законодательство. Однако законодательство допускает обращение векселей, а метод переадресации почти эквивалентен переводному векселю.

Само вексельное обращение, хотя и было у широкой публики синонимом мошенничества, может хорошо работать при жёстком контроле. Например, в Англии в 1880-е годы оно было широко распространено, и попадались вексели со 100–120 передаточными надписями! Очевидно, такое возможно только при крошечном дисконте в доли процента, а это свидетельствует о доверии к векселю.

При переадресации, как и при вексельном обращении, ваше долговое обязательство может предъявить не кредитор, а третье лицо. Это не противоречит законодательству. Разница лишь в одном: третье лицо может отказаться от приобретения данного векселя или согласиться при значительном дисконте; а при переадресации ему фактически навязывают должника. Парные переговоры при этом нереальны, ибо компьютерная система выдаёт согласованное решение сразу для группы предприятий. Отказ одного из них аннулирует весь расчёт.

Здесь требуется разработать юридическую процедуру уценки долгов и согласования. Её можно сделать приемлемой для предприятий, включив льготы согласным и налоги на непогашенные долги как на имущество для несогласных.

6.5. Уценка долгов. Продажа долгов (разумеется, со скидкой) по частным соглашениям между юридическими лицами происходит не первый год. Величина уценки до сих пор определялась на основании экспертных оценок. В § 5 показано, как можно объективно рассчитать эту величину.

Для этого проводится расчёт погашения долгов методом фракционирования, но не осуществляется само погашение. Полученные вероятности являются хорошими оценками вероятности *срочного* возврата долгов.

Если же фракционный взаимозачёт осуществлён, то у предприятия может не остаться оборотных средств или дебиторов, так что оставшиеся долги ему нечем гасить. В этом случае надо экспертно оценить, каково непрофильное имущество данного предприятия (персональные автомобили, дачи для руководства и т. п.), и его стоимость задать в качестве оборотных средств b_n . С этими величинами b_n для всех предприятий-должников надо провести новый расчёт фракционирования. Он даст новые вероятности p_n , соответствующие более медленному возврату долгов (после продажи непрофильного имущества).

Аналогично можно действовать при процедуре банкротства. Там в качестве b_n надо брать полную стоимость основных фондов предприятия, которую можно выручить при их продаже.

6.6. Банковские проводки. Если долговое обязательство между парой предприятий состоит из суммы многих отдельных договоров, то есть полезное видоизменение метода фракционирования [6]. При нём доля суммарного долга p_n , рассчитанная к выдаче, набирается из отдельных договоров. Эти договоры оплачиваются целиком, а остальные целиком не оплачиваются. Такой алгоритм особенно важен для покупок с предоплатой, выполняемых с переводом денег из банка покупателя в банк продавца. Он позволяет выполнять проводки, даже когда банковские остатки на счетах межбанковского расчётного центра малы или отсутствуют.

В [6] для расчёта величин p_n использовались классические подходы, основанные на симплекс-методе. Они требуют $\sim N^3 - N^4$ операций, что очень трудоёмко для больших систем. Однако в этой задаче также можно использовать предложенный здесь итерационный алгоритм (9), что делает весь метод в целом сверхбыстрым.

6.7. Итоги. Описанные здесь методы позволяют построить федеральную и региональную системы взаимозачёта долгов на базе недорогой современной вычислительной техники. Во время отладки и опытной эксплуатации таких систем взаимозачёт будет проводиться время от времени. В штатном режиме возможен взаимозачёт раз в сутки. При этом информация, требующая уточнения, выпадает из данного зачёта, но после её исправления лежит менее суток.

Такая система может принести большую пользу России.

Список литературы

1. *Калиткин Н.Н.* Задача зачёта взаимных долгов предприятий // Доклады Академии Наук. 1995. Т. 343, № 1. С. 12–14.
2. *Калиткин Н.Н.* Оптимальный взаимозачёт долгов предприятий // Математическое моделирование. 1995. Т. 7, № 1. С. 11–21.
3. *Калиткин Н.Н., Кузьмина Л.В.* О зачёте взаимных долгов предприятий // Математическое моделирование. 1995. Т. 7, № 4. С. 64–72.
4. *Калиткин Н.Н., Михайлов А.П.* Идеальное решение задачи зачёта взаимных долгов // Математическое моделирование. 1995. Т. 7, № 6. С. 111–117.
5. *Калиткин Н.Н., Кузьмина Л.В., Черненко М.В.* Взаимозачёт долгов методом фракционирования // Доклады Академии Наук. 2001. Т. 378, № 1. С. 29–32.
6. *Голосной И.О., Гужев Д.С., Калиткин Н.Н.* Алгоритмы расчёта предприятий через банки // Доклады Академии Наук. 1996. Т. 349, № 3. С. 299–301.
7. *Калиткин Н.Н., Кузьмина Л.В., Черненко М.В.* Метод фракционирования в клиринговых расчётах // Математическое моделирование. 2003. Т. 15, № 2. С. 102–112.

Глава IV

СТАБИЛЬНАЯ ПЕНСИОННАЯ СИСТЕМА

Сформулированы важнейшие требования к пенсионной системе: 1) она должна охватывать всех инвалидов и пожилых граждан; 2) размер пенсии должен зависеть от трудового вклада за всю жизнь; 3) система должна быть долгоживущей, не требующей законодательных изменений при сильной инфляции, обвалах экономики и банкротствах всевозможных компаний или фондов.

Показано, что такая система должна быть 1) государственной, а не коммерческой; 2) распределительного типа, а не накопительного; 3) рассчитанной на граждан со средними заработками, а не сверхвысокими. Рассматриваемые сейчас Правительством проекты во многом не удовлетворяют этим требованиям, и в них заложены недостатки, которые через несколько лет станут источниками социальной напряжённости.

Разработаны основы пенсионной системы требуемого типа. Даны простые формулы расчёта различных составляющих частей пенсии (социальной, трудовой, специальных надбавок) для разных категорий пенсионеров. Формулы содержат небольшое число коэффициентов, зависящих от демографической структуры населения и состояния экономики страны, и определяемых на базе данных Пенсионного фонда, Госкомстата и налоговой службы.

По статистическим данным найдена функция распределения населения России по доходам. Для основной массы населения она оказалась бистепенной, а учёт сверхвысоких доходов лишь незначительно изменил её. Это распределение исследовано, найден его децильный коэффициент. Расчёт баланса Пенсионного фонда с этим распределением оказался очень близким к расчёту по средним величинам. Это упрощает управление Пенсионным фондом.

Собраны и проанализированы статистические данные, потребовавшиеся для данных расчётов. Даны рекомендации по распространению данных подходов на медицинское страхование населения.

Введение

Хорошо продуманная пенсионная система является важным элементом законодательства страны, необходимым для поддержания социальной устойчивости общества. Она должна обеспечивать достойный уровень жизни пенсионеров. Перечислим наиболее веские причины этого.

Во-первых, примерно треть избирателей нашей страны — пенсионеры, а у большинства остальных избирателей есть близкие родственники пенсионного возраста; ни одно правительство не может пренебрегать их мнением. Во-вторых, пенсионеры составляют значительную часть покупателей внутреннего рынка страны. Их нищета является серьезным препятствием для развития экономики. И наконец, порядочное правительство не может игнорировать свои моральные обязательства перед народом.

Какие же требования предъявляет общественное мнение к хорошей пенсионной системе? Она должна

- обеспечивать пенсионерам достойный уровень жизни (насколько это позволяет экономика страны);

- быть социально справедливой (никто не останется без пенсии, но чем больше гражданин работал, тем больше должна быть его пенсия);

- быть простой и понятной среднему образованному человеку (чтобы каждый мог сам уверенно рассчитать размер своей пенсии);

- быть стабильной (то есть применяться без поправок десятки лет, невзирая на инфляцию или биржевые спекуляции с валютой и акциями).

В нашей стране примерно 40–35 лет действовал пенсионный закон 1957 г. [1]. Он не учитывал медленную инфляцию, и поэтому устарел уже к середине 1980-х годов, а развал экономики из-за реформ 1992 г. доканал его. Потребовалась регулярная корректировка его цифр. Это сделал закон 1992 г. [2], предусматривавший регулярную индексацию пенсий по мере нарастания инфляции. Новый закон был принят в 1997 г. [3], но он также оказался недолговечным и фактически не пережил кризиса 1998 г. Последний пенсионный закон вступил в действие с 2002 г. [4], но оказался, во-первых, настолько сложным, что его с трудом понимали даже работники социального обеспечения. Во-вторых, он начал быстро устаревать и обраться поправками (имеются сведения о подготовке новых проектов).

В данной главе изложена пенсионная система, разрабатывавшаяся авторами с начала 2001 г. Она удовлетворяет всем указанным выше требованиям, а также другим важным, но не осознаваемым обществом условиям:

- позволяет отменить все ведомственные пенсионные системы, а также персональные пенсии, назначаемые сейчас отдельными Указами Президента РФ;

- хорошо согласована с другими законами (например, Семейным кодексом РФ);

- настолько проста, что прозрачна для общественного контроля;

- аналогично ей можно строить систему медицинского страхования.

Принятие данной системы могло бы решить многие проблемы, волнующие Россию.

§ 1. Принципы пенсионной системы

1.1. История. Ещё у небольших первобытных племён, живших компактной группой, существовало обеспечение нетрудоспособных — стариков, больных и, разумеется, детей. Это хорошо изучено этнографами в конце XIX — начале XX веков, когда таких племён жило ещё много. Общая добыча племени делилась на всех, так что это была система распределительного типа. Доли были неравными: они учитывали потребности человека и его нужность для племени (при сильном голоде стариков могли оставить без еды), но не его прошлый трудовой вклад.

Позднее, по мере развития общества и повышения благосостояния, социальное обеспечение в основном переместилось в семью, где стариков кормили выросшие дети («дети — капитал на старость»). Оптимальная численность семьи зависела от трудности условий жизни; на русском Севере она доходила до 30–40 человек, живших в одной избе (женатые сыновья не отделялись). Лишь при гибели основных кормильцев община брала на себя обеспечение нетрудоспособных. Система также являлась распределительной и не учитывала прошлый трудовой вклад.

Индивидуальное накопление на старость стало возможным с появлением денег. Его очевидный недостаток — человек может оценить свои потребности, но не знает времени смерти. Поэтому он не может рассчитать нужную ему сумму. Для преодоления этого недостатка была изобретена покупка пожизненной ренты, очень распространённая в Европе (особенно во Франции) с начала XIX века. В обоих случаях размер обеспечения определялся самим получателем по накопленной им на старость сумме, то есть прошлому трудовому вкладу.

Однако по психологическим причинам далеко не каждый человек в состоянии много лет копить деньги на старость и не истратить их. Поэтому к концу XIX века в Европе получили распространение так называемые эмеритальные кассы: человек ежемесячно вносил в них часть своего заработка, а после достижения пенсионного возраста получал пожизненную пенсию. Её размер зависел от суммарного вноса. В России также стали быстро возникать такие кассы. Именно на этой системе основаны современные накопительные страховые фонды на Западе.

Параллельно возникла система государственных пенсий. В России она была учреждена в первой четверти XIX века по инициативе вдовы Павла I Марии Фёдоровны (отсюда название «Ведомство императрицы Марии»). Первоначально предполагалось обеспечить пенсиями отставных неимущих офицеров, но затем это было распространено на других государственных служащих: чиновников, преподавателей высших учебных заведений, казённых гимназий и училищ, врачей государственных больниц и т. п. Все они после 25 лет службы имели право на полную пенсию, а после 20 лет — на половинную. Для финансирования

этой системы использовались государственные средства: поступления от налогов и доходы от определённых государственных предприятий (тогда в России было немало казённых заводов, приносивших хорошие доходы). Свободная часть этих средств вкладывалась в ценные бумаги под проценты (частично — в государственные же бумаги). Следовательно, формально это была смешанная распределительно-накопительная система, но фактически она была почти чисто распределительной. Размер пенсии зависел от заработка, причём по последней должности.

Некоторые предприниматели также назначали пенсии рабочим и служащим своих предприятий, но далеко не всем, а лишь наиболее заслуженным. Кроме того, они создавали бесплатные больницы и школы, обслуживавшие гораздо большее число работников. Тем не менее, в начале XX века во всём мире различные пенсионные системы охватывали лишь меньшую часть населения, причём средне- и высокооплачиваемую.

Для рабочих и крестьян Европы, где семейные связи в это время уже ослабели, вопрос обеспечения в старости был болезненным. Во Франции возникла семейная рента — пожилой фермер передавал ферму сыну, обговорив в нотариальном договоре размер пожизненного содержания. В Англии, где фермы обычно не принадлежали фермеру, а брались в аренду, возникли муниципальные работные дома. В них пожилой или просто безработный могли получить ночлег и еду, но должны были посильно (а то и через силу) трудиться на общественных работах. Условия там были таковы, что работного дома боялись, как огня. Возникла ещё Армия спасения с её бесплатными супами и ночлежками для безработных и бездомных (финансировалось это на общественные пожертвования). Всё это было нищенской помощью.

Всеобъемлющая пенсионная система была создана впервые в СССР; она была подобием старой системы государственных пенсий, но распространялась сначала на рабочих и служащих, а затем и на колхозников, то есть охватила всё население. Старые же эмеритальные кассы России не пережили национализации банков и разрухи хозяйства страны за время гражданской войны.

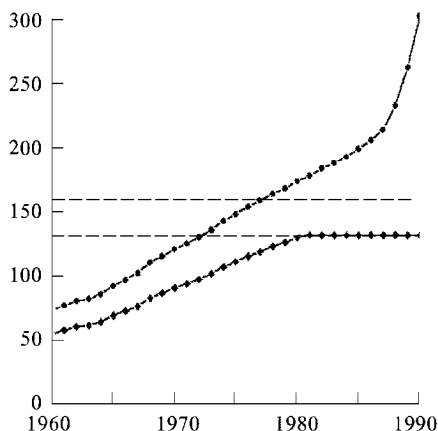


Рис. 1. Средние зарплаты в России в рублях/месяц за указанные годы точки; ромбики — пенсия при средней зарплате; нижняя горизонтальная линия — «потолок» пенсии при предельном рабочем стаже для нелюбимых профессий, верхняя — для льготных

Таблица 1. Средние зарплаты z и пенсии p при максимальном рабочем стаже по пенсионной системе СССР 1957 г.; предельный размер пенсий для льготных профессий 132 р/м, для льготных 160 р/м

Год	z руб/мес.	p руб/мес.
1960	73,1	54,82
1961	77,1	57,82
1962	80,9	60,68
1963	82,4	61,80
1964	86,0	64,50
1965	92,5	69,38
1966	97,2	72,90
1967	102,5	76,88
1968	110,9	83,18
1969	115,6	86,70
1970	121,2	90,90
1971	125,6	94,20
1972	130,4	97,80
1973	136,2	102,15
1974	143,2	107,40
1975	148,7	111,52
1976	154,2	115,65
1977	159,1	119,33
1978	164,4	123,30
1979	168,4	126,30
1980	174,0	130,50
1981	178,3	132,00
1982	184,0	132,00
1983	188,3	132,00
1984	193,2	132,00
1985	199,2	132,00
1986	206,1	132,00
1987	214,4	132,00
1988	233,2	132,00
1989	263,0	132,00
1990	303,0	132,00

Пенсионная система СССР, введённая в конце 50-х годов, была тогда в целом хорошей: минимальный размер пенсии при полном трудовом стаже обеспечивал в те годы возможность удовлетворительного проживания, а максимальный — хорошего (более чем вдвое превышал среднюю зарплату в стране!). Не было «уравниловки». Имелись недостатки, но существенными являлись лишь два: 1) крошечная пенсия,

если рабочий стаж хоть на месяц не дотягивал до полного; 2) все цифры в рублях были абсолютными, так что к концу 80-х годов инфляция сделала пенсии фактически небольшими, а вновь назначаемые пенсии стали почти уравнительными (рис. 1 и табл. 1.)

Пенсионная система СССР, несмотря на её недостатки, создала атмосферу социальной защищённости основной мало- и среднеоплачиваемой части населения. Это произвело настолько сильное психологическое воздействие на весь мир, что ведущие западные страны были вынуждены под давлением своих профсоюзов ввести у себя систему государственных социальных пенсий. Их система сейчас носит смешанный характер, включая государственную социальную пенсию распределительного типа и страховые фонды накопительного типа (государство не управляет ими, но контролирует).

1.2. Зарубежные пенсии. В развитых государствах пенсиями охвачено всё нетрудоспособное население страны, включая иностранцев недавно приехавших в страну и получивших гражданство (а в некоторых странах — даже не получивших). Приведём сведения по некоторым государствам, взятые из обычной периодической печати (не было возможности пользоваться официальными статистическими данными).

Германия. Пенсионный возраст для мужчин и женщин одинаков 65 лет: при определённых условиях можно выйти на пенсию с 60 лет, но тогда размер пенсии будет заметно ниже.

Величина отчислений в пенсионный фонд $\approx 15\%$ зарплаты (половину платит работник, половину — предприятие). Преуспевающие предприятия нередко привлекают работников тем, что платят дополнительно в пенсионный фонд сверх установленного законом процента.

Размер пенсии зависит от рабочего стажа и заработка. Социальный минимум пенсии составляет ≈ 500 евро в месяц (и многие «восточные» немцы получают именно такие пенсии). При полном 45-летнем стаже (одинаковом у мужчин и женщин!) она в 2000 г. составляла 70 % от среднего заработка за последние 3 года. Средний размер пенсий по стране составлял тогда 2500 DM/мес. для мужчин и 2000 DM/мес. для женщин (их зарплаты в среднем меньше). У чиновников и журналистов пенсии составляли 5000 DM/мес. и более, у руководителей доходили до 10–15 тыс. DM/мес., а у канцлера — более 30 тыс. DM/мес. Это в 13 раз больше средней пенсии.

Эти средние пенсии лишь по российским меркам кажутся большими. В Германии жильё и социально-медицинский уход дороги. А взрослые дети обычно живут отдельно от родителей и далеко (там, где сумели найти хорошую работу), так что не ухаживают за стариками и не помогают им деньгами. Поэтому заметная часть пенсионеров вынуждена покупать себе на пенсию место в доме престарелых, отказываясь от собственного жилья. Правда, эти дома престарелых существенно лучше российских, и есть разные категории, зависящие от финансовых возможностей пенсионера.

Кроме того, оборотная сторона хорошей медицины — население страны стареет. В 2000 г. отношение чисел работающих и пенсионеров было 4:1. По демографическим прогнозам за 10 лет оно может упасть до 2:1. Достоверность этих прогнозов неясна, но настораживает и значительная безработица — 9%. Поэтому уже в 2000 г. в Германии обсуждался болезненный вопрос: сократить размер пенсий с 70% до 67% зарплаты.

США. Пенсионный возраст для мужчин — 65 лет. Обязательное по закону пенсионное страхование в последние годы 12,4–12,7% (и медицинское страхование 2,9%); из них половину платит работник, половину — работодатель. Наряду с обязательным распространено добровольное страхование; платит только работник, и размеры отчислений устанавливает сам. Обычно в сумме набирается от 60% до 90% зарплаты.

Государство жёстко контролирует деятельность пенсионных фондов. Им разрешается вкладывать аккумулированные средства только в высоко надёжные ценные бумаги. Но такие бумаги приносят скромный доход, обычно не превышающий 4–5% (а при застое экономики он может быть ещё ниже). Высокодоходные, но рискованные вложения категорически запрещаются. Крах любого пенсионного фонда вызовет взрыв социального возмущения.

По официальным данным, инфляция в США составляет 1,5–2% даже в благополучные годы. По общему мнению экономистов, фактическая инфляция — оценка по стоимости товаров с одинаковыми потребительскими свойствами — даёт скорее 3–4% в среднем по последним 30–40 годам. Таким образом, фактическая доходность этих вложений в сопоставимых ценах еле покрывает расходы на содержание аппарата фонда. Реально это означает нулевую доходность, то есть фонд лишь предохраняет пенсионные деньги от вялой инфляции, но не приумножает их. Накопительность оказывается иллюзией.

Очевидно, от обвальной инфляции такие условия не защищают.

Англия. Пенсионный возраст 65 лет. В связи с увеличением продолжительности жизни (старением населения) обсуждается вопрос об увеличении пенсионного возраста до 67 лет.

Франция. Пенсионный возраст примерно 60 лет, а полный рабочий стаж — 40 лет (37 лет для государственных чиновников, но к ним относится достаточно узкий круг лиц). Размер пенсии — 50% среднего заработка по последним 11 годам.

Швеция. Около 2002 г. на пенсию выходил президент крупного концерна, аккуратно плативший пенсионные отчисления с огромного оклада. Его пенсия, пропорциональная отчислениям, оказалась ≈ 90 млн крон в год! Это вызвало оживлённую дискуссию в печати и было неблагоприятно воспринято общественным мнением.

Япония. Неизвестны сведения о пенсионной системе, но известна средняя продолжительность жизни — 77 лет у мужчин и 83 года у

женщин. Такова же продолжительность жизни коренного населения на Гавайях, а во всех других странах она меньше.

Итоги. Сводка количественных данных по пенсиям в некоторых странах приведена в табл. 2 (они не вполне точно совпадают с приведёнными в тексте цифрами, поскольку использовались данные разных лет). Во всех развитых странах можно выделить общие черты пенсионных систем:

— Государство законодательно устанавливает обязательные пенсионные отчисления, то есть им придан статус налога, только поступающего не в государственную казну, а на счета пенсионных фондов.

— Пенсионные фонды фактически не являются накопительными.

— Пенсиями обеспечены все нетрудоспособные, в том числе и никогда не работавшие.

— Пенсионный возраст для мужчин и женщин обычно одинаков. Это понятно. Средняя продолжительность жизни женщин на 5–10 лет выше, чем у мужчин, так что их здоровье при одинаковом возрасте в среднем лучше.

— Чем больше средняя продолжительность жизни в стране, тем больше должен быть пенсионный возраст.

— Прямая пропорциональность пенсии заработку может увеличить социальную напряжённость.

Таблица 2. Пенсии разных стран

Страна	Пенсионный возраст, лет		Пенсионные отчисления, %	Пенсия			
				в USD в месяц		% зар-платы	среднее по годам
	М	Ж		соц. мин	средняя		
США	65	65	12–15	280	700	50–80	вся жизнь
Великобритания	65	60	13–20	600			
Германия	65	65	20,3	300		≤70	3 последних
Франция	60	60	16,4		700	50	11 последних

1.3. Отечественные системы. В довоенные и первые послевоенные годы колхозники не имели ни паспортов, ни пенсий. Они составляли тогда большую часть населения СССР, так что пенсионная система была не всеобъемлющей. «Оттепельные» годы дали колхозникам полные гражданские права. Они стали получать также пенсии, хотя размер их был существенно меньше, чем у рабочих и служащих. Общий уровень всех пенсий был установлен в те годы, когда экономика страны ещё восстанавливалась, и был невысоким.

Система 1957 г. Подъём экономики позволил СССР ввести в конце 1950-х годов новую пенсионную систему, охватывающую всё население. На тот момент эта система была в целом хорошей. Пенсии при полном рабочем стаже (25 лет для мужчин и 20 для женщин) составляли 55 % зарплаты последнего года работы. При превышении стажа они увеличивались на 1 % за каждый год, а поскольку этот стаж (но не более 20 %) был невысок по сравнению с зарубежными нормами (см. п. 1.2), то большая часть пенсионеров имела 75 % от последнего заработка, что также соответствовало системам развитых стран. Предельный размер пенсии был ограничен, но достаточно большой величиной — он вдвое превышал среднюю зарплату в стране в те годы. Таким образом, максимальная пенсия обеспечивала хороший уровень жизни, и даже минимальная (70 % от минимальной зарплаты) — удовлетворительный.

Разброс зарплат в стране был тогда небольшим, так что на «потолок» в пенсии выходил только небольшой слой высокооплачиваемых людей (но для ряда их категорий были установлены надбавки). Поэтому не было «уравниловки» — размер пенсии существенно зависел от зарплаты (хотя гораздо слабее — от стажа).

Однако, абсолютность всех цифр и наличие «потолка» сказались со временем. Рост средних зарплат (рис. 1) первые 10–15 лет был связан с повышением реального уровня жизни, а в последующие годы — с инфляцией. Этот рост оказался значительным, и уже к 1980 г. средняя пенсия сравнялась с максимальной. Возникла «уравниловка», а сам уровень максимальной пенсии обеспечивал уже лишь скромное существование. Положение сильно ухудшилось к 1990 г., когда уровень жизни пенсионеров стал существенно уступать 1958 году. Особенно это сказалось на тех, кто вышел на пенсию много лет назад, когда абсолютные цифры зарплат и пенсий были существенно ниже.

Наконец, если пенсионер хотел продолжать работу, то пенсия ему не выплачивалась. Только для непрестижных неквалифицированных работ делалось исключение (инженер мог работать гардеробщиком, но не инженером). Это дискриминационная мера, во-первых, придавала пенсии характер милостыни, что возмущало общественное мнение. Во-вторых, она лишала государство многих квалифицированных специалистов.

Пенсионный возраст 60 лет для мужчин был ниже, чем в развитых странах, но это было разумно, поскольку медицинское обслуживание и продолжительность жизни также были ниже. Однако то, что пенсионный возраст для женщин был на 5 лет меньше, не укладывалось в экономические и медицинские рамки: средняя продолжительность жизни у женщин была больше, чем у мужчин. Это была чисто пропагандистская акция, результаты которой приходится расхлёбывать и сейчас.

Эта система имела ещё один недостаток — если рабочий стаж хотя бы на месяц не дотягивал до полного, то трудовая пенсия не

полагалась. Можно было получить лишь нищенскую пенсию по случаю потери кормильца (в 5–6 раз ниже максимальной). Это обычно касалось женщин, ведущих домашнее хозяйство при многих детях, то есть косвенно подталкивало к уменьшению рождаемости.

Начало 1990-х. В РСФСР действовала пенсионная система СССР. Во время самоутверждения Российской Федерации, к 1990 г., ясно выявились все недостатки описанной выше пенсионной системы. Однако именно она, без переработки, была включена в законодательство РФ в 1990 г. [1].

«Шоковая терапия» реформ 1992 г. сопровождалась катастрофической инфляцией (более 30-кратной в 1992 г. и 10-кратной в 1993 г.). Экономика страны была подорвана, сбережения населения фактически пропали, абсолютные размеры пенсий выглядели неприличными. Поэтому была введена система индексации пенсий на инфляцию. Всё это носило характер латания дыр: за 1991–1997 гг. было принято 26 законодательных поправок, не считая Указов Президента и Постановлений Правительства [2]. Одна из поправок позволила работающим пенсионерам получать пенсию.

Была также разумно определена пенсия при неполном рабочем стаже — пропорционально имеющемуся стажу.

Однако абсолютный размер даже максимальной пенсии с учётом индексации оказался настолько малым (360 % от минимальной пенсии), что далеко не обеспечивал прожиточного минимума. Жизненный уровень пенсионеров многократно уступал 1990 году, не говоря уже о 1958 г. Около 1/3 части пенсионеров, не имевших других источников существования, продолжали работать легально, а многие прирабатывали уличной торговлей и другими способами, не поддающимися государственному учёту и контролю.

Это было одной из важных причин недовольства в стране, а голосование на президентских выборах 1996 г. показало, насколько оно сильно. Требовались перемены.

Система 1997 г. [3], введённая с января 1998 г., разрабатывалась в то время, когда инфляция существенно замедлилась и казалось, что экономика стабилизируется. Это была попытка вернуться к системе 1957 г., одновременно улучшив её кое в чём.

Во-первых, в заработок для исчисления пенсий включили все выплаты, на которые начисляются страховые взносы в Пенсионный фонд (ранее учитывались лишь выплаты по основному месту работы, но не по совместительствам, а это нередко составляло большую долю заработка).

Во-вторых, заработок для исчисления пенсий стали делить на среднюю зарплату в России (для старых лет — в РСФСР) за тот же период времени. Это отношение, названное индивидуальным коэффициентом пенсионера (ИКП) автоматически вводило поправку на инфляцию, какой бы сильной та не была. Далее этот ИКП умножался на тот процент, который фигурировал в системах 1957 г. и 1990 г. (напомним, что он не

превышал 75%). Полученная величина множилась на среднюю зарплату в стране. Поэтому и далее при инфляционном росте номинальной зарплаты соответственно повышалась пенсия.

Основным недостатком этой системы было установление невысокого «потолка» пенсии: ИКП ограничивался величиной 1,2, так что 75% от него давали максимум пенсии, равной всего 0,9 средней зарплаты. Даже этот предел был установлен не сразу, в 1998 г. он был ещё в 1,7 раза ниже. Это много хуже, чем в 1958 г., когда отношение максимальной пенсии к средней зарплате составляло примерно 2. В итоге значительная часть пенсионеров получила максимальную пенсию, что означало «уравниловку».

Объяснялось это плачевным состоянием экономики и казны Пенсионного фонда. По той же причине новая система была распространена только на неработающих пенсионеров. Работающие сохраняли право получать пенсии лишь по системе 1990 г., где «потолок» был в 3–4 раза ниже. Эта дискриминация вызвала отрицательную реакцию общества.

Описанная система существенно улучшила положение пенсионеров на момент её введения — январь 1998 г. Она могла стать хорошей и долгоживущей, если увеличить предельный ИКП до ≈ 3 и распространить её на работающих пенсионеров. Однако финансовый кризис августа 1998 г. резко ухудшил положение всего населения страны, в том числе — пенсионеров. За 2–3 года экономика снова стабилизировалась, но пенсионеры обратили внимание, что абсолютные цифры пенсий в рублях заметно меньше, чем можно ожидать по средним зарплатам в стране. Официальное разъяснение дал председатель правления Пенсионного фонда РФ. Он сообщил, что для исчисления пенсий ИКП умножается на среднюю зарплату по стране, но не ту, которая начислена в данном месяце, а на ту, пенсионные отчисления с которой поступили в Пенсионный фонд. А реально выплаченная зарплата существенно (в 1,5–2 раза) меньше начисленной.

На самом деле это разъяснение является дезинформацией по двум причинам.

Во-первых, если зарплата выплачивается, например, с запозданием на месяц, то апрельская зарплата выплачивается в конце мая вместо конца апреля. Но ведь тогда в конце апреля поступает зарплата за март, так что в казну Пенсионного фонда всё равно ежемесячно поступает положенная сумма отчислений. Только в одной ситуации могла быть разница: если бы суммарная задолженность по зарплате в стране нарастала. Однако эта задолженность последние 4 года неуклонно сокращалась с ≈ 80 млрд руб. в начале послекризисного 1999 года до 29–34 млрд руб. в начале 2003 г.

Во-вторых, средняя задержка зарплаты по стране все эти годы была небольшой. Это легко подсчитать по балансам. Например, в апреле 2003 г. средняя начисленная зарплата составляла 5035 руб./мес, количество работающих — 75–90 млн чел. (эти данные наименее точны), а суммарная задержанная зарплата — 34 млрд руб., то есть

≈500 руб/мес на одного работающего. Это означает задержку на 0,1 месяца, всего 3 дня. Даже в начале 1999 г., когда суммарная задолженность по зарплате составляла 80 млрд руб., а средняя зарплата была в 4 раза меньше, средняя задержка зарплаты по стране составляла 1 месяц. Другое дело, что на большинстве предприятий платили аккуратно, а на меньшей части — задерживали по полгода.

Отсюда видно, что истинные поступления в Пенсионный фонд РФ практически соответствовали начисленной зарплате. Исчисляя трудовые пенсии по иным цифрам, Пенсионный фонд в 1,5–2 раза обманывал пенсионеров. Ниже будет показано, что этот обман был вынужденным (то есть недоплаченные деньги расходовались на законные цели). Поэтому упрекать правление Пенсионного фонда следует не за малый размер пенсий, а за сокрытие истинных причин этого.

Система 2001 г. До сих пор все отечественные системы были чисто налогово-распределительными. Для построения неплохой системы достаточно было лишь снова стабилизировать экономику и модернизировать систему 1997 г., как указано выше. Однако в правительственных кругах возобладала идея накопительных пенсий по образцу Запада и была принята система 2001 г. [4–6], введённая в действие с 1 января 2002 г. Эта система состоит из трёх законов — «О государственных пенсиях», «О трудовых пенсиях» и «О пенсионном страховании». Проанализируем её.

Обязательные страховые пенсионные начисления на зарплату составляют, как и ранее, 28%. Из них 14% непосредственно поступают в Пенсионный фонд для финансирования трудовых пенсий. Остальные 14% идут на финансирование государственных пенсий и сначала зачисляются в госбюджет, а оттуда переводятся в Пенсионный фонд (что не имеет смысла, и лишь затрудняет проверку целевого использования денег).

Трудовая пенсия состоит из трёх частей — *базовой, страховой и накопительной*. Базовая — это тот минимум, который получают все, независимо от зарплаты и стажа. Поэтому её правильней называть социальной. Страховая часть пропорциональна суммарному заработку за всю жизнь, а сведения о нём накапливаются на индивидуальном счёте каждого работника в Пенсионном фонде. На её финансирование направляется 14% страховых взносов граждан 1956 г. рождения и старше, 12% для 1957–1966 гг. рождения и 8% 1967 г.р. и моложе. Остальная часть из поступающих в Пенсионный фонд 14% направляется на накопительный индивидуальный счёт, деньги которого вкладываются в какие-то доходные ценные бумаги. Накопленная сумма с доходами идёт на выплату накопительной части пенсии. Наличие социального минимума и то, что страховая и накопительная части пенсии пропорциональны суммарному заработку за всю жизнь, на первый взгляд разумны.

Однако вырисовываются недостатки. Во-первых, накопительная часть (а это почти половина поступлений в Пенсионный фонд) слабо

защищена от биржевых крахов, обесценивающих акции и доходные бумаги, в которые они вложены. Если гражданин сам выбрал компанию, управляющую его накоплениями, он и отвечает за последствия. Если он не сделал выбор, то его средствами управляет Внешэкономбанк. Но этот банк является ответчиком по долгам СССР, и при дефолте его крах неизбежен. В обоих случаях гражданин теряет пенсионные накопления.

Таким образом, накопительная часть пенсии фактически не защищена ни от инфляции, ни от расхищения. Однократный сильный «обвал» рубля, многолетняя заметная инфляция или крах инвестиционного фонда могут практически уничтожить накопление всей жизни. Первая же такая серьёзная встряска поднимет вопрос о совершенствовании Закона.

Страховая часть пенсии довольно надёжно защищена от хищений; крахи инвестиционных фондов её не затрагивают, ибо она имеет налоговый тип «собрал–раздал». Однако её размер зависит от условия накопленного за всю жизнь капитала, измеренного в рублях. При быстрой инфляции сумма заработка за все предыдущие годы оказывается малой по сравнению с вкладом последних лет и даже последнего года. Тогда размер пенсии будет фактически определяться заработком последнего года, что не оптимально. Ползучая или обвальная инфляция тоже могут её обесценить. В Законе предусмотрен некоторый механизм индексации пенсии при инфляции. Однако, его счётчик включается только с момента назначения пенсии; а назначаться она будет по обесцененному условному капиталу (сам этот условный капитал не индексируется при инфляции).

Во-вторых, уже назначенную пенсию обесценивает инфляция. Поэтому введена индексация пенсий, но её механизм нечётко и зависит от решений Правительства, а волюнтаризм в таких вопросах опасен. Базовая часть пенсии в той же мере защищена от хищений. Механизм её индексации довольно разумно защищает её от инфляции. Но базовая часть — это пособие на бедность. В результате через ряд лет накопительная и страховая части существенно сократятся, и всем пенсионерам останется только это пособие, независимо от их трудового вклада. Можно оценить этот срок точнее: при 5%-ной годовой инфляции это 20–30 лет, при 10%-ной это 10–16 лет, при 20%-ной это 6–9 лет (что наиболее близко к нынешней ситуации), а при обвале рубля — немедленно.

Кроме трудовых пенсий, есть государственные. К ним относятся скромные пенсии никогда не работавшим старикам и инвалидам, устанавливаемые примерно на уровне базисной; это вполне естественно. Но в эту же категорию входят пенсии военным и за выслугу лет — государственным служащим. Пенсии за выслугу лет выплачиваются дополнительно к трудовым и могут достигать 75 % оклада. Они существенно превышают средние трудовые пенсии и финансируются за счёт

первых 14 % страховых начислений, переводимых в Пенсионный фонд через госбюджет.

Таким образом, высокие пенсии государственных служащих финансируются за счёт обычных категорий работающих. То же относится к категориям, имеющим право раннего выхода на пенсию (вредные профессии, артисты балета и т. д.). Поэтому уровень трудовых пенсий невысок, хотя процент пенсионных отчислений вдвое выше, чем в развитых странах. То же скрытое дотирование привилегированных категорий за счёт обычных было в предыдущих пенсионных системах. Именно это стеснялись открыто сказать официальные лица.

Если бы ранее установленные пенсии при переходе с системы 1997 г. на систему 2001 г. были пересчитаны строго по законам, их размеры были бы существенно выше. Так, в 2001 г. средняя начисленная зарплата составляла 3578 руб/мес (а по последнему кварталу ещё выше), что давало «потолок» пенсии 3220 руб/мес. Зачтена же была лишь средняя зарплата ≈ 1700 руб/мес и максимальная трудовая пенсия ≈ 1530 руб/мес, что в 2,1 раза меньше.

Таким образом, недостатки действующей пенсионной системы велики. Общественное мнение не одобряет эту систему в целом. Особенные нарекания вызывает накопительная часть пенсии, так как все помнят печальный пример финансовых пирамид. Уже сейчас в правительственных кругах идёт речь о необходимости усовершенствования пенсионной системы. В частности, предлагается отменить накопительную часть пенсии и соответственно уменьшить страховые отчисления.

Поэтому предлагаемая авторами система является очень актуальной. Она содержит автоматическую регулировку на случай сильной инфляции и даже обвала. Она предусматривает также ручную настройку всего одним-двумя управляющими параметрами. Это несложно. Поэтому она очень скоро может оказаться востребованной.

1.4. Основные требования. Любая пенсионная система является компромиссом между пожеланиями граждан и возможностями экономики страны.

Пожелания отдельных граждан обычно эгоистичны: поменьше отчислять в пенсионный фонд, получать пенсию пораньше и побольше, причём приближающуюся к последней зарплате и не обесценивающуюся до самой смерти (а это в среднем 15–20 лет). Граждане болезненно реагируют ещё на два фактора. Они против «уравниловки» — чем больше рабочий стаж и выше зарплата, тем больше должна быть пенсия. Однако одновременно они против слишком большого — в десятки раз — отличия пенсий (как и против слишком большого отличия зарплат). Невыполнение этих пожеланий приводит к социальной напряжённости. Политик или партия, игнорирующие это, рискуют потерять голоса избирателей.

Напряжённость особенно усиливается, если система противоречит сложившимся в обществе понятиям социальной справедливости (кото-

рые формируются политическими партиями, профсоюзами, экономическими магнатами и церковью с помощью средств массовой информации). Сейчас каждый знает, что с его фонда зарплаты всю жизнь отчисляли в пенсионный фонд определённый процент. Поэтому достаточно болезненно воспринимаются не только невысокий уровень пенсий или «уравниловка», но также дискриминация работающих пенсионеров («пенсия — не милостыня, я её заработал»).

Система 2001 г. ликвидировала большинство дискриминационных ограничений для работающих пенсионеров, но некоторые всё же остались. Например, установлены надбавки к пенсиям лауреатам престижных премий (Государственной и т.п.), но лишь после прекращения работы. Поскольку такие люди работают, как правило, до самой смерти, такая льгота выглядит издевательством.

До сих пор медленная инфляция с 1960 по 1990 г., быстрая инфляция начала 90-х годов или мгновенные «обвалы» типа 1998 г. приводили к обесцениванию пенсий и требовали внесения поправок в Закон о пенсиях. Это также серьёзный источник напряжённости. Вдобавок в такой ситуации нелегко разумно оценить пенсионный взнос за всю трудовую жизнь. Но нельзя исключить такие ситуации в будущие годы, а хороший пенсионный закон должен работать без поправок десятки лет. Необходимо продумать такую систему, причём она должна согласовываться с возможностями экономики и быть достаточно простой.

Простота закона желательна во всех аспектах. Нынешний Закон содержит много разных видов пенсий (по инвалидности, потере кормильца, работе в тяжёлых условиях и т.п.) и большое количество частных случаев. Надо объединить их насколько возможно, упростить формулировки и сократить текст. Он должен быть настолько простым, чтобы работник пенсионного фонда смог бы легко рассчитать пенсию и объяснить это пенсионеру, а гражданин со средним образованием мог бы легко понять это объяснение.

Очевидно также, что система должна учитывать возможности экономики страны, то есть автоматически увеличивать или уменьшать пенсии при подъёмах или спадах экономики. Для этого в неё надо включить управляющие параметры, определяемые по ежемесячным или ежеквартальным поступлениям в Пенсионный фонд. Параметров может быть 2–3, но не более, иначе управление станет недостаточно прозрачным.

Наконец, систему следует согласовать с остальными законами, в первую очередь с Законом о семье и браке. Типична ситуация, когда жена сидит дома с детьми и долго не работает. Она теряет рабочий стаж, а её продвижение по служебной лестнице замедляется, что приводит к потере в зарплате. Всё это существенно уменьшает её пенсию, но на пенсии мужа не отражается. Но пенсия является неким имущественным правом, а согласно Семейному кодексу имущество, нажитое в браке, поровну принадлежит супругам (если они не заключали брачного договора с иными условиями). Значит, на пенсионные счета

жены и мужа за время нахождения в браке должны поступать одинаковые суммы — отчисления от половины их суммарного заработка.

Следует поощрять также наличие 2–3 детей в семье, либо путём надбавок к пенсиям, либо путём снижения страховых начислений или налогов. Сейчас это особенно важно из-за катастрофической демографической ситуации (естественная убыль населения в 1994–2002 гг. составляла почти 0,6 % в год!).

В следующих двух пунктах рассмотрено, какое устройство пенсионной системы должно максимально обезопасить её от финансовых потрясений.

1.5. Государственная или частная?. Западные страны пришли к смешанной пенсионной системе: государство собирает налоги, и за их счёт обеспечивает социальную пенсию всем нетрудоспособным; частные или государственные фонды принимают взносы, помещают эти средства в высоконадёжные акции, а доход обращают на выплаты пенсий. Поэтому пенсия каждого гражданина содержит социальную и трудовую части.

Например, выше говорилось, что в США обязательные пенсионные отчисления составляют $\approx 13\%$ зарплаты, из которых половину уплачивает предприятие, половину — сам работающий (для сравнения, в России в 2000 г. эти цифры составляли 28 % и 1 %). Могут быть дополнительные добровольные отчисления, а престижные предприятия могут привлекать работников своими дополнительными отчислениями. Всё это обеспечивает пенсию без «потолка» размером 70–100 % от индивидуальной зарплаты гражданина.

Эти цифры обычно приводят сторонники частных пенсионных фондов. Однако они умалчивают о многом. Высоконадёжные акции мало доходны (в США это 4–5 %), а организаторы частных фондов хотят получить доход и для себя. Они приобретают более доходные акции и рискуют обанкротиться. Возможно и прямое мошенничество. Банкротство частного пенсионного фонда вызывает потерю пенсий у большого числа лиц. Возникает взрыв социального недовольства, причём в России он будет традиционно направлен против Правительства.

На Западе пенсионные фонды существуют около полутора веков. Недобросовестные или неумелые организаторы в основном исчезли. Возраст фонда в десятки лет является весомой гарантией его надёжности. Дополнительную гарантию даёт обязательное государственное страхование фондов, хотя это снижает доходность фондов. Вдобавок, оно позволяет вновь организуемым фондам попробовать свои силы без ущерба для граждан. В западных условиях умеренный страховой процент покрывает риски фондов и, тем самым, государства. Кроме того, фонды страхуют друг друга, так что совокупность частных фондов всё более напоминает единый фонд, причём тесно связанный и почти сросшийся с государством.

В нынешней России старых частных пенсионных фондов нет. Мошенничество сильно развито. Государству пришлось бы или вводить огромный страховой процент для государственного страхования рисков фондов, или нести огромные убытки при банкротстве частных фондов, или столкнуться с ещё большими социальными взрывами, чем при крахе финансовых пирамид.

Поэтому сейчас разумно лишь следующее решение. Основной пенсионный фонд должен быть государственным. Частные пенсионные фонды можно разрешить, но жёстко подходить к выдаче лицензий, а в печати постоянно разъяснять, что государство не отвечает по их обязательствам (вплоть до неперемennого включения этих слов в текст частных пенсионных договоров).

1.6. Распределение или накопление? Сейчас узаконили и усиленно пропагандируют фонды накопительного типа (бывшие эмеритальные кассы), ссылаясь на западный опыт. Как решающий аргумент приводят то, что внесённые в фонд отчисления приносят прибыль. Это увеличивает сумму к моменту выхода на пенсию, что позволяет уменьшить процент пенсионных отчислений (13 % в США, 28 % в России).

На самом деле это заблуждение. Накопить на старость можно квартиру и запас одежды. Но человеку нужны еда, топливо, медицинская помощь и другие вещи и услуги, которые невозможно запасти на 15–20 лет вперёд. Запас денег не решает проблему — деньги есть не будешь. Нужно то, что производит подросшее новое поколение. Часть его труда идёт на содержание пенсионеров, что фактически является введением налога. Этот налог взимается с работающих и *немедленно* расходуется. Таким образом, суть так называемой накопительной пенсии та же, что в первобытном обществе: вырасти молодое поколение, а в старости оно накормит тебя.

Возрастание вклада за счёт прибыли — тоже миф. Деньги сами не размножаются (только Буратино мог поверить лисе Алисе и коту Базилио, что деньги растут). Эта прибыль есть недооплачиваемый труд, то есть тоже налог, но не прямой, а менее заметный косвенный. Вдобавок любой накопительный фонд обесценивается при финансовом кризисе, затрагивающем ряд государств или хотя бы данное. А такие кризисы повторяются последние годы, и причиной их являются современные фондовые спекуляции (что серьёзно беспокоит даже такого фондового игрока, как Дж. Сорос).

Для не вполне убеждённого читателя приведём такой пример. Пусть всё работоспособное население исчезнет (уедет за границу), а все накопительные фонды останутся в сохранности. Встанут все заводы и электростанции, замрёт транспорт и трубопроводы, рубль снова станет полностью «деревянным» — неконвертируемым. Что тогда будут делать пенсионеры?

Формирование пенсионных фондов частично через прямые налоги (отчисления), а частично через косвенные на самом деле не влияет на

размеры пенсий, а лишь усложняет расчёты и делает процедуру менее прозрачной.

Кроме того, это увеличивает армию чиновников и дельцов, управляющих фондами; а им нужны не только высокие оклады (которые у чиновников много больше, чем у рядовых работников), но и дорогостоящие офисы с техникой, что стоит гораздо дороже. Поэтому существенно возрастают накладные расходы фонда, а это приводит к уменьшению пенсий.

Наконец, ещё страшнее мошенничество. Гражданин 40 или более лет платит в накопительный фонд страховые взносы и даже получает оттуда подтверждения в их получении. А выходя на пенсию, он может узнать, что где-то в середине его трудового стажа, 20 лет назад, нечистый на руку делец запустил лапу в казну фонда и сбежал за границу или даже благополучно умер, так что взыскать не с кого и невозможно: срок исковой давности истёк. При распределительном же фонде всё собранное за месяц тут же и расходуется, за исключением резервного переходящего остатка. Возможна ежемесячная проверка финансов, что существенно затрудняет мошенничество.

Отсюда следует важный вывод: *целесообразен пенсионный фонд налогово-распределительного типа, а не накопительного. Но величина пенсий должна определяться с учётом трудового вклада за всю жизнь, причём инфляция не должна обесценивать вклад прошлых лет.*

Необходима разработка соответствующей системы и методики расчёта её параметров.

1.7. Виды и размеры. В современном законодательстве пенсия зависит от многих обстоятельств. Это надо в основном сохранить, упростив формулировки и методику расчёта. Размеры пенсий необходимо согласовать с экономическими возможностями страны, которые в основном можно охарактеризовать тремя величинами: отношением прожиточного минимума к среднему заработку, процентом отчислений в Пенсионный фонд и отношением числа работающих к числу пенсионеров.

Трудовой вклад за всю жизнь надо определять следующим образом: заработок человека за каждый месяц (точнее, величину его пенсионных отчислений за месяц) делить на среднюю зарплату по стране за тот же месяц и полученные коэффициенты суммировать. Тогда результат будет справедливым даже при очень сильной инфляции. Нынешний метод пенсионной системы 1997 г., когда заработок в рублях суммируется за ряд лет, явно хуже: при сильной инфляции вклад старых лет практически не учитывается. То же относится к пенсионной системе 2001 г., где все отчисления разных лет суммируются на едином лицевого счёте застрахованного гражданина без поправки на инфляцию. Какой-либо «обвал» рубля немедленно обесценит все уже накопленные взносы.

Размер поступлений в Пенсионный фонд — социальный налог — зависит от числа работающих и средней зарплаты в стране, то есть от состояния экономики (фактически то же было бы для фондов накопительного типа). Поэтому величины пенсий в Законе невозможно обозначать в абсолютных денежных суммах. Их следует устанавливать в долях средней зарплаты на данный момент. Тогда они будут автоматически повышаться при укреплении экономики (или понижаться при кризисе).

Государственную пенсионную систему следует ориентировать на людей с низкими и средними заработками: высокооплачиваемые более инициативны и сами способны дополнительно обеспечить себе старость. Может быть, для пенсий придётся вводить «потолки», иначе Пенсионный фонд не выдержит сверхвысоких выплат частным нотариусам и т. п. Этот вопрос требует количественного анализа.

1. Минимальной должна быть социальная пенсия, выплачиваемая тем людям, которые достигли пенсионного возраста, не являясь инвалидами и никогда не работали (без уважительных причин). Её размер следовало бы привязать в основном к величине прожиточного минимума. Однако при нынешнем состоянии экономики он будет заметно меньше прожиточного минимума.

2. При инвалидности по общему заболеванию или травме в любом возрасте назначается социальная пенсия плюс надбавка на лечение. Размер надбавки зависит от группы инвалидности и измеряется в фиксированных долях социальной пенсии.

3. Работавшим людям по достижении пенсионного возраста назначаются социальная плюс трудовая пенсии. Размер трудовой части пенсии берётся по вкладу за всю жизнь, но не линейно, а по более слабо возрастающей кривой. Абсолютная величина трудового вклада привязывается в основном к среднему заработку по стране.

4. Ситуация при инвалидности по производственному заболеванию или травме требует более детальной проработки. Социальная пенсия и надбавка на лечение очевидны, но трудовой вклад следует увеличить. Если бы не травма, человек мог работать ещё ряд лет и достичь более высокого заработка.

5. Требуется проработки вопрос о пенсиях по случаю потери кормильца. Для не работавшего супруга (и вообще супругов) имеется разумное решение: из пенсионных отчислений с заработка каждого супруга половина заносится на его пенсионный счёт, а половина — на счёт другого супруга. Это хорошо согласуется с Семейным кодексом. Для детей и других иждивенцев решение не очевидно, но возможно такое решение: каждому иждивенцу — половина (поскольку второй кормилец жив) социальной пенсии, плюс $1/(n+1)$ -я доля трудовой пенсии кормильца, где n — число иждивенцев. Если кормилец умер вследствие профессионального заболевания или травмы, то трудовая часть пенсии увеличивается аналогично п. 4.

6. Пенсии на вредных производствах также требуют дополнительной проработки. Разумно устанавливать для таких предприятий повышенный процент пенсионных страховых отчислений, а людям предоставить право выходить на пенсию раньше. Возможно при этом выплачивать только трудовую часть пенсии, и лишь по достижении общего пенсионного возраста — социальную часть.

Такой подход возможен и к пенсиям военнослужащих, которых увольняют из армии в 45 лет, то есть задолго до достижения общего пенсионного возраста (и они зачастую находят менее тяжёлую работу).

7. Надбавки за особые заслуги перед страной (Герой и др.) следует устанавливать не в процентном повышении пенсии, а кратно социальной пенсии или в долях средней зарплаты.

8. Для нормального роста народонаселения необходимо иметь 2–3 ребёнка в семье. Чтобы ребёнок полноценно развивался, мать должна уделять ему повышенное внимание хотя бы до школьного возраста. Это приведёт к сокращению её трудового стажа и пенсионного счёта. Поэтому трудовая пенсия женщин, родивших и вырастивших детей, должна содержать существенные льготы за первых 1–3 детей и небольшие за четвёртого-пятого (слишком большой рост народонаселения нецелесообразно поощрять по многим причинам).

Выводы. Конечная цель работы — разработка гибкой системы пенсий и простых методов расчёта, позволяющих оперативно определять основные параметры формул по текущим данным о состоянии Пенсионного фонда. Черновой проект такой системы изложен в следующих разделах, хотя для «доводки» ряда деталей потребуется привлечение узких специалистов — юристов и экономистов.

Данная концепция существенно отличается от предлагаемых сейчас проектов. Она позволяет построить более справедливую и долговечную пенсионную систему, которая ослабит социальную напряжённость в обществе.

§ 2. Балансные расчёты средних величин

2.1. Обозначения. В данной главе используются следующие обозначения:

t — средний нормативный трудовой период гражданина за всю жизнь (в годах);

T — среднее нормативное время жизни после оформления пенсии (в годах);

N — численность (в млн чел.);

z — «нормированная зарплата», то есть отношение индивидуального заработка гражданина за месяц к средней зарплате по стране за тот же месяц;

ζ — индивидуальный заработок гражданина (в руб/мес);

q — пенсионное начисление на зарплату (в формулах — в виде десятичной дроби, в таблицах — в %);

Q — полное нормированное отчисление за всю жизнь данного гражданина в Пенсионный фонд (в виде десятичной дроби);

s — отношение социальной части пенсии к средней зарплате;

g — надбавка на инвалидность (в долях средней зарплаты);

h — надбавка к пенсии за заслуги (в долях средней зарплаты);

r — надбавка за рождение и выращивание детей (в долях средней зарплаты);

p — отношение трудовой части пенсии к средней зарплате;

m — отношение прожиточного минимума к средней зарплате;

ν — число иждивенцев умершего кормильца (целое);

μ — коэффициент накладных расходов Пенсионного фонда: зарплата аппарата, помещения, оборудование (десятичная дробь или %);

P — отношение средней по стране пенсии к средней зарплате; будем также называть его базовой пенсией;

k — отношение числа пенсионеров к числу работающих;

Z — безразмерная (нормированная, отнесённая к средней зарплате по стране) индивидуальная зарплата гражданина за всю жизнь;

$m, ж$ — индексы, обозначающие мужчин или женщин (отсутствие индекса — среднее по стране);

n — индекс, обозначающий категорию инвалидности, заслуги или трудности профессии ($n = 0$ — обычные условия, отсутствие индекса — среднее значение по стране);

j — индекс месяца;

α, β — параметры пенсионной формулы;

$f(\zeta)$ — функция распределения по зарплатам;

a, b, c — параметры этой функции распределения.

Таким образом, T есть среднее расчётное время жизни всех пенсионеров страны, безотносительно к полу или профессии; $T_{ж}$ — только женщин; T_{nm} — мужчин n -й категории профессии и т. д.

2.2. Принципы расчёта. Для расчёта инвестиционного накопительного фонда потребовался бы аппарат дифференциальных уравнений, так как надо находить динамику роста капитала и доходов фонда. Кроме того, для достоверного расчёта нужны надёжные демографические и экономические прогнозы, причём не на ближайшие 2–4 года, а на 20–40 лет. Такие прогнозы нереальны, а без них развитие инвестиционного фонда будет путём слепого. Это ещё один аргумент против таких фондов.

Для налогово-распределительного фонда динамика относится только к резервам (переходящему остатку). Основная часть деятельности описывается балансными соотношениями — ежемесячным поступлением пенсионного налога и выплатой пенсий. В этих соотношениях надо лишь предусмотреть выделение некоторой части средств в резерв. Такая задача несравненно проще, она не требует прогнозов, и все

исходные данные для неё содержатся в базе данных Пенсионного фонда и вполне достоверны.

Для расчётов высокой точности нужна подробная база данных, включающая функции распределения ныне работающих по зарплатам, пенсионеров — по их бывшим зарплатам (из которых им назначена пенсия), численности работников с особыми условиями труда, имеющими льготные условия для выхода на пенсию (вредные производства, армия, балет и т. п.), демографические кривые распределения смертности по возрастам и многое другое. На основе такой базы данных можно разрабатывать автоматизированную систему управления Пенсионным фондом.

Нам эти данные недоступны. Поэтому возможно лишь произвести оценки на основе публикуемых данных по средним зарплатам, средней продолжительности жизни, прожиточному минимуму и собственному опыту жизни. Этими данными надо пользоваться с осторожностью. Например, средняя продолжительность жизни мужчин в России сейчас составляет 58,5 лет. Однако из этого нельзя делать вывод, что все мужчины умирают до выхода на пенсию, так что им пенсии вообще не нужны. На самом деле эта средняя продолжительность включает смертность взрослых от ранних инфарктов и т. п., а также детскую смертность. Количество пенсионеров при этом остаётся значительным.

Оценки. В пенсионную систему 2001 г. заложена средняя расчётная продолжительность жизни пенсионера после достижения пенсионного возраста $T = 18$ лет. Одна и та же цифра установлена для мужчин и женщин, для всех профессий (хотя на самом деле они должны сильно отличаться). Пройгнорируем пока различия профессиональных условий и рассмотрим отдельно мужчин и женщин. Введём для них численности N_m , N_j и средние продолжительности жизни на пенсии T_m , T_j , а также полную численность пенсионеров N . Запишем для этих величин очевидные балансные соотношения:

$$N_m + N_j = N, \quad \frac{T_m N_m + T_j N_j}{N} = T; \quad (1a)$$

Введём также две разности:

$$N_j - N_m = \nu, \quad T_j - T_m = \vartheta. \quad (16)$$

Из соотношений (1) следуют выражения

$$\begin{aligned} N_m &= \frac{N - \nu}{2}, & N_j &= \frac{N + \nu}{2}, \\ T_m &= T - \vartheta \frac{N + \nu}{2N}, & T_j &= T + \vartheta \frac{N - \nu}{2N}. \end{aligned} \quad (2)$$

По официальным данным Пенсионного фонда, $N = 38,6$ млн чел., $T = 18$ лет. Величину ν оценим из следующих соображений. По данным переписи 2002 г. разница полного числа мужчин и женщин в стране равна 10,07 млн чел. Мужчин рождается несколько больше, но их

смертность в детском и предпенсионном возрасте выше, так что можно принять $\nu \approx 10,1$ млн чел. (женщин больше, чем мужчин). Сложнее оценить ϑ . По той же переписи средняя продолжительность всей жизни у женщин 72,0 года, а у мужчин 58,5 лет. Учитывая, что женщины выходят на пенсию на 5 лет раньше, получим $\vartheta \approx 72,0 - 58,5 + 5 = 18,5$ лет. Подставляя эти цифры в (2) получим такие величины:

$$N_m \approx 14,3 \text{ млн}, \quad N_{\text{ж}} \approx 24,3 \text{ млн}, \quad T_m \approx 6,4 \text{ года}, \quad T_{\text{ж}} \approx 24,9 \text{ лет}. \quad (3)$$

Обсудим эти оценки. То, что женщин-пенсионеров почти вдвое больше, чем мужчин, и что они живут дольше, хорошо известно. Но полученная цифра для $T_{\text{ж}}$ неправдоподобно велика. Она больше, чем разница между средней продолжительностью жизни и пенсионным возрастом женщин $72,0 - 55,0 = 17$ лет, что в принципе невозможно! Более того, средняя расчётная жизнь *всех* пенсионеров $T = 18$ лет также превышает эту цифру.

Это противоречие объясняется «двойной бухгалтерией» Пенсионного фонда. В одну кучу складываются все пенсионеры: работники нелёгких профессий, пенсионеры силовых ведомств с их гораздо более ранним выходом на пенсию, а также инвалиды, получающие пенсию всю жизнь или значительную часть её. Число пенсионеров по инвалидности составляет $N_{\text{и}} = 1,9$ млн чел. Численность пенсионеров силовых ведомств не публиковалась, но можно грубо оценить её как $N_{\text{с}} \approx 8$ млн чел. (текущая численность только органов МВД составляет не менее 2 млн чел., а возраст выхода на пенсию ранний). К ним следует присоединить спортсменов и балерин, выходящих на пенсию ещё раньше.

Аккуратные оценки сделать невозможно, но правдоподобные цифры даёт разбиение на три группы (т — обычную трудовую, с — силовую, и — инвалидов) с такими параметрами:

$$\begin{aligned} N_{\text{т}} \approx 28,7 \text{ млн.}, \quad N_{\text{с}} \approx 8,0 \text{ млн.}, \quad N_{\text{и}} \approx 1,9 \text{ млн.}, \\ T_{\text{т}} \approx 13,5 \text{ лет}, \quad T_{\text{с}} \approx 26 \text{ лет}, \quad T_{\text{и}} \approx 57 \text{ лет}; \end{aligned} \quad (4)$$

их стандартное усреднение как раз даёт $N = 38,6$ млн и $T = 18$ лет. При этом примерно 4 млн пенсионеров-силовиков и 0,4 млн инвалидов достигли общепенсионного возраста, так что полное число стариков-пенсионеров составляет ≈ 33 млн чел.

Но для стариков хорошо выполняются соотношения $N_{\text{ж}} - N_{\text{м}} \approx 10$ млн., $T_{\text{ж}} \approx 72,0 - 55 = 17$ лет. Отсюда легко получить для них, аналогично (1)–(2), балансные оценки

$$N_{\text{м}} = 11,5 \text{ млн.}, \quad N_{\text{ж}} = 21,5 \text{ млн.}, \quad T_{\text{м}} = 7 \text{ лет}, \quad T_{\text{ж}} = 17 \text{ лет}. \quad (5)$$

Это даёт среднее время жизни стариков на пенсии $T = 13,5$ лет. Данные цифры выглядят гораздо более правдоподобно. Разумеется, и здесь число женщин-пенсионеров и их продолжительность жизни на пенсии

больше, чем для мужчин. Данные (4), (5) берутся за основу при оценках предлагаемой пенсионной системы.

Принимаются для оценок следующие средние цифры и другие данные.

1. Мужчина работает до выхода на пенсию в среднем 40 лет (с 20 после армии до 60) и живёт на пенсии в среднем 7 лет. Эти цифры относятся к не вредным профессиям. Для профессий с пенсионными льготами продолжительность работы уменьшается на 5–20 лет в зависимости от категории льготы, а продолжительность жизни на пенсии примерно настолько же увеличивается.

2. Женщина не служит в армии, но она рождает и воспитывает детей. Поэтому она работает в среднем 30 лет (с 18 до 55 минус 7 лет на детей) и живёт на пенсии в среднем 17 лет. Для профессий со льготными условиями пенсии цифры изменяются аналогично приведённым выше.

3. Число мужчин и женщин трудоспособного возраста примерно одинаково, так что в среднем гражданин до пенсии работает при общих профессиях 35 лет и живёт на пенсии 13,5 лет. При льготных профессиях цифры соответственно изменяются.

4. В 1-м квартале 2001 г. средняя зарплата в России, засчитываемая Пенсионным фондом, составляла 1522 руб. Максимальная пенсия по пенсионной системе 1997 г. была 90 % от средней зарплаты (1370 руб). Минимальная пенсия была установлена в размере 600 руб (это 40 % средней зарплаты) и стала играть роль социальной. Пенсии по пенсионной системе 1991 г. оказались меньше минимальной, так что эта система фактически была упразднена, хотя формально она существовала.

Зависимость существующих сейчас пенсий от индивидуальной зарплаты гражданина показана на рис. 2. По системе 1991 г. пенсия вообще не зависит от зарплаты и равна минимальной. По системе 1997 г. для зарплат менее 53 % от средней зарплаты величина пенсии не зависит от зарплаты и равна минимальной. В диапазоне 53–120 % средней она пропорциональна минимальной; при зарплатах более 120 % средней она выходит на «потолок». Видно, что эта пенсионная система устарела и толкала граждан к сокрытию даже умеренных заработков порядка двукратного среднего, не говоря о высоких.

5. Прожиточный минимум сейчас существенно различается в разных регионах. В Москве он в начале 2001 г. вдвое превышал среднюю «пенсионную» (то есть засчитываемую Пенсионным фондом) зарплату и раза в полтора — начисленную зарплату. В начале 2003 г. он стал близок к средней начисленной зарплате. Состав потребительской корзины, употребляемый при его расчёте, является скорее благим пожеланием, чем реальностью. При этом в него не входит даже полная оплата коммунальных услуг, не говоря о стоимости самого жилья (учёт всего этого поднял бы прожиточный минимум до небес). Поэтому увязывать

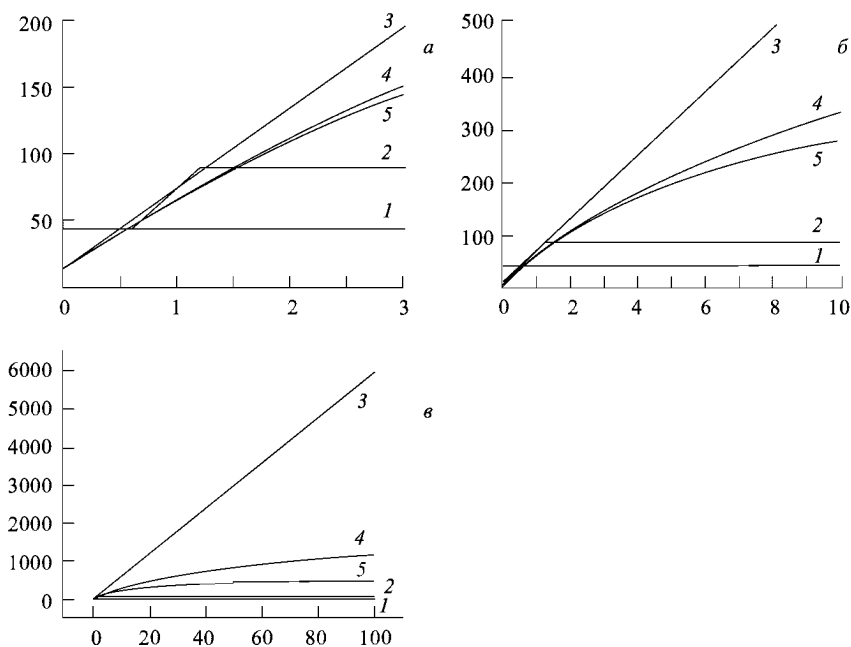


Рис. 2. Зависимость пенсии $(p + s)\%$ от базовой пенсии P . Обозначения линий: 1 — система 1991 г., 2 — система 1997 г., 3 — линейная формула (11), 4 — корневая формула (12), 5 — рациональная формула (13) с параметрами $\alpha = 0,6$ и $\beta = 8 \dots$ (см. табл. 3). Три рисунка показывают диапазоны средних, больших и очень больших заработков

размер социальной и, тем более, трудовой пенсии с прожиточным минимумом пока нецелесообразно и даже невозможно.

6. Авторы не располагали данными о профессиональной или общей инвалидности, а также по случаям потери кормильца. Эти виды пенсии разрабатывались по аналогии с существующими системами и старой системой СССР, но вносились изменения для увязки с Семейным кодексом и понятиями социальной справедливости.

2.3. Средняя пенсия. Пренебрежём пока безработицей, смертностью и инвалидностью граждан, не достигших пенсионного возраста. Тогда получим первое балансное соотношение для отношения числа пенсионеров к числу работающих:

$$k = \frac{T'}{t} \quad \left(k < \frac{T'}{t} \right). \quad (6)$$

При учёте безработицы, смертности и инвалидности равенство надо заменить приведённым в скобках неравенством. Неравенство мало полезно для расчётов, но всех требуемых статистических данных у авторов нет.

Баланс сбора пенсионного налога и выплаты пенсий даёт второе соотношение для общих (не льготных) профессий:

$$kP = (1 - \mu)q. \quad (7)$$

Оно позволяет определить среднюю пенсию P , если задан пенсионный налог q . Наоборот, если надо подогнать пенсию P под заданное в Законе значение, оно показывает необходимый для этого пенсионный налог q .

Подставляя (6) в (7), получим третье балансное соотношение:

$$\frac{PT}{t} = (1 - \mu)q. \quad (8a)$$

Его можно применять не только к общим пенсиям, но и к льготным отраслевым. Для этого вместо общего трудового стажа t надо подставлять уменьшенный отраслевой, вместо средней жизни обычного пенсионера T — настолько же увеличенный срок, а вместо среднего пенсионного налога q — отраслевой q_n :

$$\frac{PT'_n}{q_n} = (1 - \mu)q_n. \quad (8б)$$

Из (8б) видно, что для обеспечения льготникам той же средней пенсии P надо увеличивать пенсионный налог на предприятиях этой отрасли.

2.4. Социальная пенсия. Это та милостыня, которую общество подаёт здоровому человеку, достигшему общего пенсионного возраста (60/55 лет) и всю жизнь бездельничавшему. Общество не имеет моральных обязательств перед таким человеком. Поэтому уровень социальной пенсии должен быть низким, отбивающим охоту бездельничать, но позволяющим иметь очень скромное жильё, питание и одежду. Если прожиточный минимум должен обеспечивать достойное существование, то социальная пенсия может быть существенно меньше.

На 1-й квартал 2001 г. минимальная пенсия составляла 450руб/мес, то есть 40% средней «пенсионной» зарплаты; от настоящей же начисленной зарплаты это было 20–27%. К апрелю 2003 г. средняя начисленная зарплата превысила 5000 руб/мес, а социальная пенсия возросла до 636руб/мес, что составляет около 13% зарплаты. Учитывая ближайшую перспективу роста средней зарплаты не только в денежном, но и реальном исчислении, можно ориентировочно предложить социальную пенсию на уровне $s = 0,15 = 15\%$ средней начисленной зарплаты по стране. Возможна и меньшая цифра, так как подавляющее большинство граждан будут иметь ещё и трудовую часть пенсии, поскольку они проработали хотя бы несколько лет или имеют другие основания для дополнительной пенсии.

Гражданин, заработавший трудовую часть пенсии, должен получать и её, и социальную пенсию. Если лишать его социальной части, то возможны ситуации, когда немного работавший человек будет иметь пенсию меньше, чем вообще не работавший.

Здесь есть одна оговорка. При профессиях с пенсионными льготами трудовая часть пенсии сейчас выдаётся не с общего пенсионного возраста (60/55 лет), а с меньшего. Но социальная часть должна выдаваться не с льготного возраста, а с общего. Очень многие пенсионеры-льготники на самом деле вполне трудоспособны и продолжают работать в более лёгких отраслях до общего пенсионного возраста, а их трудовая часть пенсии существенно больше социальной. Поэтому предлагаемое решение разумно.

Возможно и другое, менее привычное решение этой проблемы. Оно будет рассмотрено ниже.

2.5. Трудовая пенсия. Для простоты пока примем, что гражданин всю жизнь работал в n -й отрасли, где установлены нормативный трудовой стаж t_n , нормативное время жизни пенсионера T_n и отчисления в Пенсионный фонд q_n (за время жизни человека эти отчисления могут законодательно изменяться). Трудовая пенсия определяется по суммарному отчислению в Пенсионный фонд за всё время работы (J месяцев). Для этого за каждый j -й месяц определяются индивидуальная зарплата z_j , нормированная на среднюю зарплату по стране в тот же месяц, и пенсионное отчисление q_{nj} той отрасли, где гражданин тогда работал. Составляется сумма по всем месяцам (а при совместительстве — и по всем местам работы):

$$\sum_{j=1}^J q_{nj} z_j; \quad (9)$$

она показывает полное нормированное отчисление в Пенсионный фонд за всю жизнь.

Проще всего было бы определить пенсию следующим образом. Всю сумму (9), внесённую гражданином в Пенсионный фонд, последний обязан выплатить ему за расчётное время жизни пенсионера T_n . С учётом содержания самого Фонда это даёт размер пенсии:

$$P = \frac{1 - \mu}{12T_n} \sum_{j=1}^J q_{nj} z_j; \quad (10)$$

именно эту пенсию целесообразно называть *базовой* (коэффициент 12 появился потому, что T_n измеряется в годах). Разумеется, отдельные пенсионеры умрут раньше или позже расчётного периода, а кто-то вообще не доживёт до пенсии, но долю тех, кто умер раньше, выплатят должителям. Формула (10) лежала в основе расчёта эмеритальных касс, только в неё вводились поправки на доходность накопленных средств, но не учитывалась инфляция (расцвет этих касс приходился на время обращения золотых и серебряных монет).

Пенсия (10) не обесценивается инфляцией, и это её главное преимущество перед всеми ранее предлагавшимися способами. Это объясняется тем, что в неё входят не абсолютные цифры в рублях, которые

могут обесцениться, а ежемесячные отношения к средней зарплате по стране на те же моменты, а не на данный. Поэтому значительное отчисление, поступившее на лицевой счёт в Пенсионном фонде 20 лет назад, будет и сейчас оставаться столь же значительным. Разумеется, это большой страховой риск, и его может взять на себя только государственный распределительный фонд. Ни для накопительного, ни для частного фонда такое непосильно.

Формула (10) проста, но она не учитывает три обстоятельства. Во-первых, человек может сменить профессию. Тогда какое T_n надо подставлять в формулу? Не случайно эмеритальные кассы были обычно отраслевыми, а не всеобщими. Во-вторых, формула отдаёт все деньги тем, кто платил взносы, а надо часть денег перераспределить на социальные пенсии тем, кто не работал. В-третьих, полная трудовая пенсия не должна быть меньше социальной. Это также требует перераспределения денег от более богатых к более бедным. Заметим, что перераспределение может сделать только государственный фонд, но не частный.

Забудем пока про отраслевой фактор и рассмотрим перераспределение. Каждый застрахованный должен получать социальный минимум s плюс трудовую часть пенсии p . Трудовая часть пенсии должна монотонно возрастать с ростом базовой пенсии P . Она должна начинаться с нуля и либо асимптотически выходить на «потолок», либо расти неограниченно, но не быстрее линейной зависимости. Приведём три подходящие аппроксимации:

$$p = \alpha P, \quad \frac{dp}{dP} = \alpha, \quad P(\infty) = \infty; \quad (11)$$

$$p = \frac{\alpha P}{\sqrt{1 + 2P/\beta}}, \quad \frac{dp}{dP} = \frac{\alpha(1 + P/\beta)}{(1 + 2P/\beta)^{3/2}}, \quad P(\infty) = \infty; \quad (12)$$

$$p = \frac{\alpha P}{(1 + P/\beta)}, \quad \frac{dp}{dP} = \frac{\alpha}{(1 + P/\beta)^2}, \quad P(\infty) = \alpha\beta. \quad (13)$$

Здесь α , β — численные параметры закона, которые надо разумно подбирать. Формула (11) даёт линейный рост без «потолка», как принято в пенсионной системе 2001 г. В формуле (12) рост медленней линейного, но «потолок» пенсии тоже нет. Последняя формула (13) имеет «потолок» пенсии, но выход на него довольно медленный, так что при малых и средних зарплатах все три зависимости близки к линейным.

Производные dp/dP показывают, на сколько копеек увеличивается трудовая пенсия при увеличении базовой пенсии на очередной рубль. Параметры α и β подбирают для получения необходимых значений средней пенсии. Проанализируем предложенные аппроксимации.

Очевидно, надо выбирать $\alpha < 1$, чтобы в трудовую часть пенсии $p(P)$ входила лишь часть базовой, а остальная часть шла на перераспределение. В системе 2001 г. на перераспределение и лицевой счёт идут одинаковые страховые отчисления по 14 %, так что $\alpha = 0,5$.

Здесь предлагается немного увеличить это значение и взять $\alpha = 0,6$ (это означает начисление $\approx 11\%$ в Госбюджет и 17% прямо в Пенсионный фонд).

Подобрано такое значение коэффициента $\beta \approx 8$, чтобы значение «потолка» пенсии в формуле (13) составляло ≈ 5 . На апрель 2003 г. при средней начисленной зарплате ≈ 5000 руб/мес это составляет 25000 руб/мес. Примерно такие пенсии получают министры и депутаты Государственной Думы, так что этот «потолок» (а он будет расти в рублях при повышении средней зарплаты) не удовлетворит лишь удачливых дельцов, а для подавляющей части населения вполне приемлем.

Все эти законы представлены в табл. 3 и на рис. 2 для выбранных значений параметров. Для удобства показано полное значение пенсии $p + s$ с социальной добавкой. Для сравнения в табл. 3 и на рис. 2 приведены также трудовые пенсии по системе 1991 г. и 1997 г.

Таблица 3. Зависимость пенсии $(p + s)\%$ от базовой пенсии P (в долях средней зарплаты) для разных систем и данного проекта (11)–(13) с параметрами $\alpha = 0,6$ и $\beta \approx 8$; в действующей системе 2001 г. первое число соответствует отсутствию банкротств, а второе — банкротству накопительных фондов

P	Система 1991 г.	Система 1997 г.	Система 2001 г.	Проект (11)	Проект (12)	Проект (13)
			$P_{\max} \dots P_{\min}$			
0,0	45	45	45 ... 45	15	15	15
0,1	45	45	46 ... 45	21	20	20
0,2	45	45	47 ... 46	27	26	26
0,3	45	45	49 ... 47	33	32	32
0,4	45	45	50 ... 48	39	37	37
0,5	45	45	52 ... 49	44	43	43
0,6	45	45	53 ... 49	51	48	48
0,7	45	52	54 ... 50	56	53	53
0,8	45	59	56 ... 51	62	58	58
0,9	45	67	57 ... 52	69	63	63
1,0	45	74	59 ... 53	74	68	68
1,2	45	90	61 ... 54	87	78	77
1,4	45	90	64 ... 56	99	87	86
1,6	45	90	67 ... 57	110	96	95

Продолжение табл. 3

P	Система 1991 г.	Система 1997 г.	Система 2001 г.	Проект (11)	Проект (12)	Проект (13)
			$P_{\max} \dots P_{\min}$			
1,8	45	90	70 ... 59	122	104	103
2,0	45	90	73 ... 61	134	113	111
2,2	45	90	75 ... 62	146	121	118
2,4	45	90	78 ... 64	159	129	126
2,6	45	90	81 ... 65	171	136	133
2,8	45	90	84 ... 67	183	144	139
3,0	45	90	87 ... 69	195	151	146
3,5	45	90	94 ... 73	225	168	161
4,0	45	90	101 ... 77	254	185	175
4,5	45	90	108 ... 81	284	200	188
5,0	45	90	115 ... 85	315	215	200
5,5	45	90	122 ... 89	345	229	211
6,0	45	90	129 ... 93	374	243	221
6,5	45	90	136 ... 97	405	256	231
7,0	45	90	143 ... 101	435	269	240
7,5	45	90	150 ... 105	465	281	248
8,0	45	90	157 ... 109	495	293	256
8,5	45	90	164 ... 113	525	304	263
9,0	45	90	171 ... 117	555	315	270
9,5	45	90	178 ... 121	585	326	277
10,0	45	90	185 ... 125	615	336	283
20,0	45	90	325 ... 205	1215	507	360
40,0	45	90	605 ... 365	2415	742	418
100,0	45	90	1445 ... 845	6015	1197	463

Система же 2001 г., не показанная авторам, даёт прямую, с коэффициентом $\alpha = 0,5$, которая при отсутствии биржевых крахов будет лежать немного ниже линии, соответствующей формуле (11). Но если

крах обанкротит накопительные пенсионные фонды (что в российских условиях почти неизбежно), то наклон этой прямой уменьшится до $\alpha = 8/28 \approx 0,3$, и она пойдёт вдвое ниже прямой (11).

Видно, что при очень малом заработке (базовой пенсии $P \rightarrow 0$) все три кривые начинаются в одной точке ($p = 0$ у всех) и при этом касаются друг друга. Они очень близки для низких базовых пенсий $P \leq 1$, то есть для низкооплачиваемого населения, и лишь при $P \geq 2$ начинают заметно различаться. Далее выше всех идёт линейная кривая (11), ниже — корневая кривая (12) и ещё ниже — дробь (13). Последние две кривые в начальной точке имеют одинаковую вторую производную. Поэтому они очень близки на начальном участке при $P \leq 2$ и лишь для наиболее состоятельных граждан с $P \geq 10$ существенно различаются.

Какой из вариантов (11)–(13) предпочтительней? Для низкооплачиваемого населения все кривые дают практически одинаковые пенсии. Высокооплачиваемые граждане предпочтут линейную зависимость (11). В этом есть определённая выгода для государства, поскольку людям становится выгодным легализовать доходы. При кривых (12) и (13) появляется больше стимулов к утаиванию высоких доходов. Вдобавок расчёты пенсии с нелинейными формулами труднее объяснить среднеобразованным слоям населения, а это достаточно важный фактор (их можно заменить ломаными, но это громоздко и также хуже воспринимается населением). Зато кривые (12)–(13) сильнее перераспределяют средства от богатых к бедным, что позволяет уменьшить процент пенсионных начислений. Надо произвести количественный расчёт последнего эффекта, для чего требуется функция распределения населения по доходам. Ниже будет сделан такой расчёт.

Ныне действующая пенсионная система 2001 г. даёт пенсии несколько ниже линии (11) даже при благополучной работе накопительных пенсионных фондов. При их крахе она даёт почти вдвое меньшие пенсии. Поэтому она опасна — может нарушить стабильность в обществе.

2.6. Пенсионный возраст женщин. Во всех пенсионных системах СССР и России, включая ныне действующую систему, он на 5 лет меньше мужского. Это было установлено ещё в первые послереволюционные годы из пропагандистских соображений, без серьёзного медицинского и экономического обоснования. Рассмотрим, к чему это приводит.

Сейчас всё население обслуживается единым Пенсионным фондом и отчисляет на пенсии единый страховой процент $q = 28\%$. Предположим, что организовано два таких фонда — мужской и женский. Какой процент q должен установить каждый фонд?

Пока ограничимся не льготными профессиями с пенсионным возрастом 60/55 лет. Данные по численности и нормативной продолжительности жизни таких пенсионеров оценены выше (4). Нормативный трудовой стаж мужчин $t_m = 40$ лет. Для женщин по действующей

системе 2001 г. было $t_{\text{ж}} = 30$ лет; но одновременно с этим рассмотрим вариант стажа $t_{\text{ж}} = 35$ лет для бездетных женщин. Накладные расходы Пенсионного фонда примем $\mu = 10\%$. Потребуем, чтобы средняя (базовая) пенсия $P = 75\%$. Тогда по формуле (8а) легко рассчитать страховые отчисления q .

Результаты расчёта приведены в табл. 4. Видно, что чисто мужской пенсионный фонд обеспечивает высокую пенсию даже при скромных отчислениях $q_{\text{м}} < 15\%$ (первый столбец). Если же учесть, что только 3/4 мужчин доживают до пенсионного возраста, а их взносы достаются остальным, то отчисления можно снизить до уровня $q_{\text{м}} \approx 11\text{--}12\%$.

Таблица 4. Страховые отчисления мужчин и женщин

	Мужчины	Женщины		
		с детьми	без детей	Проект
N млн чел	11,5	21,5	21,5	
пенс. возраст лет	60	55	55	60
t лет	40	30	35	40
T лет	7	17	17	12
P %	75	75	75	75
q %	14,6	47,2	40,5	25,0
$q_{\text{ж}}/q_{\text{м}}$	—	3,2	2,8	1,7

Женщинам с детьми пришлось бы платить в свой фонд огромные отчисления $q_{\text{ж}} \approx 47\%$ (второй столбец). У бездетных женщин рабочий стаж больше, но всё равно отчисления остаются большими $q_{\text{ж}} \approx 40$ (третий столбец). Это больше «мужского» отчисления соответственно в ≈ 3 и 2 раза. Таких отчислений не выдержат фонды зарплаты предприятий, и начнётся или сокращение зарплаток, или дискриминация женщин при приёме на работу. Если же уменьшить процент отчислений до уровня мужского фонда, то женские пенсии станут в 2–3 раза ниже мужских при одинаковой зарплате. Это тоже неприемлемо.

При общем Пенсионном фонде и одинаковом отчислении $q_{\text{м}} = q_{\text{ж}}$ это означает следующее. Мужчины в среднем работают на 1/3 дольше женщин, а живут на пенсии в 2,5 раза меньше! Поэтому в Пенсионный фонд мужчины вносят больше, а расходует он в основном на женщин. Это считается социально необходимым. Однако провести в жизнь такое решение может только государственный фонд. Любой частный пенсионный фонд быстро поймёт, что пенсии женщин выгодно страховать только при большом проценте пенсионных отчислений (при учёте допенсионной смертности коммерческий процент для женщин

примерно вчетверо выше, чем для мужчин!) и найдёт лазейки в законах для дискриминации женщин.

Таким образом, экономика заставляет увеличивать пенсионный возраст женщин до уровня мужского. В табл. 4 приведён такой расчёт, где для женщин взят пенсионный возраст 60 лет (четвёртый столбец). При этом одновременно на 5 лет сокращается нормативная продолжительность жизни на пенсии и настолько же увеличивается рабочий стаж (он взят таким же, как у мужчин, а причина этого будет объяснена в пункте о семье). Поэтому страховой процент сразу резко сокращается до $q_{ж} = 25\%$. Это ещё существенно выше мужского, но уже меньше даже нынешних единых 28% . Такое решение является социально более справедливым, хотя даже оно даёт преимущество женщинам.

К такому же решению подталкивают объективные медицинские показатели. Средняя продолжительность жизни женщин в России на 13,5 лет больше, чем у мужчин. Это свидетельствует о лучшем здоровье и более длительном сохранении трудоспособности.

Именно таким путём были вынуждены решать аналогичную проблему развитые страны. Если сначала в ряде стран женский пенсионный возраст был ниже мужского, то сейчас он почти во всех странах одинаков (см. табл. 2). Совсем недавно этот процесс прошёл в бывших социалистических странах. Например, в ГДР пенсионный возраст был, как и в СССР, 60 лет у мужчин и 55 — у женщин. Сейчас он стал по законам ФРГ 65 лет у тех и других.

Переход. Предлагаемое изменение неизбежно, но оно будет очень непопулярным. Его могут позволить себе только Правительство и Государственная Дума, пользующиеся достаточным доверием народа. Однако есть два обстоятельства, позволяющие внести такое изменение сравнительно безболезненно.

Во-первых, экономическая ситуация в стране пока ещё нелёгкая. Из 38,6 млн пенсионеров около 11 млн продолжают легально работать, чтобы сносно жить. Из них многие уже достигли предлагаемой возрастной планки, так что им подъём пенсионного возраста безразличен. Если учесть, что работают в основном те пенсионеры, которые ещё недавно перешагнули возрастную пенсионную планку, то окажется, что лишь примерно 2,5 млн женщин будет задето предлагаемым подъёмом их пенсионного возраста до 60 лет.

Во-вторых, это изменение можно увязать с ревизией Пенсионного фонда и выявлением его скрытых резервов. Выше показывалось, что это должно привести к увеличению пенсий в 1,5–2 раза. Но это повышение надо проводить совместно с повышением женского пенсионного возраста и не сразу, а поэтапно, растянув переходный процесс на 5–10 лет. Это ослабит общественное недовольство.

Для должной постепенности надо каждую половину календарного года увеличивать общий пенсионный возраст женщин на 3 месяца. За 10 календарных лет он увеличится на требуемые 5 лет. При этом

каждое отдельное повышение будет задевать не более 150 тыс. женщин, хотящих бросить работу по достижении пенсионного возраста.

К сожалению, ещё более благоприятным моментом для такой реформы были тяжёлые 1999–2000 гг., но этот момент упущен. Чем благополучнее будет экономика, тем меньше пенсионеров станет работать, и переход затруднится.

Старение. Все развитые страны сталкиваются с проблемой старения населения. Увеличивается продолжительность жизни. Растёт отношение числа пенсионеров к числу работающих. Так, в Германии сейчас оно составляет 1:4, а через 10 лет ожидается 1:2. Выходов из этого положения три: 1) уменьшать размеры пенсий; 2) увеличивать проценты пенсионных отчислений; 3) увеличивать возраст выхода на пенсию.

Большинство стран считает нежелательным заметно уменьшать пенсии или увеличивать проценты пенсионных отчислений во избежание социальных взрывов. Постепенное небольшое увеличение пенсионного возраста (поскольку время жизни на пенсии много меньше трудового стажа, даже маленькое увеличение пенсионного возраста сильно влияет). Это оправдывается улучшением здравоохранения и повышением трудоспособности немолодых людей.

Правда, на небольшое уменьшение пенсий им также приходится идти. В Германии полная пенсия в 2000 году составляла 70% от средней зарплаты за последние 3 года работы, а в 2001 году она снижена до 67%.

Сегодня в России проблема старения кажется неактуальной: здравоохранение и условия жизни настолько плохие, что в ближайшие годы повышение общего пенсионного возраста сверх 60 лет немыслимо. Хотя демографы и делают пессимистические прогнозы о старении, их оценки крайне ненадёжны (например, перепись 2002 г. показала, что общую численность населения они предсказали с ошибкой в 2,5 млн чел). Однако мы строим стабильную пенсионную систему, рассчитанную на десятки лет. Поэтому необходимо допустить возможность увеличения продолжительности жизни и соответственное увеличение общего пенсионного возраста.

Очевидно, состояние здоровья населения потребует повышения общего пенсионного возраста не ранее, чем через 10 лет, когда пенсионные возрасты женщин и мужчин уже сравняются. Тогда их общий пенсионный возраст можно будет повышать также небольшими ступеньками, на 3 месяца раз в календарный год или в полгода.

По-видимому, хорошим индикатором повышения будет расчётная продолжительность жизни пенсионера: целесообразно поддерживать её на уровне 15–17 лет. Если статистические данные укажут, что она становится больше этой величины, пенсионный возраст надо соответственно увеличивать. Разработать такой критерий и включить его в систему управления Пенсионным фондом несложно.

Оценки. В формулах (4) приведена оценка нынешнего состава пенсионеров с грубой разбивкой на группы, а в формулах (5) даны цифры только для тех, кто перешагнул нынешний пенсионный возраст $M/J = 60/55$ лет. Как изменятся эти цифры, если ввести для женщин пенсионный возраст 60 лет, и ликвидировать все профессиональные льготы по более раннему выходу на пенсию?

Количество пенсионеров-инвалидов не изменится. Из пенсионеров-льготников (силовиков и др.) примерно половина, то есть 4 млн., отпадёт сразу: их возраст меньше общепенсионного. Из остальных пенсионеров число мужчин и их расчётное время жизни не изменятся. Но для женщин расчётное время жизни на пенсии сократится на 5 лет, а число уменьшится соответственно на 5/17 долю. В итоге получим

$$N_{\text{и}} = 1,9 \text{ млн}, \quad N_{\text{тм}} = 11,5 \text{ млн}, \quad N_{\text{тж}} = 15,2 \text{ млн},$$

$$N_{\text{т}} = 26,7 \text{ млн}, \quad N = 28,6 \text{ млн}, \quad (14)$$

$$T_{\text{и}} = 57 \text{ лет}, \quad T_{\text{м}} = 7 \text{ лет}, \quad T_{\text{ж}} = 12 \text{ лет}, \quad T_{\text{т}} = 10 \text{ лет}, \quad T = 13 \text{ лет}.$$

Здесь приведены также полные численности и расчётные времена жизни трудовых пенсионеров и всех пенсионеров. Видно, что учёт даже небольшого числа инвалидов с их большой продолжительностью жизни на пенсии существенно увеличивает T .

Общее число пенсионеров при предлагаемом изменении уменьшается почти на четверть. Оценим единые страховые отчисления q , которые нужны для обеспечения того же высокого уровня пенсии $P = 75\%$, что и в табл. 4:

$$q = 27\%, \quad \tilde{q} = 20,8\%; \quad (15)$$

здесь $\tilde{q}$ описывает ситуацию, когда трудовые пенсионеры обеспечивают только себя, а q включает обеспечение инвалидов за счёт трудящихся. Оба норматива ниже, чем нынешние 28%, и обеспечивают уровень пенсии примерно вдвое выше, чем сейчас.

Это иллюстрирует эффект предлагаемой реформы.

2.7. Отраслевые льготы. Действующее законодательство предусматривает несколько категорий труда со льготными условиями выхода на пенсию. Льготы состоят в более раннем сроке выхода на пенсию, а в ряде профессий — в получении дополнительной пенсии за выслугу лет. Эту систему можно изложить в более простой форме, чем она ранее записывалась. Все профессии разбиваются на шесть категорий по вредности или тяжести условий труда:

— нулевая — обычные условия труда, пенсионный возраст 60 лет (у мужчин; сейчас у женщин он на 5 лет ниже, но всюду далее, согласно п. 2.6, будем предполагать, что женский пенсионный возраст подтянут до мужского);

— первая — повышенная интенсивность, вредность или ухудшенные условия труда (химическое производство, полевые экспедиции и т. п.), пенсионный возраст 55 лет;

— вторая — высокая интенсивность, вредность или плохие условия труда (подземные работы, радиоактивное производство и т. п.), пенсионный возраст 50 лет;

— третья — особо тяжёлые или опасные условия труда (армия и другие силовые структуры), пенсионный возраст 45 лет;

— четвёртая — сверхвысокая интенсивность труда (балет), пенсионный возраст 40 лет;

— пятая, которой формально пока нет — экстремальные условия труда (профессиональный спорт), пенсионный возраст 35 лет.

Потребуем, чтобы для всех категорий труда средняя пенсия была одной и той же $P = 75\%$, как принято в табл. 4. Тогда для $n > 0$ надо увеличить пенсионные отчисления q_n . Расчёт по формуле (86) даёт результаты, приведённые в табл. 5.

Таблица 5. Отраслевые страховые начисления для получения пенсии
 $P = 75\%$

Категория отрасли n	Пенсионный возраст, годы	Трудовой период t_n , годы	Пенсионный период T_n , годы	Страховые начисления q_n , %	q_n/q_0
0	60	40	10	20,8	1,00
1	55	35	15	35,7	1,71
2	50	30	20	55,6	2,67
3	45	25	25	83,3	4,00
4	40	20	30	125,0	6,00
5	35	15	35	194,4	9,33

Прокомментируем результаты. Чтобы обеспечить равный уровень пенсии при сокращённом трудовом периоде, приходится увеличивать процент пенсионных отчислений, причём для высоких категорий — очень сильно. Если этого не сделать, то пенсионеры высоких категорий будут получать свои пенсии за счёт нулевой категории. Это означало бы государственное дотирование определённых отраслей, причём не явное, а скрытое. Такое недопустимо, поскольку искажает истинную экономику.

Это ещё можно было понять в условиях социалистической экономики: когда государство является собственником всех предприятий, оно может платить и отчислять как в явной форме, так и в косвенной. Но для частных предприятий, находящихся в руках разных хозяев, нельзя открыто заставлять работников одного предприятия оплачивать пенсии чужих работников.

Таким образом, традиционная для СССР практика составления государственного списка вредных профессий и льгот никогда не была

по настоящему хорошей, а в современных условиях совершенно непригодна. Нужны иные решения. Обсудим их.

Трудность, вредность или опасность профессии довольно относительны. В среднем работа в химическом цехе более опасна, чем в бухгалтерии, а работа под землёй ещё опаснее. Однако многое зависит от организации безопасности труда как вообще на производстве, так и на конкретном рабочем месте. Хорошие фильтры и герметизация на химическом заводе сильно уменьшают вредность. Снабжение геологических экспедиций хорошими рациями и вертолётами существенно увеличивает безопасность. Оборудование рабочих помещений кондиционерами и хорошим освещением облегчает условия труда. Все эти нюансы невозможно отразить в общегосударственных списках.

В общих списках есть и немалые пробелы. Так, работа профессиональных спортсменов (к ним относится и «любительская» спортивная элита) до сих пор вообще не отнесена к тяжёлым, хотя спортсмен в 25 лет считается зрелым, в 30 лет — пожилым и в 35 лет — стариком. Студенческим стройотрядам не давали никаких льгот, хотя там работали по 12–14 часов в день без выходных, то есть рабочая неделя составляла 80–100 часов вместо стандартных 40. Такова же интенсивность труда многих учёных, конструкторов и руководящего состава производств, также не имеющих льгот за это.

Поэтому государство должно (совместно с профсоюзами) составить такой перечень категорий трудности и общую инструкцию об аттестации отраслей, предприятий и рабочих мест. Отнесение предприятий и рабочих мест к той или иной категории должно производиться совместной комиссией Государственной инспекции по труду и профсоюзов, аттестующей каждое предприятие раз в 5 лет. При существенном ухудшении или улучшении условий труда проводится внеочередная аттестация. Результаты аттестации доводятся до сведения всех работников.

Страховые взносы. Отнесение труда к более тяжёлой категории не означает автоматического увеличения пенсий или сокращения рабочего стажа. Трудовая пенсия каждого работника определяется только его личным счётом в Пенсионном фонде. Чтобы этот счёт был больше, предприятие должно отчислять больший страховой процент q_n . Если оно этого не делает, никакого увеличения пенсии работник не получит.

Поэтому государство должно законодательно устанавливать только минимальный процент q , который предприятия перечисляют в Пенсионный фонд за каждого работника. Повышенный процент q_n за тяжесть труда должен устанавливаться в коллективных договорах дирекции и профсоюза на каждом предприятии с дифференциацией по цехам и рабочим местам. Этот список должен доводиться до сведения всех работников и включаться в трудовой договор при приёме на работу.

Очевидно, эти проценты будут неодинаковыми даже у однотипных предприятий одной отрасли. Процветающее предприятие будет привлекать работников не только высокой зарплатой, но и больши-

ми пенсионными отчислениями (именно так делают на Западе, что упоминалось в § 1). Это относится и к нулевой категории трудности: платная гимназия может установить более высокий процент, чем муниципальная школа. Такое может быть даже в чисто государственной отрасли: сравним финансовые условия службы в армии по призыву и по контракту.

Однако надо ясно понимать, что даже самое преуспевающее предприятие не в состоянии заплатить те огромные отчисления, которые соответствуют высоким категориям в табл. 5. Это значит, что повышенных отчислений хватит лишь для маленького сокращения рабочего стажа, гораздо меньшего, чем в табл. 5.

Пенсионный возраст. Выше отмечалось, что большинство отраслевых пенсионеров, достигнув льготного пенсионного возраста, остаются вполне здоровыми и трудоспособными. Они обычно продолжают работать на более лёгких работах, а нередко и на старой работе.

Это неудивительно: бытовые условия, условия труда и здравоохранение в среднем неуклонно улучшаются. Конечно, «реформа» 1990-х годов или Великая Отечественная война приводили к ухудшению условий, но после таких срывов наступает подъём. Но, несмотря на это, жизнь и работа сейчас существенно легче, чем в 1950-е годы и, тем более, в 1930-е. А 25-летний срок службы офицеров в армии как был установлен при Петре I, так и не изменялся 300 лет. Поэтому разумным является такое решение: *никакая профессия не даёт права на снижение пенсионного возраста.*

Если же здоровье работника существенно ухудшилось (неважно, по причине общего заболевания, профессионального заболевания или травмы), то медицинская комиссия устанавливает ему группу инвалидности, а *при инвалидности пенсионный возраст уменьшается.*

Детальнее последнее рассмотрено ниже.

Смена профессии. Если профессия стала непосильной для работника, а его пенсионный возраст не наступил, он вынужден искать более лёгкую работу. Но для новой работы у него может не быть профессиональной подготовки. Это серьёзная проблема, но хорошее решение для неё найдено в контрактной армии США.

Там вовсе не обещают немедленного выхода на пенсию по окончании армейской службы. Но по договору, который заключает контрактник, после окончания службы он имеет право на оплату его образования (в том числе высшего) за счёт бюджета армии. Таким образом, предприятие оплачивает получение новой профессии и даёт бывшему работнику возможность найти новую хорошо оплачиваемую работу. Именно такой пункт об оплате переподготовки должен включаться в трудовые договоры всех предприятий, где установлена повышенная тяжесть труда.

Нередко переход на другую профессию бывает психологической травмой (каково звезде балета стать репетитором или олимпийскому

чемпиону — тренером!). Но человек сам выбирает профессию и несёт ответственность за этот выбор.

Чиновники. Особо льготные условия выхода на пенсию имеют государственные служащие и депутаты высоких рангов. Их число на душу населения в несколько раз выше, чем было в СССР (и в десятки раз больше, чем в царской России) и продолжает расти. Зарплаты у них существенно выше, чем у других бюджетников, а при выходе на пенсию они могут получать за выслугу лет до 75 % последнего оклада, что нередко в несколько раз превышает среднюю пенсию в стране. Напомним, что выплачиваются эти пенсии не за счёт повышения пенсионных отчислений данной социальной группы, а за счёт половины пенсионных отчислений рядовых пенсионеров.

Ярким примером являются члены Государственной Думы. Во-первых, их число почти вдвое превышает число членов Конгресса США, хотя население России вдвое меньше, чем в США. Во-вторых, они законодательно установили себе оклады, равные окладу министра, что в 7–10 раз превышает зарплату профессора университета (а в США — лишь в 2–3 раза). В-третьих, они законодательно установили себе пенсии в размере 75 % оклада, если их пребывание в Государственной Думе длилось хотя бы 3 года (даже если всю предыдущую жизнь работали на малооплачиваемых должностях, так что их лицевой счёт в Пенсионном фонде ничтожен).

Ради объективности заметим, что так ведёт себя чиновничья элита всех стран. Аналогичные законы приняли ещё раньше депутаты Бундестага ФРГ. А во Франции государственные служащие — это особый слой вроде номенклатуры СССР: человека, попавшего в эти списки, нельзя исключить из них иначе, как по суду.

Высокие пенсии чиновников, с учётом их многочисленности и способа финансирования, серьёзно подрывают экономику Пенсионного фонда. Единственный разумный выход — лишить чиновников этих привилегий: *все пенсионеры, включая экс-президентов страны, должны получать пенсии на общих основаниях*, то есть по тому, какая сумма накопилась на их лицевом счёте в Пенсионном фонде. При этом краткосрочное пребывание на высоких должностях не сильно увеличит пенсию.

2.8. Особые заслуги. В действующем законодательстве установлено увеличение пенсии за некоторые выдающиеся заслуги — Героям России и СССР, героям социалистического труда, лауреатам Государственных премий. Эти надбавки увеличивают пенсию на определённый процент, что неразумно: для солдата-Героя надбавка окажется мизерной. Есть также персональные пенсии выдающимся артистам, спортсменам и т. п., устанавливаемые Указом Президента РФ. Их размеры велики, но они установлены в рублях, и поэтому обесцениваются инфляцией. Эти бессистемные надбавки целесообразно упорядочить.

Для этого надо ввести градацию заслуг перед государством. Например, будем считать отсутствие Правительственных наград нулевой категорией заслуг, и введём ещё 5 категорий:

- первая — наличие любых наград, имеющих статус Правительственных, в том числе, любые медали (включая медаль «Ветеран труда»);
- вторая — лица, награждённые любыми орденами, заслуженные работники различных отраслей хозяйства, заслуженные артисты СССР и России, бронзовые призёры Олимпийских игр и т. п.;
- третья — кавалеры высоких орденов страны (по списку), народные артисты СССР и России, заслуженные деятели науки, серебряные призёры Олимпийских игр и т. п.;
- четвёртая — кавалеры высших орденов России и СССР (по списку), лауреаты Государственных и Ленинских премий СССР и России, Олимпийские чемпионы;
- пятая — Герои России и СССР, Герои социалистического труда.

Попадание в категорию определяется не суммированием орденов, медалей и прочих заслуг, а по наивысшей заслуге.

Надбавки за заслуги следует устанавливать в долях средней зарплаты, а не в процентах от индивидуальной пенсии гражданина (иначе для солдата-срочника, ставшего Героем России, эта надбавка может оказаться мизерной, а удачливый бизнесмен даже за медаль получит много). Разумна шкала, приведённая в табл. 6. Она подчёркивает разницу между категориями — чем выше категория, тем больше разница между соседними категориями.

Таблица 6. Надбавки за заслуги

Категория n	0	1	2	3	4	5
Надбавка h_n	0,00	0,2	0,5	0,9	1,4	2,0

2.9. Члены семьи. В семью входят супруги, дети, родители, другие родственники и неродные люди. Из них некоторые могут быть нетрудоспособными (несовершеннолетние, старики, инвалиды) или не работать и вести домашнее хозяйство. При этом другие зарабатывают и считаются кормильцами. Существует ряд статей законов, определяющих пенсии нетрудоспособным членам семьи в случае потери одного или всех кормильцев. Эти статьи можно усовершенствовать следующим образом.

Супруг. Обычно женщина тратит несколько лет на рождение детей и уход за ними. Это существенно уменьшает её рабочий стаж, замедляет служебное повышение и рост зарплаты. В итоге её лицевой счёт в Пенсионном фонде оказывается гораздо меньшим, чем у мужчины. Если же жена преуспевающего работника вообще бросает работу и занимается только домашним хозяйством, то по нынешним законам

она становится иждивенцем и может претендовать только на мизерную социальную пенсию по потере кормильца.

Однако по Семейному кодексу всё имущество, нажитое в браке (за исключением подарков и наследства) является совместной собственностью обоих супругов. Лицевой пенсионный счёт является имущественным правом, так что он тоже в равной мере должен принадлежать супругам. Поэтому *половина пенсионных отчислений каждого супруга должна поступать на его счёт, а половина — на счёт второго супруга*.

Тогда за время брака супруги зарабатывают одинаковые приращения пенсий, даже если один из них не работает, а занимается только детьми и домашним хозяйством. Тем самым нет необходимости искусственно вводить в закон уменьшенный рабочий стаж для женщин.

Данное правило не применяется в двух случаях. Во-первых, если заключён брачный контракт (что допускается Семейным кодексом), в котором оговорены иные условия в отношении пенсий. И во-вторых, если брак не оформлен официально. Но это уже воля самих людей.

Всё это уравнивает женские пенсии с мужскими. Поэтому не требуется назначать какие-то специальные пенсии одному из супругов по случаю смерти второго супруга-кормильца.

Дети. Несовершеннолетние дети ещё не могли сами заработать себе пенсию. Поэтому в случае смерти каждого из родителей необходимо назначать детям пенсию до достижения ими совершеннолетия. Предлагается социально справедливый способ определения её размера.

Пусть умерший родитель имел n несовершеннолетних детей и успел заработать себе трудовую часть пенсии $p = \alpha P$ согласно формуле (11). Если он сам получал пенсию и тратил её на себя и детей, то на долю каждого пришлось бы $p/(n+1)$. Кроме того, каждый нетрудоспособный должен получать ещё социальный минимум s . Поэтому на долю каждого ребёнка придётся

$$s + \frac{p}{n+1} = s + \frac{\alpha P}{n+1}. \quad (16a)$$

В случае смерти обоих родителей выплачиваются трудовые доли обоих:

$$s + \alpha \frac{P_{\text{отца}} + P_{\text{матери}}}{n+1}; \quad (16б)$$

разумеется, социальный минимум прибавляется лишь один раз.

Когда один из детей достигает совершеннолетия, выплата ему пенсии прекращается. Но при этом уменьшается число несовершеннолетних детей n . Поэтому пенсии остальным детям соответственно увеличиваются.

Родители и другие взрослые, фактически пользовавшиеся финансовой или иной помощью умершего, сами имеют право на пенсию соответственно своему трудовому вкладу. В любом случае им гарантирован социальный минимум s . Только эту пенсию они и получают. Если

кормилец хочет обеспечить им дополнительную пенсию после своей смерти, он должен либо делать на их имя отчисления в Пенсионный фонд, либо застраховать свою жизнь в их пользу.

Заметим, что Пенсионный фонд с его огромными денежными потоками мог бы стать очень надёжным страховщиком жизни для таких ситуаций. Для этого целесообразно законодательно разрешить ему такую деятельность.

2.10. Инвалидность. Инвалидность должна давать право более раннего выхода на пенсию. Но для того, чтобы этот инструмент хорошо работал, надо ввести более детальную шкалу инвалидности — примерно шесть групп вместо нынешних трёх. Для единообразия будем нумеровать их в порядке ухудшения здоровья: нулевая группа будет соответствовать нормальному здоровью, а шестая — практически полной утрате трудоспособности.

Группа инвалидности устанавливается медицинской комиссией. Если причина инвалидности неустранима (например, ампутирована рука), инвалидность устанавливается пожизненно. Если болезнь в принципе излечима, то инвалидность устанавливается на определённый срок, по истечении которого надо повторно проходить комиссию. Предприятия не имеют права предоставлять инвалидам работу с тяжёлыми (выше определённой степени) условиями труда.

Следует различать общую инвалидность (вследствие общего заболевания или травмы не по вине предприятия) и профессиональную (по вине предприятия либо работодателя, либо ещё чьей-то). Сроки выхода на пенсию при этом одинаковые, но размер пенсии может сильно различаться.

Общая инвалидность наступает вследствие болезни или травмы, в которой виноват сам гражданин: простудился на улице, свалился на даче с дерева и сломал руку, нарушил правила безопасности на предприятии и т. п. В табл. 7 представлены, в зависимости от группы инвалидности: возраст, с которого этот гражданин получает право выйти на пенсию; надбавка на лечение к социальной части пенсии q_n ; категории тяжести работ, на которые предприятиям разрешено его принимать, если он продолжает трудиться. Трудовую часть его пенсии следует по-прежнему определять по формуле (11) через базовую пенсию (12). Но для расчётного времени жизни на пенсии T_n следует рассмотреть два варианта:

$$T_n = T_0, \quad (17a)$$

$$T_n = T_0 + 60 - T, \quad (17б)$$

где T — возраст фактического выхода на пенсию. Здесь T_0 — расчётное время для здоровых граждан (сейчас оно должно быть $T_0 = 10$, как отмечалось выше). Второй вариант соответствует ситуации, когда заболевший получает только ту пенсию, которую в среднем обеспечивает его собственный лицевой счёт. Это обоснованно, но может приводить к ничтожной трудовой части, если инвалидность наступила в молодом

возрасте. Первый вариант означает пенсию существенно больше, хотя всё равно заметно меньшую, чем у здорового человека. Но при нём инвалида существенно дотируют здоровые.

Выбор можно сделать, подсчитав по статистическим данным, во что обойдётся казне Пенсионного фонда каждый из вариантов. По предварительным оценкам, если поднять пенсионный возраст женщин до мужского и «отлучить от кормушки» отраслевых пенсионеров, то денег хватит даже на вариант (17а).

Возможно, лучше окажется такой вариант. Инвалид может выйти на пенсию немедленно после установления инвалидности, не дожидаясь указанного в табл. 7 возраста. Тогда размер трудовой части рассчитывается по формуле (17б) и будет малым. Если же он доживается своего пенсионного возраста, то трудовая часть рассчитывается согласно (17а) и будет существенно больше.

Таблица 7. Пенсии по инвалидности

Группа инвалидности n	Пенсионный возраст, годы	Надбавка на лечение g_n	Разрешена работа категории трудности
0	60	0,00	любой
1	55	0,1	4-й и ниже
2	50	0,2	3-й и ниже
3	45	0,3	2-й и ниже
4	40	0,4	1-й и ниже
5	35	0,5	0-й
6	любой	0,6	только на спец. местах

Профессиональная инвалидность наступает при болезни или травме по вине предприятия (ненадлежащая охрана труда), другого учреждения (падение на обледенелом городском тротуаре), аварии на общественном транспорте и т. п. Пенсионный возраст, надбавка на лечение и права работать далее здесь также определяются табл. 7. Но размер трудовой части пенсии рассчитывается иначе. Человек проработал J месяцев, а при отсутствии инвалидности мог бы проработать нормативный срок 40 лет=480 месяцев и существенно увеличить свой пенсионный счёт и базовую пенсию P . Поэтому с учётом (11) следует

$$p = \alpha P \cdot 480 / [\min(J, 480)]; \quad (18)$$

выбор минимума нужен на тот случай, если инвалидность получена в пожилом возрасте, когда человек проработал больше 40 лет. При расчёте P нормативный срок жизни берётся общепенсионный T_0 .

Если инвалидность наступила до общепенсионного возраста, то собственный пенсионный счёт гражданина не покрывает его пенсию (18).

Недостающую сумму не следует покрывать за счёт взносов здоровых пенсионеров. Её должен внести в Пенсионный фонд виновник — предприятие-работодатель, муниципальное управление, транспортное предприятие и т. п. Величину соответствующей суммы нетрудно рассчитать.

Гражданин, получивший профессиональную травму или профессиональное заболевание, должен получать бесплатное лечение и 100%-ную выплату по больничному листу до тех пор, пока его состояние здоровья не стабилизируется (эти расходы ложатся не на Пенсионный фонд). Если он выздоровел, то продолжает работать. Если работоспособность не полностью восстановилась, оформляется профессиональная инвалидность одной из шести групп.

Замечание. Если травма или заболевание явились следствием того, что руководство предприятия не обеспечило охрану труда, то гражданин должен иметь право на возмещение ущерба. Но это делается в судебном порядке, и ущерб возмещает предприятие, а не Пенсионный фонд.

2.11. Налог на бездетность. В России сложилась очень тревожная ситуация. Из-за резкого ухудшения условий жизни, связанного с экономическим шоком 1992–1993 гг. (300-кратная инфляция за 2 года!), возросла смертность в среднем возрасте и резко упала рождаемость. С 1993 г. по 2002 г. демографическая убыль стабильно составляла по 0,8–0,9 млн чел. в год. Первые годы она частично компенсировалась миграционным приростом за счёт приезда русских, выдавливаемых из ближнего зарубежья. Но сейчас этот источник иссяк — выдавлены все, кто был в состоянии сняться с места. Поэтому можно надёжно предсказать, что в ближайшие 5 или более лет (если не произойдёт резкой смены правительственного курса) население России будет ежегодно убывать на $0,85 \pm 0,10$ млн чел.

Фундамент любой пенсионной системы — молодое поколение. Необходимо поощрять рождение и воспитание детей. Для устойчивого роста населения желательно иметь 2–3 ребёнка в семье, но более 4 детей поощрять нецелесообразно, ибо нормальные семьи всё равно вряд ли заведут много детей, а поощрять семьи пьяниц и наркоманов бессмысленно — их дети часто рождаются неполноценными и становятся обузой для общества.

Нынешнее поощрение в форме крошечного пособия на ребёнка малоэффективно. Целесообразней вернуться к практике, существовавшей ещё в 1960-е годы — налог на бездетность. Только этот налог должен поступать не в госбюджет, а в Пенсионный фонд, то есть выражаться в форме увеличенного процента пенсионных отчислений. Это вполне оправданно: те, кто растят детей, несут дополнительные расходы; а ведь эти дети потом будут кормить всех стариков, в том числе бездетных.

Психологически снижение налога (отчисления), причём на всю оставшуюся жизнь, действует сильнее, чем пособие на детей (тем

более, что пособие платится ограниченное время). Вдобавок, это стимулирует обзаведение детьми в более раннем возрасте. Но размер снижения должен быть существенным. Ориентируясь на нынешние отчисления 28 % и нашу оценку 21 % для 60-летнего возраста выхода женщин на пенсию, можно предложить табл. 8. Заметим, что в старом законодательстве налог на бездетность брался со всех бездетных мужчин, но только с замужних женщин. Здесь же предлагается не делать никаких исключений, даже если человек не может иметь детей по медицинским причинам.

Таблица 8. Пенсионные отчисления

Количество детей	0	1	2	3 и более
Отчисления q , %	35	27	21	18

Пенсионный счёт. Согласно табл. 8, при малом количестве детей берутся пенсионные отчисления $q > 18\%$. Однако на личный пенсионный счёт гражданина при этом следует зачислять только $q = 18\%$, которые он получил бы при большом количестве детей. Остальная часть суммы идёт в общую часть Пенсионного фонда и формально расходуется на всех пенсионеров (фактически это компенсация за то, что данный гражданин не несёт расходы по воспитанию детей).

Если гражданин (гражданка) состоит в зарегистрированном браке, то отчисления на пенсионный счёт его супруги (супруга) происходит именно с этой суммы $q = 18\%$. Отчисления на бездетность на эти счета не попадают. Такая система разумно работает и в том случае, когда у каждого из супругов не одинаковое количество детей.

Альтернатива. Опишем и второй вариант, хотя он кажется менее действенным по психологическим причинам. При нём пенсионные отчисления для всех одинаковы, но за рождение и воспитание n детей полагается добавка к пенсии r_n . Она измеряется не в процентах от пенсии, а в долях от средней зарплаты (табл. 9). Тем самым, она больше стимулирует не высокооплачиваемых работников, а среднеоплачиваемых, составляющих основную часть населения. Если ребёнка вырастил один родитель, вся надбавка достаётся ему; если двое (супруг, платящий алименты, при этом считается), то надбавка делится поровну между ними. Усыновлённые дети учитываются наравне с собственными. Родитель, лишённый родительских прав, теряет и право на надбавку.

Таблица 9. Надбавка к пенсии на детей

Число детей n	1	2	3 и более
Надбавка r_n при одном родителе	0,20	0,34	0,40
Надбавка r_n каждому из двух родителей	0,10	0,17	0,20

2.12. Работа пенсионера. Если гражданин получил право на пенсию по общему возрасту или инвалидности, но продолжает работать на этой или других работах, то пенсия ему должна выплачиваться полностью. Это не подачка, он заработал её своими пенсионными отчислениями.

Поскольку он продолжает работать и выплачивать отчисления в Пенсионный фонд, его трудовая пенсия увеличивается. Но обращаться за очередным пересмотром размера пенсии он имеет право не раньше чем через 1 год.

Здесь есть одна деталь. Если пенсия назначена по старости, или по общей инвалидности с нормативной жизнью по формуле (17б), то она «самодостаточная» — в среднем её финансируют собственные страховые взносы данного пенсионера. Поэтому трудовая часть пересчитывается по тем же формулам, по которым была назначена.

При профессиональной инвалидности, а также при общей инвалидности с расчётом по формуле (17а), эту пенсию дотируют другие источники. Поэтому для её пересчёта необходимо разработать специальные формулы.

§ 3. Управление Пенсионным фондом

3.1. Распределение по доходам. Балансные оценки по средним величинам в § 2 были ориентировочными. Первое, что надо знать для аккуратных расчётов, — это функция распределения населения по доходам. Её нельзя брать «с потолка», а надо опираться на достаточно надёжные статистические данные.

На основании данных Госкомстата за I квартал 1994 г., приведённых в форме гистограммы в табл. 10 и на рис. 3, видно число человек (включая детей), чьи средние доходы на одного человека лежат в пределах одной денежной ступеньки. Таких ступенек равной ширины 15, а 16-я градация охватывает «хвост» распределения — всех граждан с более высокими доходами.

Способ получения этих данных авторам неизвестен, но это не перепись. По-видимому, это выборочное обследование, так что его точность не слишком высока. Если обследовалось 1500 семей (как это обычно делается) численностью по 3–4 человека, то общее число

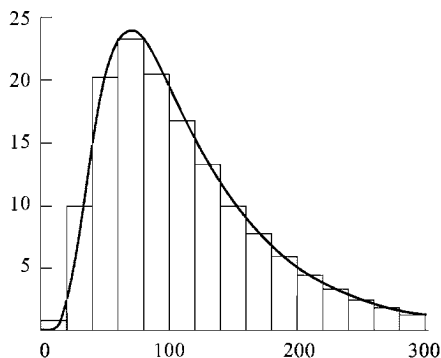


Рис. 3. Гистограмма — данные Госкомстата о распределении населения по доходам (табл. 10); кривая — распределение (19) с параметрами (25)

обследованных человек ≈ 5000 . Отсюда нетрудно рассчитать средне-квадратичную погрешность каждой ступеньки гистограммы. В самом деле, пусть в группу j_m доходов попадает n_j человек. Поскольку в выборке ≈ 5000 человек, а всё население составляет 148,4 млн человек, то из выборки в данную группу попадает $5000n_j/(148,4 \cdot 10^6)$ чел. Тогда абсолютная статистическая погрешность равна корню квадратному из последней дроби, а абсолютные погрешности численностей в табл. 10 составляют $\delta_j = \sqrt{148,4 \cdot 10^6 n_j / 5000} \sim \sqrt{30000 n_j}$ чел. Эти величины также приведены в табл. 10.

Таблица 10. Распределение населения России по доходам за I квартал 1994 г.

Диапазон доходов, тыс. руб/чел	Число человек (в млн.)	
	Данные Госкомстата	Аппроксимация (19) с параметрами (25)
0–20	$0,7 \pm 0,15$	0,558
20–40	$10,0 \pm 0,55$	8,550
40–60	$20,7 \pm 0,80$	19,834
60–80	$23,4 \pm 0,85$	23,698
80–100	$21,1 \pm 0,80$	21,710
100–120	$17,2 \pm 0,72$	17,675
120–140	$13,4 \pm 0,63$	13,628
140–160	$10,2 \pm 0,55$	10,257
160–180	$7,7 \pm 0,49$	7,654
180–200	$5,7 \pm 0,41$	5,710
200–220	$4,3 \pm 0,36$	4,277
220–240	$3,2 \pm 0,31$	3,226
240–260	$2,5 \pm 0,27$	2,452
260–280	$1,9 \pm 0,24$	1,879
280–300	$1,4 \pm 0,21$	1,453
больше 300	$5,0 \pm 0,39$	6,040
Всего	148,4	

Аппроксимация. Кривая распределения на рис. 3 начинается в нуле, и сначала растёт примерно как степенная функция. Затем рост замедляется и сменяется убыванием. Убывание явно не настолько

быстрое, как у экспоненты, и похоже на обратное степенное. Поэтому естественно аппроксимировать функцию распределения (такие аппроксимации экономисты называют моделями) отношением простейших многочленов:

$$f(\zeta) = \frac{a}{b} \frac{(\zeta/b)^n}{(1 + \zeta/b)^{n+m}}. \quad (19)$$

Здесь параметры n, m определяют форму кривой, a задаёт масштаб по оси ординат, b — по оси абсцисс. Очевидно должно быть $n \geq 0$; далее увидим, что всегда $m > 2$.

Назовём это распределение « n/m -бистепенное». Исследуем его.

Качественный вид кривой (19) для некоторых n, m показан на рис. 4. При $n = 0$ это монотонно убывающая кривая. При $n > 0$ она имеет максимум, величина и положение которого определяются формулами

$$\begin{aligned} \zeta_{\max} &= \frac{n}{m} b, \\ f(\zeta_{\max}) &= \frac{an^m m^m}{b(n+m)^{n+m}}. \end{aligned} \quad (20)$$

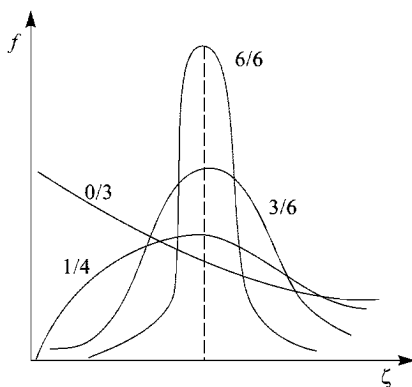


Рис. 4. n/m -бистепенное распределение при разных степенях

Чем больше n, m , тем выше и острее пик вблизи максимума. При $n, m \rightarrow \infty$ распределение превращается в δ -функцию. Это означает ситуацию, когда все граждане имеют одинаковый заработок $\zeta_{\max}$. Поэтому бистепенное распределение (19) позволяет исследовать разные экономические ситуации, включая предельные случаи. Величину $\zeta_{\max}$ следует называть наиболее вероятной зарплатой — её получает наибольшее количество человек.

Значения n, m могут в принципе быть нецелыми. Для простоты ограничимся целыми n . Тогда все выкладки удастся довести до конца в явном виде. Будем нормировать функцию распределения не на единицу, а на полное число человек N :

$$N = \int_0^{\infty} f(\zeta) d\zeta = a \int_0^{\infty} \frac{(\zeta/b)^n}{(1 + \zeta/b)^{n+m}} d\frac{\zeta}{b} = a \cdot A_{nm}, \quad (21a)$$

где

$$A_{nm} = \sum_{k=0}^n \frac{(-1)^k}{k+m-1} C_n^k = [(m-1)C_{n+m-1}^n]^{-1} = \frac{n!(m-2)!}{(n+m-1)!}; \quad (21б)$$

Таблица 11. Величины $1/A_{nm}$

$n \backslash m$	2	3	4	5	6	7	8
2	3	12	30	60	105	168	252
3	4	20	60	140	280	504	840
4	5	30	105	280	630	1260	2310
5	6	42	168	504	1260	2772	5544
6	7	56	252	840	2310	5544	12012
7	8	72	360	1320	3960	10296	24024
8	9	90	495	1980	6435	18018	45045
9	10	110	660	2860	10010	30030	80080

таким образом, $1/A_{nm}$ есть целое число (по-видимому, для нецелых n, m надо заменить все факториалы Γ -функциями, но доказать это не удалось). Значения $1/A_{nm}$ приведены в табл. 11. Отсюда

$$a = N/A_{nm}. \quad (21b)$$

Средняя зарплата $\zeta_{\text{ср}}$ равна полному заработку населения, делённому на его численность. Для распределения (19) это даёт

$$\zeta_{\text{ср}} = \frac{1}{N} \int_0^{\infty} \zeta f(\zeta) d\zeta = \frac{a}{N} b B_{nm} = \frac{b B_{nm}}{A_{nm}}, \quad (22a)$$

где

$$\begin{aligned} B_{nm} &= \sum_{k=0}^{n+1} \frac{(-1)^k}{k+m-2} C_{n+1}^k = \frac{m(n+1)!(m-3)!}{n(m+n-1)!} = \\ &= A_{n+1, m-1} = \frac{n+1}{m-1} A_{nm}; \end{aligned} \quad (22b)$$

величина $1/B_{nm}$ также оказывается целой (она извлекается из табл. 11), а замечание о переходе к Γ -функции аналогично. Подставляя (21b) и (22b) в (22a) и (20), получим

$$\frac{\zeta_{\text{ср}}}{b} = \frac{m(n+1)}{n(m-2)} > 1, \quad \frac{\zeta_{\text{ср}}}{\zeta_{\text{max}}} = \frac{m^2(n+1)}{n^2(m-2)}. \quad (23)$$

Когда $n, m \rightarrow \infty$ и распределение (19) переходит в δ -функцию, выполняется $\zeta_{\text{ср}} = b$; это очевидно: если все зарплаты одинаковы, то наиболее вероятная равна средней. В общем же случае из (22b) следует такое соотношение для наиболее вероятной и средней зарплат:

$$\zeta_{\text{ср}} > \zeta_{\text{max}} \quad \text{при} \quad m > \frac{n}{2(n+1)} \left(n + \sqrt{n^2 + 8n + 8} \right) \xrightarrow{n \rightarrow \infty} n+1; \quad (24)$$

границная линия этого неравенства показана на рис. 5.

Заметим также, что для сходимости интегралов в (21a) и (22a) необходимо $n > -1$ и $m > 3$. Но в экономике разумно усилить ограничение и требовать $n \geq 0$, чтобы функция распределения не обращалась в бесконечность. Даже значение $n = 0$ показывает наличие большого числа очень низкооплачиваемых работников (см. рис. 4), то есть соответствует нищему обществу со сверхбогатой верхушкой и её окружением.

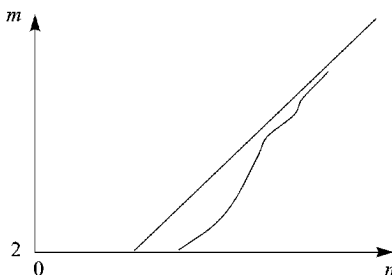


Рис. 5. Выше кривой $\zeta_{\text{ср}} > \zeta_{\text{max}}$, ниже — наоборот

Параметры для аппроксимации гистограммы рис. 3 и табл. 10 бистепенным распределением (19) подбирались вручную. Наилучшие результаты получились при следующих значениях параметров, точно передающих нормировку:

$$n = m = 6, \quad b = \zeta_{\text{max}} = 70, \quad \zeta_{\text{ср}} = 123, \quad a = 2310N. \quad (25)$$

Кривая аппроксимирующего распределения приведена на рис. 3. Видно, что она очень хорошо описывает гистограмму. Расчёт гистограммы по этой кривой приведён в последнем столбце табл. 10. Он превосходно согласуется с данными Госкомстата: отклонение не превышает статистических погрешностей во всех строчках таблицы за исключением последней строки $\zeta > 300$. Там рассчитанное значение выше экспериментального. Но этого следовало ожидать: это «хвост» распределения, описывающий высокооплачиваемых работников; а люди стараются скрывать высокие доходы.

Напомним, что параметры a, b — масштабные. Форма гистограммы определяется относительными высотами её столбцов, то есть 15 числами. Таким образом, очень точно описаны 15 экспериментальных точек с помощью всего 2 параметров n и m . Такое совпадение не может быть случайным. Значит, найден действительно удачный способ описания распределения по доходам, по крайней мере для среднеоплачиваемых работников с доходами $\zeta \leq 2,5\zeta_{\text{ср}}$. Из рис. 3 видно, что можно ожидать хорошего описания даже вдвое больших зарплат, хотя далёкий «хвост» распределения остаётся загадкой. Но области $\zeta < (3 \div 5)\zeta_{\text{ср}}$ достаточно для основных экономических расчётов: дальнейший «хвост» охватывает очень малую часть населения.

Использованные данные Госкомстата относились к 1994 г. Очевидно, величина b (денежный масштаб) будет заметно изменяться со временем хотя бы из-за инфляции. Однако вряд ли изменение сильно коснётся качественной формы кривой. Поэтому далее примем гипотезу, что $n = m = 6$ и в другие годы. Величина a согласно (21в) при этом зависит только от N , которое почти не изменяется со временем. Поэтому будут выполняться соотношения

$$n = m = 6, \quad \varsigma_{\max} = b, \quad \varsigma_{\text{ср}} = \frac{7}{4} \varsigma_{\max}, \quad a = 2310N; \quad (26)$$

заметно изменяться от года к году будет только b , легко определяемое по средней зарплате (данные о которой регулярно публикуются).

Олигохвост. Легальный доход бюджетных служащих, включая чиновников среднего ранга (с учётом не только оклада, но возможных надбавок и дотаций на лечение, зачастую составляющих второй оклад) обычно не превышают 5 средних зарплат. До 10 средних зарплат дотягивают доходы лишь высших чиновников. Однако у крупных дельцов и их окружения доходы могут в десятки и сотни раз превышать среднюю зарплату. Эта группа лиц образует далёкий «хвост» распределения $f(\zeta)$. Назовём его «олигохвост» (от слова олигарх) и попробуем найти его.

В 2000 г. средняя «пенсионная» зарплата составляла 16257 руб/год = 1355 руб/мес. Средняя начисленная (то есть истинная легальная) зарплата не публиковалась, но по словам председателя ПФ была в 1,5–2 раза выше, то есть составляла 2000–2700 руб/мес. Кроме того, налоговая инспекция сообщила прессе данные о числе человек с особо высокими доходами; они приведены в табл. 12. Число легально работающих не публиковалось, но ИНН уже получили 75 млн чел. Поэтому оцениваем его в 80–85 млн чел. Сводка всех данных представлена в табл. 12.

Цифры первого столбца табл. 12 неточны — они оценочные. Но остальные три столбца вполне надёжны — это данные налоговой инспекции. Они относятся к доходам в десятки и более раз выше средних, то есть к олигохвосту. Число человек с доходом больше ζ определяется интегралом от функции распределения

$$N(\zeta) = \int_{\zeta}^{\infty} f(\xi) d\xi. \quad (27)$$

Бистепенное распределение (19) на хвосте имеет обратную степенную асимптотику $f(\zeta) \sim \zeta^{-m}$. Подстановка её в (27) даёт $N(\zeta) \sim \zeta^{-(m-1)}$, это тоже обратный степенной закон. В последних трёх столбцах таблицы числа $N(\zeta)$ действительно зависят от ζ по степенному закону, но степень $m \neq 6$: по II и III столбцам с учётом статистической погрешности она даёт наклон $m - 1 = -2,12 \pm 0,03$ (по III и IV наклон будет

–2,5, но определять наклон по одному человеку крайне ненадёжно). Таким образом, фактический наклон указывает на значение $m \approx 3$.

Таблица 12. Данные за 2000 г.

Годовой доход ζ	Средний по ПФ 16257 р. Начислено 24–32 т.р.	Более 1 млн руб.	Более 10 млн руб.	Наиболь- ший 95 млн руб.
Число человек	80–85 млн чел.	более 35 тыс. чел	267 чел	1 чел.
Расчёт по (28а)	Параметры (30)	30,68 тыс. чел.	307 чел.	3 чел.

Для диапазона средних зарплат зависимость (19) с $m = 6$ весьма хороша. Можно предложить два следующих расширения 6/6-бистепенного распределения на сверхвысокие заработки:

$$f(\zeta) = \frac{a}{b} \frac{(\zeta/b)^n}{(1 + \zeta/b)^{n+m}} \cdot \begin{cases} [1 + (c\zeta/b)^k], & (28a) \\ [1 + (c\zeta/b)]^k, & (28б) \end{cases}$$

$$n = m = 6, \quad k = 3, \quad c \ll 1,$$

первое из них назовём жёстким, второе — мягким, по характеру влияния дополнительного члена $c\zeta/b$. При достаточно больших ζ оба варианта имеют асимптотику

$$f(\zeta) \approx \frac{a}{b} \left(\frac{cb}{\zeta} \right)^3 \quad \text{при } \zeta \gg \frac{b}{c}. \quad (28)$$

Добавочные множители в (28а)–(28б) слабо влияют на нормировку и другие соотношения, связанные со средними величинами. Поэтому можно принять $a \approx N_{\text{полн}}/A_{66} = 2310 \cdot N_{\text{полн}}$ и $\zeta_{\text{ср}} \approx 7b/4$ согласно (21в) и (22а). Учитывая эти соотношения и подставляя (28) в (22), получим:

$$c^3 \approx \frac{N(\zeta)}{1155 \cdot N_{\text{ср}}} \left(\frac{7\zeta}{4\zeta_{\text{ср}}} \right)^2. \quad (29)$$

Наибольшую неопределённость вносит сюда выбор $\zeta_{\text{ср}}$. По «пенсionной» зарплате и значениям $N(\zeta)$ из II и III столбцов получается $c \approx 0,16$. Если же взять оценку начисленной зарплаты, то получим $c \approx 0,108$. Обсудим это.

Значение $c \approx 0,16$ поднимает хвост функции распределения гораздо сильнее, чем значение $c \approx 0,11$. То же можно сказать о мягком варианте (28б) по сравнению с жёстким вариантом (28а). Но гистограмма на рис. 3 хорошо описывается 6/6-бистепенным распределением и не требует подъёма хвоста. Это заставляет отдать предпочтение жёсткому варианту (28а) с параметром $c \approx 0,11$.

Таким образом, олигохвост удаётся удовлетворительно описать распределением (28а), содержащим два дополнительных параметра: безразмерный коэффициент c и показатель степени. Для 2000 г. принимаем следующие значения параметров:

$$N_{\text{полн}} = 82,5 \text{ млн.чел.}, \quad b = 16 \text{ т.р./год}, \quad \zeta_{\text{ср}} = 28 \text{ т.р./год}, \quad c = 0,108. \quad (30)$$

Расчёт олигохвоста по ним приведён в табл. 12 и удовлетворительно описывает эксперимент (если учесть, что сверхвысокие доходы обычно скрывают, то расхождения естественны). Это распределение можно использовать для уточнения экономических расчётов.

Уточнение. Для улучшенного распределения (28а) можно получить точные выражения для нормировочного интеграла (21а) и средней зарплаты (22а).

$$N = a(A_{nm} + c^k A_{n+k,m-k}) \rightarrow \frac{a}{2310} (1 + 21c^3), \quad (31a)$$

$$\zeta_{\text{ср}} = ab(B_{nm} + c^k B_{n+k,m-k}) \rightarrow \frac{ab}{1320} (1 + 40c^3), \quad (31b)$$

$$\frac{\zeta_{\text{ср}}}{b} = \frac{A_{nm} + c^k A_{n+k,m-k}}{B_{nm} + c^k B_{n+k,m-k}} \rightarrow \frac{7}{4} \frac{1 + 40c^3}{1 + 21c^3}; \quad (31в)$$

стрелкой показаны значения для $n = m = 6$, $k = 3$. Аналогичные, но более громоздкие выражения можно получить для распределения (28б). Использование соотношений (31) даёт следующее уравнение для определения c :

$$\frac{c^3(1 + 21c^3)}{(1 + 40c^3)^2} = \frac{7N(\zeta)}{2640N_{\text{полн}}} \left(\frac{\zeta}{\zeta_{\text{ср}}} \right)^2. \quad (32)$$

Оно квадратное относительно c^3 , и его решение для оценочной начисленной зарплаты $\zeta_{\text{ср}} \approx 28$ тыс. руб./год даёт опять $c = 0,111$ (небольшое увеличение c не превышает 3%). Таким образом, данное уточнение оказалось небольшим.

По величине c видно, что олигохвост начинает доминировать при доходах, вдесятеро превышающих среднюю зарплату в стране (это доходы, превышающие оклад Президента России).

3.2. Децильный коэффициент. Так называют отношение суммарного дохода 10% самых богатых граждан и 10% самых бедных. Это то же самое, что отношение средних доходов одного человека в этих двух группах. Считается, что этот коэффициент хорошо передаёт меру имущественного неравенства граждан. Если его значение больше 10, в обществе возникает опасное социальное напряжение.

Пусть N — полная численность населения и, как и выше, $N(\zeta)$ — число людей с доходом больше ζ . Тогда определим границы высоких $\zeta_{\text{в}}$

и низких $\zeta_{\text{н}}$ доходов соотношениями

$$N(\zeta_{\text{в}}) \equiv \int_{\zeta_{\text{в}}}^{\infty} f(\zeta) d\zeta = 0,1 \cdot N, \quad N(\zeta_{\text{н}}) \equiv \int_{\zeta_{\text{н}}}^{\infty} f(\zeta) d\zeta = 0,9 \cdot N; \quad (33)$$

это означает, что доходы более $\zeta_{\text{в}}$ или менее $\zeta_{\text{н}}$ имеет по 10% населения. Децильный коэффициент равен отношению их доходов:

$$\varkappa = \int_{\zeta_{\text{в}}}^{\infty} \zeta f(\zeta) d\zeta / \int_0^{\zeta_{\text{н}}} \zeta f(\zeta) d\zeta. \quad (34)$$

Для произвольно заданной функции распределения надо сначала определить значения границ $\zeta_{\text{в}}$, $\zeta_{\text{н}}$ из уравнений (33) и затем подставить их в (34). Как правило, явно решить эту задачу не удаётся, и приходится использовать численные методы интегрирования и решения уравнений.

Были проведены такие расчёты для n/m -бистепенного распределения (19). Легко видеть, что децильный коэффициент не зависит от масштабных коэффициентов a , b , остаётся зависимость только от показателей степени n , m . Расчёты проводились как для целых, так и для нецелых степеней. Вычисленные значения $\varkappa_{nm}$ приведены в табл. 13. Обсудим эти результаты.

По своему определению $\varkappa_{nm} > 1$. При $n, m \rightarrow \infty$ распределение переходит в «монозарплатную» δ -функцию, для которой $\varkappa_{\infty\infty} = 1$. При уменьшении n , m коэффициент монотонно возрастает по каждой из степеней. При $n = 0$ он очень велик (≈ 100 и более) а при $m \rightarrow 3$ он обращается в бесконечность. Для реального 6/6-распределения, исследованного выше, его значение $\varkappa_{66} = 8,8$ лежит ниже социально опасной границы $\varkappa = 10$.

Замечание. Отметим, что для бистепенного распределения (19) при $n = 0$ нетрудно выполнить все интегрирование в (33)–(34) и при этом явно решить уравнения для определения границ и получить несложное выражение для децильного коэффициента при произвольных m . Приведём эти формулы:

$$\frac{\zeta_{\text{в},n}}{b} = N_{\text{в}}^{-1/(m-1)} - 1, \quad \zeta_{\text{ср}} = \frac{b}{m-2}, \quad \text{где } N_{\text{в}} = 0,1, \quad N_{\text{н}} = 0,9; \quad (35a)$$

$$\varkappa_{0m} = \frac{(m-1)N_{\text{в}}^{r(m-2)/(m-1)} - (m-2)N_{\text{в}}}{1 + (m-2)N_{\text{н}} - (m-1)N_{\text{н}}^{r(m-2)/(m-1)}} \quad (35b)$$

Они были использованы для проверки численных расчётов.

Расхождение. Последняя фраза кажется странной. Социальное напряжение в России все эти годы было и остаётся большим. Интересны оценки значений $\varkappa$ в разных странах, публикуемые экономистами и статистиками (годы не всюду указаны). Они приведены в табл. 14.

Расхождение этих цифр по современной России с рассчитанными здесь данными большое. Возникает вопрос — кто как считал? В данных

Таблица 13. Децильный коэффициент κ_{nt} для n/m -бистепенного распределения (19)

$\frac{m}{n}$	2,50	3,00	3,50	4,00	4,50	5,00	5,50	6,00	6,50	7,00	7,50	8,00	8,50	9,00	9,50	10,00
0,0	376,97	202,68	151,78	127,80	114,00	105,04	98,74	94,09	90,55	87,72	85,46	83,56	81,97	80,63	79,48	78,45
0,5	180,31	93,65	68,18	56,26	49,41	44,99	41,91	39,64	37,91	36,54	35,43	34,51	33,74	33,08	32,52	32,03
1,0	127,03	64,03	45,69	37,15	32,27	29,13	26,95	25,34	24,12	23,15	22,37	21,72	21,18	20,72	20,32	19,98
1,5	102,83	50,87	35,80	28,80	24,82	22,26	20,48	19,17	18,18	17,39	16,76	16,24	15,80	15,43	15,11	14,83
2,0	89,35	43,63	30,38	24,25	20,76	18,52	16,97	15,83	14,97	14,28	13,73	13,28	12,89	12,57	12,29	12,05
2,5	80,94	39,09	27,00	21,41	18,24	16,21	14,80	13,77	12,98	12,36	11,86	11,45	11,10	10,81	10,55	10,33
3,0	75,10	36,00	24,70	19,49	16,53	14,64	13,33	12,37	11,63	11,06	10,59	10,21	9,89	9,62	9,38	9,18
3,5	70,98	33,77	23,04	18,10	15,30	13,51	12,27	11,36	10,67	10,12	9,68	9,32	9,02	8,76	8,54	8,34
4,0	67,81	32,08	21,79	17,06	14,38	12,66	11,47	10,60	9,94	9,42	9,00	8,65	8,36	8,12	7,90	7,72
4,5	65,33	30,76	20,81	16,24	13,65	12,00	10,85	10,02	9,38	8,87	8,47	8,13	7,85	7,62	7,41	7,23
5,0	63,31	29,70	20,03	15,59	13,08	11,47	10,36	9,54	8,92	8,44	8,04	7,72	7,45	7,22	7,02	6,84
5,5	61,74	28,83	19,39	15,06	12,60	11,04	9,95	9,16	8,55	8,08	7,70	7,38	7,11	6,89	6,70	6,53
6,0	60,33	28,11	18,86	14,61	12,21	10,68	9,61	8,84	8,25	7,78	7,41	7,10	6,84	6,62	6,43	6,26
6,5	59,16	27,50	18,41	14,24	11,88	10,37	9,33	8,57	7,99	7,53	7,16	6,86	6,60	6,39	6,20	6,04
7,0	58,18	26,97	18,02	13,92	11,59	10,11	9,09	8,34	7,76	7,31	6,95	6,65	6,40	6,19	6,01	5,85
7,5	57,30	26,52	17,69	13,64	11,35	9,88	8,87	8,13	7,57	7,13	6,77	6,48	6,23	6,02	5,84	5,68
8,0	56,55	26,12	17,39	13,39	11,13	9,69	8,69	7,96	7,40	6,97	6,61	6,32	6,08	5,87	5,69	5,54
8,5	55,90	25,77	17,13	13,18	10,94	9,51	8,53	7,80	7,25	6,82	6,47	6,18	5,94	5,74	5,56	5,41
9,0	55,31	25,45	16,90	12,99	10,77	9,36	8,38	7,67	7,12	6,69	6,35	6,06	5,82	5,62	5,45	5,29
9,5	54,75	25,18	16,70	12,82	10,62	9,22	8,25	7,54	7,00	6,58	6,24	5,95	5,72	5,52	5,34	5,19
10,0	54,30	24,92	16,51	12,66	10,49	9,10	8,14	7,43	6,90	6,48	6,14	5,86	5,62	5,42	5,25	5,10

Таблица 14. Децильный коэффициент в разных странах

Страна	СССР	Соврем. Запад	Россия 2000 г.	Москва 2000 г.	Россия (год?)
κ	4,4	6–8	14	56	88

расчётах κ не учитывался олигохвост. Но выше было показано, что он включается слишком далеко, чтобы ощутимо повлиять на средние величины. Переход к распределению (28а) с параметрами (30) практически не изменяет κ_{66} , а само это распределение хорошо описывает официальную информацию о легальных доходах.

Возможно два объяснения такого расхождения. Экономисты не говорят о своей методике расчёта κ в России. Но если они используют косвенные оценки нелегальных доходов олигархов (о чём можно судить по вывозу капитала на Запад), это должно увеличить κ в несколько раз, ибо увеличиваются оценки доходов высокооплачиваемой группы.

Кроме того, авторы неявно считали безработных равномерно распределёнными по семьям, тем самым подразумевая доход на душу населения. Если же говорить только о трудоспособном населении (а его 80–90 млн чел.) и в низкооплачиваемую группу включить безработных (а их ≈ 6 млн чел. при оценках по методике МОТ), то в низкооплачиваемой 10%-й группе (8–9 млн чел.) остаётся только 2–3 млн работающих, и доход этой группы падает в 3–4 раза, соответственно возрастает κ .

Видно, что второй метод подсчёта неразумен, ибо безработные распределены по семьям случайно и в среднем равномерно. А в семье все доходы делятся на всех членов семьи. Поэтому приводимая здесь оценка κ для легальных доходов в современной России состоятельна.

Значения κ . На Западе олигархи тоже утаивают доходы, но в гораздо меньшей степени, чем в России (по оценкам западных налоговых служб, утаивается от четверти до трети доходов). Поэтому их значения κ не сильно занижены. Но какое реальное значение имеет этот показатель в экономике?

Для 6/6-бистепенного распределения решение уравнений (33) даёт следующие значения высокой и низкой зарплат (соответствующие части под кривой распределения на рис. 3 заштрихованы).

$$\zeta_b = 1,8\zeta_{cp} = 3,2\zeta_{max}, \quad \zeta_{ll} = 0,36\zeta_{cp} = 0,64\zeta_{max}, \quad \kappa = 8,8. \quad (36)$$

Из величины ζ_b видно, что за κ ответственна ближняя, а не дальняя часть хвоста, то есть высокие, а не сверхвысокие доходы. Это высокооплачиваемые служащие и средние предприниматели с доходами в 2–5 раз выше средних (то есть в 3–8 раз выше наиболее массовой зарплат ζ_{max}). Они пользуются обычно теми же (или почти теми же) магазинами и рынками, что и среднеоплачиваемое население. Но они в состоянии заплатить за товар, не торгуясь. Их количество велико, так что они покупают значительную часть товаров. Именно это побуждает торговцев поднимать цены, что ударяет по всему населению.

Олигархи и их окружение много богаче, но их число невелико. Много товаров им не нужно. Поэтому их доходы не приводят к подъёму цен (исключение — рынок недвижимости: дома и хорошие квартиры). Однако основным раздражающим народ фактором являются именно олигархи. Поскольку децильный коэффициент их не выявляет, то желательно придумать другой показатель.

3.3. Варианты трудовой пенсии. Трудовая часть пенсии гражданина p является некоторой функцией его базовой пенсии P — нормативного отчисления в Пенсионный фонд за всю жизнь. Три варианта $p(P)$ предложены в § 2. Все балансы Пенсионного фонда и оценки необходимого процента пенсионных отчислений q выполнялись выше со средними по стране значениями p и P . Так было бы, если зарплаты всех граждан одинаковы, то есть функция распределения по доходам есть δ -функция. Но каждый гражданин получает свою зарплату. Насколько это изменяет расчёты балансов?

В предложенном здесь методе расчёта прошлого трудового вклада используется отношение индивидуальной зарплаты к средней по стране. Тем самым результаты не зависят от масштабных множителей a, b функции распределения (19), которые могут сильно изменяться при инфляции. Остаётся только зависимость от показателей степени n, m , описывающих форму $f(\zeta)$. Эти параметры должны гораздо слабее зависеть от времени: их заметное изменение означало бы качественную экономическую перестройку. Поэтому будем далее считать их постоянными и пользоваться для прошлых лет одной и той же функцией распределения. Тогда трудовая часть пенсии, вычисленная по средней зарплате, равна $p(\zeta_{\text{ср}})$. А усреднение индивидуальных пенсий, найденных по индивидуальным зарплатам, даёт

$$p_{\text{ср}} = \int_0^{\infty} p(\zeta) f(\zeta) d\zeta, \quad \zeta_{\text{ср}} = \int_0^{\infty} \zeta f(\zeta) d\zeta. \quad (37)$$

Надо выяснить, насколько они отличаются, то есть каково отношение

$$l = \frac{p_{\text{ср}}}{p(\zeta_{\text{ср}})}. \quad (38)$$

Исследуем это для вариантов (11)–(13) пенсионного закона и бистепенной функции распределения (19).

Очевидно, если выбран линейный закон трудовой пенсии $p = \alpha \zeta$ (11), то для любой функции распределения $f(\zeta)$ выполняется $l = 1$. Таким образом, для этого варианта пенсионного закона баланс Пенсионного фонда можно точно рассчитывать по средней зарплате. Это очень упрощает управление Пенсионным фондом.

Для корневого (12) и дробного (13) пенсионных законов величина l зависит от параметра β этих законов, но не зависит от α . Численным интегрированием были проведены расчёты l для бистепенных распределений (19) с разными n, m и при том значении β , которое

было подобрано в табл. 3. Результаты расчётов представлены в табл. 15 и табл. 16. Обсудим их.

Поскольку на рис. 2 кривая (12) лежит ниже (11), а кривая (13) — ещё ниже, то $1 = l_{(11)} > l_{(12)} > l_{(13)} > 0$. Когда $n, m \rightarrow \infty$, величины $l_{(12)}$ и $l_{(13)}$ стремятся к единице. При уменьшении показателей степени n, m оба коэффициента монотонно убывают по каждому аргументу. Но из таблиц видно, что оба коэффициента практически равны 1 как при экспериментальных степенях $n = m = 6$, так и при других разумных их значениях. Только при нереально малых m отличие коэффициентов от 1 становится значительным.

Объяснить это нетрудно. Параметр нелинейности β в законах (12) и (13) показывает, при каком отношении ζ к средней зарплате сильно проявляется нелинейность закона. Здесь $\beta > 8$, а при таких значениях $\zeta/\zeta_{\text{ср}}$ хвост распределения уже достаточно мал. Эта часть хвоста вносит ничтожный вклад во все интегралы, и разница пенсионных законов слабо проявляется. Заодно видно, что основное влияние оказывает малая степень m , а степень n почти не влияет на результат. Брать же малое β нецелесообразно: для (13) это означает низкий «потолок» пенсии, и в любых случаях — сильную «уравниловку». Уже отмечалось, что это толкает к сокрытию доходов.

Но если $l < 1$, это означает перераспределение страховых взносов от богатых к бедным за счёт нелинейности пенсионного закона. Из того факта, что реально $l \approx 1$, следует важный вывод. При не уравнительных пенсионных законах (больших β) перераспределение ничтожно и практической роли нелинейность не играет. Она только усложняет расчёт трудовой пенсии. Поэтому целесообразно выбрать линейную пенсионную формулу (11).

Но этот вывод можно сделать только благодаря тому, что проведены расчёты с реальной функцией распределения доходов.

3.4. Управление. Данный проект содержит несколько параметров пенсионного закона, которые можно разбить на две группы. Первая группа определяет размер пенсии: размеры социальной пенсии s , надбавок на инвалидность g_n , надбавок за заслуги h_n , коэффициент α . Вторая группа определяет пенсионные отчисления: общее отчисление q и отчисление за бездетность r_n . Эти параметры надо подобрать и изменять при необходимости, чтобы поддерживать стабильное состояние казны Пенсионного фонда. Как это целесообразнее сделать?

Оптимизация по многим параметрам — непростая задача, и управление становится при этом сложным и непрозрачным. Лучше всего один раз тщательно подобрать все коэффициенты первой группы, используя все статистические данные Пенсионного фонда, налоговой службы и Госкомстата по зарплатам, демографии и составу пенсионеров. Затем эти коэффициенты надо закрепить в законе и не изменять их десятки лет.

Таблица 15. Коэффициент l (38) для пенсионного закона (12) с $\beta \approx 8$

$\frac{m}{n}$	2,500	3,000	3,500	4,000	4,500	5,000	5,500	6,000	6,500	7,000	7,500	8,000	8,500	9,000	9,500	10,000
0,0	0,780	0,859	0,888	0,903	0,912	0,918	0,922	0,925	0,927	0,929	0,931	0,932	0,933	0,934	0,935	0,935
0,5	0,795	0,875	0,905	0,920	0,929	0,934	0,939	0,942	0,944	0,946	0,947	0,948	0,949	0,950	0,951	0,952
1,0	0,802	0,884	0,914	0,929	0,938	0,943	0,948	0,951	0,953	0,955	0,956	0,957	0,958	0,959	0,960	0,961
1,5	0,807	0,890	0,920	0,935	0,943	0,949	0,953	0,956	0,959	0,960	0,962	0,963	0,964	0,965	0,966	0,966
2,0	0,811	0,894	0,924	0,939	0,947	0,953	0,957	0,960	0,962	0,964	0,966	0,967	0,968	0,969	0,970	0,970
2,5	0,813	0,896	0,927	0,942	0,950	0,956	0,960	0,963	0,965	0,967	0,969	0,970	0,971	0,972	0,972	0,973
3,0	0,815	0,899	0,929	0,944	0,953	0,958	0,962	0,965	0,967	0,969	0,971	0,972	0,973	0,974	0,975	0,975
3,5	0,817	0,900	0,931	0,946	0,954	0,960	0,964	0,967	0,969	0,971	0,972	0,974	0,975	0,975	0,976	0,977
4,0	0,818	0,902	0,932	0,947	0,956	0,961	0,965	0,968	0,971	0,972	0,974	0,975	0,976	0,977	0,978	0,978
4,5	0,819	0,903	0,933	0,948	0,957	0,963	0,967	0,970	0,972	0,974	0,975	0,976	0,977	0,978	0,979	0,979
5,0	0,820	0,904	0,934	0,949	0,958	0,964	0,968	0,970	0,973	0,975	0,976	0,977	0,978	0,979	0,980	0,980
5,5	0,821	0,905	0,935	0,950	0,959	0,964	0,968	0,971	0,974	0,975	0,977	0,978	0,979	0,980	0,981	0,981
6,0	0,821	0,905	0,936	0,951	0,959	0,965	0,969	0,972	0,974	0,976	0,977	0,979	0,980	0,980	0,981	0,982
6,5	0,822	0,906	0,936	0,951	0,960	0,966	0,970	0,973	0,975	0,977	0,978	0,979	0,980	0,981	0,982	0,982
7,0	0,822	0,906	0,937	0,952	0,961	0,966	0,970	0,973	0,975	0,977	0,979	0,980	0,981	0,982	0,982	0,983
7,5	0,823	0,907	0,937	0,952	0,961	0,967	0,971	0,974	0,976	0,978	0,979	0,980	0,981	0,982	0,983	0,983
8,0	0,823	0,907	0,938	0,953	0,961	0,967	0,971	0,974	0,976	0,978	0,980	0,981	0,982	0,983	0,983	0,984
8,5	0,823	0,908	0,938	0,953	0,962	0,968	0,972	0,974	0,977	0,978	0,980	0,981	0,982	0,983	0,984	0,984
9,0	0,824	0,908	0,938	0,953	0,962	0,968	0,972	0,975	0,977	0,979	0,980	0,981	0,982	0,983	0,984	0,985
9,5	0,824	0,908	0,939	0,954	0,963	0,968	0,972	0,975	0,977	0,979	0,981	0,982	0,983	0,984	0,984	0,985
10,0	0,824	0,909	0,939	0,954	0,963	0,969	0,973	0,975	0,978	0,979	0,981	0,982	0,983	0,984	0,985	0,985

Таблица 16. Коэффициент l (38) для пенсионного закона (13) с $\beta \approx 8$

$\frac{m}{n}$	2,500	3,000	3,500	4,000	4,500	5,000	5,500	6,000	6,500	7,000	7,500	8,000	8,500	9,000	9,500	10,000
0,0	0,740	0,829	0,864	0,882	0,893	0,900	0,905	0,909	0,912	0,915	0,916	0,918	0,919	0,920	0,921	0,922
0,5	0,757	0,849	0,884	0,903	0,914	0,921	0,926	0,930	0,933	0,935	0,937	0,938	0,940	0,941	0,942	0,942
1,0	0,766	0,859	0,895	0,914	0,925	0,932	0,937	0,941	0,943	0,946	0,948	0,949	0,950	0,951	0,952	0,953
1,5	0,772	0,866	0,902	0,921	0,932	0,939	0,944	0,948	0,950	0,953	0,954	0,956	0,957	0,958	0,959	0,960
2,0	0,776	0,870	0,907	0,926	0,936	0,944	0,949	0,952	0,955	0,957	0,959	0,961	0,962	0,963	0,964	0,965
2,5	0,779	0,874	0,911	0,929	0,940	0,947	0,952	0,956	0,959	0,961	0,963	0,964	0,965	0,966	0,967	0,968
3,0	0,781	0,877	0,913	0,932	0,943	0,950	0,955	0,958	0,961	0,963	0,965	0,967	0,968	0,969	0,970	0,971
3,5	0,783	0,879	0,915	0,934	0,945	0,952	0,957	0,961	0,963	0,965	0,967	0,969	0,970	0,971	0,972	0,973
4,0	0,785	0,880	0,917	0,936	0,947	0,954	0,959	0,962	0,965	0,967	0,969	0,970	0,972	0,973	0,974	0,974
4,5	0,786	0,882	0,918	0,937	0,948	0,955	0,960	0,964	0,966	0,969	0,970	0,972	0,973	0,974	0,975	0,976
5,0	0,787	0,883	0,920	0,938	0,949	0,956	0,961	0,965	0,968	0,970	0,971	0,973	0,974	0,975	0,976	0,977
5,5	0,788	0,884	0,921	0,939	0,950	0,957	0,962	0,966	0,968	0,971	0,972	0,974	0,975	0,976	0,977	0,978
6,0	0,788	0,885	0,922	0,940	0,951	0,958	0,963	0,967	0,969	0,972	0,973	0,975	0,976	0,977	0,978	0,979
6,5	0,789	0,885	0,922	0,941	0,952	0,959	0,964	0,967	0,970	0,972	0,974	0,975	0,977	0,978	0,979	0,979
7,0	0,789	0,886	0,923	0,942	0,952	0,959	0,964	0,968	0,971	0,973	0,975	0,976	0,977	0,978	0,979	0,980
7,5	0,790	0,887	0,924	0,942	0,953	0,960	0,965	0,969	0,971	0,973	0,975	0,977	0,978	0,979	0,980	0,980
8,0	0,790	0,887	0,924	0,943	0,954	0,961	0,965	0,969	0,972	0,974	0,976	0,977	0,978	0,979	0,980	0,981
8,5	0,791	0,887	0,925	0,943	0,954	0,961	0,966	0,970	0,972	0,974	0,976	0,978	0,979	0,980	0,981	0,981
9,0	0,791	0,888	0,925	0,944	0,954	0,961	0,966	0,970	0,973	0,975	0,977	0,978	0,979	0,980	0,981	0,982
9,5	0,792	0,888	0,925	0,944	0,955	0,962	0,967	0,970	0,973	0,975	0,977	0,978	0,980	0,981	0,981	0,982
10,0	0,792	0,889	0,926	0,944	0,955	0,962	0,967	0,971	0,973	0,976	0,977	0,979	0,980	0,981	0,982	0,983

Во второй группе следует также по статистическому материалу подобрать и закрепить отношения r_n/q . Тогда останется только один свободный параметр q . Целесообразно изменить его официальный статус — сделать его не пенсионным отчислением, а пенсионным налогом (тем самым, r_n тоже станут налогами на бездетность) и пересматривать ежегодно или чаще, если это окажется возможным технически и юридически.

Прогноз q делается по текущим балансам Пенсионного фонда — ежемесячным поступлениям и расходам с учётом государственного бюджета на следующий год. При этом предусматривается обязательный резерв на случай обычных колебаний экономической ситуации (но не на случай обвала экономики — здесь никакой резерв не спасёт).

Этот способ выгоден ещё и потому, что ожидаемые изменения q будут меньше, чем при других управляющих параметрах: увеличение q не только увеличивает поступления в Пенсионный фонд, но одновременно сокращает заработную плату (так как они берутся из фонда заработной платы). А выплаты Пенсионного фонда измеряются в этих уменьшенных заработных платах.

По предложенной системе размер пенсии каждому пенсионеру назначается не в рублях, а в долях средней заработной платы. Размер этой платы берётся за очередной квартал, причём по данным налоговой службы, а не Пенсионного фонда: все налоги поступают через налоговую службу, которая затем перечисляет пенсионный налог в казну Пенсионного фонда. Соответственно денежный размер пенсий пересматривается раз в квартал автоматически на основании Закона, без дополнительных постановлений Правительства.

Учёт накладных расходов Пенсионного фонда почти несуществен: по отчёту Пенсионного фонда за 1999 год они составили 3,1 млрд руб. при выплате пенсий на 207,4 млрд руб., то есть 1,5 %. Правда, это относится к центральному аппарату и коммуникациям. Учёт содержания местных отделов социального обеспечения и расходов по доставке пенсий повышает накладные расходы до 3,4 %.

Для более точных расчётов надо использовать распределение граждан по заработкам (что фактически сделано выше), а также сведения об инвалидности, наличии детей, заслугах и потере кормильца (которых у нас нет). Для завершения темы постараемся найти дополнительные статистические данные и максимально уточнить и обосновать цифры в предлагаемом проекте.

§ 4. Статистические данные

4.1. Достоверность. Среди попадавшихся нам опубликованных данных было много заметных противоречий. Например, приводились числа работающих 65 млн чел. и пенсионеров 38,4 млн чел. Но работают в возрасте 19–58 лет (женщины до 55 лет, мужчины до 60), то есть в среднем 39 лет. Тогда получается, что пенсионеры живут в среднем

$39 \times 38,4/65 \approx 23$ года, то есть до 81 года. Но медицинская статистика указывает на 68–70 лет.

Поэтому трудно выбрать достоверные данные для построения необходимых расчётов. К сожалению, даже комиссии, разрабатывающие Закон о пенсиях, пользуются сомнительными данными (например, продолжительность жизни пенсионера берут 17 лет, то есть сильно завышенную).

Много путаницы вносит то, что значительная часть пенсионеров продолжает работать. Кроме того, лишь $\approx 1/3$ фактических доходов граждан легальна и облагается пенсионным налогом (это высказывание бывшего председателя правления Пенсионного фонда).

Официальных публикаций в легко доступных изданиях очень мало. Основным источником данных авторам служили публикации в «Российской газете». Их делили на четыре категории: «а» — официальные публикации Правительства и ведомств; «б» — публикации со ссылкой на сведения Госкомстата, Минналогов и других ведомств; «в» — интервью с гос. чиновниками ранга министра; «г» — прочие приводимые журналистами сведения. Они зачастую не совпадают друг с другом, а иногда расхождения значительны. Предпочтения отдавались данным более высокой категории надёжности.

4.2. Группы населения. Здесь достаточно подробные и надёжные результаты содержатся в переписи 2002 г. [7].

- 145,2 млн. — полное население страны (правда, к моменту выхода этой книги оно сократится на $\approx 1,7$ млн.).

- 31 млн детей, школьников и студентов (примерно 1 млн студентов фактически дополнительно работает, хотя не всегда легально).

- 84,4 млн граждан трудоспособного возраста, уже не учащих. Из них примерно поровну мужчин моложе 60 лет и женщин моложе 55 лет.

- 29 млн граждан общепенсионного возраста (9,3 млн мужчин и 20,5 млн женщин).

- Таким образом, потенциальный трудовой ресурс страны не превышает 84 млн человек. Фактически он существенно меньше. В этой возрастной группе 3,0 млн инвалидов, лишь часть из которых работает. Приблизительно 3 млн пенсионеров-силовиков и 0,7 млн пенсионеров вредных профессий, $\approx 18,2$ млн граждан трудится только в подсобном хозяйстве, не уплачивая при этом налогов или отчислений. Кроме того, немало граждан находится на иждивении других членов семьи. В итоге по данным переписи 2002 г. только

- 62,2 млн человек официально называли себя работающими. Это явно заниженная цифра.

Число работающих. Цифры по числу работающих наименее надёжны. По словам председателя ПФ, за сентябрь 2001 г. собрано 37,4 млрд руб. Не уточнено, взяты они по зарплате августа или сентября. Поэтому примем среднюю цифру зарплаты 1558,5 руб $\pm 0,54\%$; то-

гда взносы в ПФ с одного человека 28 % от зарплаты равны 436,38 руб. Это даёт число работающих $N_p = 37,4$ млрд руб./436,38 руб. = 85,7 млн чел. $\pm 0,54$ %.

Но, по словам министра налогов и сборов, за январь-сентябрь 2001 г. в ПФ собрано 328 млрд руб. (в среднем 36,44 млрд в каждый месяц, что лишь на 2,6 % отличается от цифры за сентябрь). Однако средняя зарплата за январь-сентябрь или декабрь-август составляет 1622 р. $\pm 0,05$ %. Это даёт число работающих $N_p = 80,2$ млн чел., что отличается почти на 7 %. Судя по такому расхождению, даже министры не располагают достоверной информацией (умышленно искажать именно эту цифру вряд ли стали бы).

Журналисты и неправительственные экономисты называют самые разные цифры. Налоговая инспекция сообщила, что уже 76 млн физических лиц получили ИНН. Он нужен только работающим легально. Причём, по своему опыту мы видим, что пока не все легально работающие его получили (на глазок, получили $\approx 80 - 90$ %). Такая оценка даёт 85–95 млн легально работающих. Сравнивая эту оценку с данными двух министров, примем

$$N_p \approx 85-90 \text{ млн. чел.} = 87,5 \pm 2,5 \text{ млн.чел.} \quad (39)$$

Гастарбайтеры. Возможно, число работающих включает в себя не только граждан России, проживающих на её территории. Если это так, только легально работающих иностранцев (включая граждан СНГ) должно быть 14–18 млн человек. А ведь немало ещё нелегальных. Это тревожный признак, если благосостояние страны сильно зависит от иностранных рабочих. Данный вопрос возник неожиданно, и заслуживает отдельного изучения.

Но нельзя считать надёжными используемые здесь газетные публикации. Для достоверного анализа нужны хорошие статистические данные. В частности, по мнению авторов, число работающих пенсионеров в 2–2,5 раза больше официального и составляет 10–13 млн. При этом немалая их часть работает легально (то есть платя налоги), но без трудовой книжки, так что по статистике ПФ они числятся неработающими.

Это связано с тем, что до 01.01.2002 пенсионер, работающий с трудовой книжкой, получал не полную пенсию, а лишь минимальную. Новый пенсионный закон отменяет это ограничение. Более того, легальное продолжение работы приводит к увеличению пенсионного капитала для исчисления страховой и накопительной частей пенсии. Поэтому можно ожидать, что пенсионеры постепенно перестанут скрывать факт работы, и число работающих пенсионеров по данным ПФ возрастет и приблизится к фактическому.

4.3. Группы пенсионеров. Сведения о распределении пенсионеров по разным группам неполны, а данные Пенсионного фонда, перепи-

си — 2002 и других источников заметно расходятся. Авторы составили следующую сводку на октябрь 2002 года.

- 38,4 млн человек — полная численность пенсионеров по данным ПФ РФ. Это надёжная цифра. Отдельно численности пенсионеров мужчин и женщин не попадались. По косвенным данным можно оценить ≈ 15 млн мужчин и ≈ 23 млн женщин.

- 4,7 млн инвалидов всех возрастов согласно переписи. Однако по другим источникам полное число инвалидов составляет 11 млн., что сомнительно. Из них инвалидов I и II групп — 1,1 млн., что правдоподобно. Сколько из них инвалидов по общим и сколько по профессиональным заболеваниям — не сообщалось. По косвенным данным можно оценить эти числа как 1,7 млн и 3,0 млн.

- 0,8 млн или более — пенсии по потере кормильца; из них 0,5 млн. — детям, не менее 0,3 млн. — вдовам и родителям.

- 6,6 млн. — пенсионеры силовых ведомств, выходящие на пенсию существенно раньше общего срока. Эти данные получены косвенной оценкой результатов переписи.

- 2,3 млн. — трудовые пенсии с льготным сроком. Из них 2,0 млн. — вредные производства, 0,25 млн. — районы Севера.

- 24,0 млн. — обычные трудовые пенсии без льготных условий.

- 1,6–5,0 млн официальная и неофициальная оценки числа работающих пенсионеров. При этом совершенно не учитываются пенсионеры, подрабатывающие в розничной торговле и сфере бытовых услуг.

4.4. Демография. Ряд данных переписи октября 2002 г. уже опубликован [7]. На их основе скорректирован ряд данных Госкомстата за 1989–2002 гг. (от предыдущей переписи до данной).

Очень интересны данные по изменению численности населения (табл. 17 и рис. 6). Ранее публиковались только данные по полному изменению численности населения. Из поведения соответствующей кривой на рис. 6 невозможно было дать прогноз изменения численности на ближайшие годы. Сейчас же отдельно даны естественный прирост (разница рождаемости и смертности) и миграционный прирост (разница приезда в Россию и уезда). Это проясняет картину.

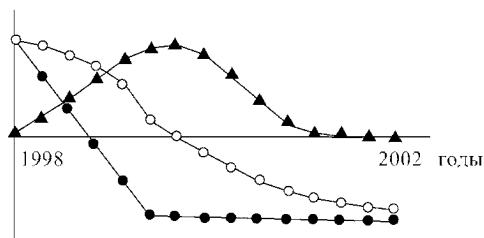


Рис. 6. Изменение численности населения (тыс. человек); точки — естественный прирост, треугольники — миграционный, кружки — полный

Таблица 17. Изменение численности населения (тыс. человек)

	Естественный прирост (убыль)	Миграционный прирост (убыль)	Общий прирост (убыль)
1989	580	83	663
1990	338	164	502
1991	110	52	162
1992	–207	176	–31
1993	–738	440	–298
1994	–870	811	–59
1995	–832	502	–330
1996	–818	346	–472
1997	–750	353	–398
1998	–697	285	–412
1999	–923	155	–768
2000	–954	214	–740
2001	–937	72	–865
2002	–935	78	–857

Видно, что с 1989 г. по настоящее время, несмотря на отъезд многих россиян на Запад, миграционный прирост был положителен. Ясно, что это связано с переездом граждан из республик (ближнего зарубежья). Особенно большим этот приток был в 1994–1997 гг., когда бывшие республики вытесняли русскоязычное население. Сейчас он упал почти до нуля, но не из-за смягчения давления, а потому, что уехали почти все, кто был в состоянии сняться с места. Поэтому можно уверенно сказать, что в ближайшие годы миграционный прирост не превысит 100 тыс. чел. в год.

Естественный прирост резко падал с 1989 по 1992 г. по мере ухудшения жизненного уровня (из рис. 6 видно, что инфляция в те годы была уже значительной). Шоковые «реформы» 1992 г. настолько ухудшили уровень жизни, что люди боялись заводить детей. И уже с 1993 года по настоящее время установился отрицательный естественный прирост на стабильном уровне -850 ± 100 тыс. чел./год. Он обусловлен в основном снижением рождаемости, а также некоторым увеличением смертности граждан среднего (работоспособного) возраста. По-видимому, последнее связано со стрессами, так как смертность среди пожилых граждан практически не возросла, несмотря на ухудшение условий жизни и здравоохранения. Поэтому на ближайшие годы

можно прогнозировать тот же отрицательный прирост (если кардинально не изменится правительственный курс, что мало вероятно).

По сумме этих факторов, в ближайшие 5 лет следует ожидать отрицательного полного прироста населения России в среднем по -800 тыс. чел./год.

4.5. Продолжительность жизни. Средняя продолжительность жизни граждан России на 1999 г. составляла у мужчин $T_m = 60,5$ лет и у женщин $T_{ж} = 72,7$ лет. В 2001 г. по данным демографической концепции она была в среднем $T = 65,9$ лет, причём у женщин она на 12 лет больше, чем у мужчин. Это означает, что средняя продолжительность жизни мужчин $T_m \approx 60$ лет (на деле несколько меньше, $T_m \approx 58,6$ лет), а женщин $T_{ж} \approx 72$ года (на деле несколько больше). Это показывает, что за 2 года средняя продолжительность жизни сократилась на 0,7 года — очень тревожный признак. Средняя длительность жизни на пенсии по прогнозу для нового Закона о пенсиях составляет $T_n = 19$ лет. Прогнозов отдельно для мужчин и женщин не публиковалось.

Цифры на первый взгляд не согласованы: возраст выхода на пенсию 60 лет у мужчин и 55 лет у женщин, то есть средний $\approx 57,5$ лет. Его разность со средней продолжительностью жизни составляет ≈ 10 лет, а не 19. Расхождение велико.

На самом деле надо не просто оперировать со средними цифрами, а проводить усреднения с учётом функции распределения смертности по возрастам. Например, по величине T_m может показаться, что мужчины вообще не доживают до пенсии. Но у мужчин в возрасте 40–50 лет довольно велика смертность от инфаркта. Те, кто пережили опасный возраст, имеют шанс прожить заметно дольше 60 лет.

Однако функции распределения не публиковались, так что нельзя проверить согласованность цифр по продолжительности жизни. Сама цифра $T_n = 19$ представляется умышленно завышенной, чтобы уменьшить размер страховой части пенсии.

4.6. Средняя зарплата. В табл. 18 приведены средние ежемесячные и поквартальные зарплаты z_c по стране для исчисления пенсий (то есть те, с которых в соответствующем месяце были отчислены 28 % в Пенсионный фонд РФ). Существуют ещё начисленная зарплата z_n и оценки полной зарплаты с учётом теневых выплат z_t . Они сильно отличаются. Так, на 1-й квартал 2001 г. председатель Пенсионного фонда приводил такие цифры:

$$z_c = 1523 \text{ р. (а)}, \quad z_n \approx 2500 \text{ р. (б)}, \quad z_t \geq 4500 \text{ р. (в)}. \quad (40)$$

Начисленная зарплата превышает среднюю «пенсионную» почти на 1000 р., т.е. более чем в 1,6 раза. Это означает, что со значительной доли легальной зарплаты не поступают положенные отчисления в Пенсионный фонд (ПФ). Отчасти это связано с задержкой выплат зарплаты на многих предприятиях. Но возникает щекотливый

вопрос. ПФ не сообщал о быстром росте задолженности по взносам. Следовательно, взносы всё же попадают в ПФ, хотя с опозданием. Тогда разница между средней и начисленной зарплатой не должна быть большой. Нет ли тут двойной бухгалтерии ПФ — собрать все взносы, пусть с задержкой, но засчитать для расчёта размера пенсий только поступления в срок? Подозрения о двойной бухгалтерии подкрепляются следующими соображениями. Во-первых, в официальном органе — «Российской газете» — до перехода на новую пенсионную систему (01.01.2002 г.) публиковались сведения только о «пенсионной» зарплате, а после — только о начисленной. Одновременно та и другая зарплаты не публиковались, что подозрительно. Во-вторых, средняя начисленная зарплата поразительно быстро выросла от ≈ 2500 руб. в 2001 году до ≈ 5000 в 2003 году, хотя ни о каком двукратном подъёме жизненного уровня никто не трубил.

В-третьих, задержки зарплаты в среднем по стране никогда не были большими, хотя на отдельных предприятиях они могли достигать полугода и более. Средняя задержка определяется легко: надо взять данные Госкомстата по суммарной задержке в стране и поделить их на среднюю зарплату и число работающих. Проведённые авторами такие расчёты показали, что даже сразу после дефолта августа 1998 г. средняя задержка не превышала одного месяца, а в 2003 году сократилась до 3-х дней.

Таблица 18. Средние зарплаты $z_c(a)$, руб/мес.

	2000 г.	2001 г.
Январь	1194	1136
Февраль	1324	1620
Март	1254	1811
Апрель	1244	1724
май	1252	1653
Июнь	1271	1635
Июль	1413	1896
Август	1411	1550
Сентябрь	1325	1567
Октябрь	1528	
Ноябрь	1457	
Декабрь	1584	
I квартал	1257	1522
II квартал	1256	1671
III квартал	1383	1671
IV квартал	1523	
Годовая	1355	

Чисто теневая зарплата равна разнице $z_t - z_n$ и весьма велика ≈ 2000 р/мес. Однако если её легализовать, то с этой суммы надо заплатить полный пакет отчислений 35,6 % ≈ 710 руб. ещё до удержания налогов. Значит, при полной легализации и своевременной выплате зарплат мы бы имели (на 1-й квартал 2001 г)

$$z'_c \approx 3800 \text{ р. (г)}, \quad (41)$$

что в 2,5 раза больше нынешнего значения (1). Соответственно возросли бы взносы в ПФ и подоходный налог. Легко видеть, что такие взносы обеспечили бы выплату пенсий выше прожиточного минимума. Таким образом, современный недостаточный уровень пенсий в первую очередь обусловлен преобладанием теневой экономики и соответственно плохой собираемостью налогов, и лишь во вторую очередь — плачевным состоянием экономики (хотя эта причина тоже весьма существенна).

Положение с поступлением взносов в ПФ может сейчас существенно улучшиться в связи с принятием нового Трудового кодекса № 197-ФЗ от 30.12.2001. В нём предусмотрены штрафные санкции к предпринимателям за задержку заработной платы. Уже в конце 2001 г. многие предприниматели поторопились выплатить задолженность. Это может привести к подъёму средней зарплаты z_c до размера начисленной z_n с соответственным увеличением взносов в ПФ. Тогда для двойной бухгалтерии не останется места. Однако теневую часть зарплаты гораздо труднее легализовать. Любые льготы, которые при этом государство обещает гражданину, будут существенно меньше налогов, которые он дополнительно заплатит. Единственный весомый аргумент здесь — страх наказания, достаточно суровые штрафные и даже уголовные санкции и активная работа налоговой полиции.

4.7. Нелегальные доходы. Разумеется, их точные размеры неизвестны. По словам бывшего председателя Пенсионного фонда М. Зурабова, в 2001 г. средняя «пенсионная» зарплата примерно 1500 руб./мес., начисленная зарплата — 2,5–3,0 тыс. руб./мес., а полный доход с учётом нелегального он оценил как 4,5 тыс. руб./мес. Вероятно, он включал сюда доходы мелких торговцев, строителей-«шабашников» и т. п., ускользающих от регистрации и уплаты налогов.

Но есть ещё олигархи. В начале 1996 г. Правительство РФ опубликовало неполные официальные данные за 1992–1995 годы. За эти 4 года инфляция составила 1800 раз, а номинальная заработная плата возросла в 300 раз, то есть покупательная способность населения уменьшилась в 6 раз. В то же время суммарная продукция промышленности и сельского хозяйства упала в 1,6–1,7 раза. Если раньше население в состоянии было купить всю произведённую продукцию, то теперь оно могло купить лишь 1/4 её часть! Что же стало с остальной произведённой продукцией? Ведь на складах она не залежалась.

Возможно лишь одно объяснение. Продукция была продана нелегально (в том числе на Запад), а вырученные за неё деньги тоже нелегально ушли на какие-то счета в зарубежных или российских банках. Но в эти нелегальные доходы входят и доходы частного предпринимательства, для которых вполне разумна оценка М. Зурабова (50% от легальных доходов). Поэтому окончательное распределение доходов на начало 1996 г. выглядит следующим образом.

Примерно 1/4 доходов легальна. Её получает основная масса населения (легальный суммарный доход олигархов ничтожен). Ещё 1/8 полного дохода — это нелегальный доход основной массы населения (мелкий бизнес к ней относится). Остальные 5/8 — это нелегальный доход олигархов, которым они делятся со своим окружением и коррумпированными чиновниками.

Оценим влияние нелегальных доходов на децильный коэффициент (в § 3 для легальных доходов найдено значение $\varkappa = 8,8$). Нелегальные доходы основной массы населения его слабо изменяют, ибо их распределение не должно заметно изменить $f(\zeta)$. Полный легальный доход «верхних» 10% населения составляет $\approx 1/3$ от легального дохода всего населения, то есть 1/8 полного дохода всей страны. Учёт нелегального дохода неолигархических групп увеличивает эту цифру в 1,5 раза, то есть даёт 3/16. Учёт же олигархов добавляет сюда 5/8, то есть увеличивает ещё в 4,33 раза. Поскольку на доход «нижних» 10% это не влияет, то во столько же раз увеличивается $\varkappa$.

Итак, с учётом нелегальных доходов можно оценить децильный коэффициент как $\varkappa \approx 38$. Это разумно согласуется с оценками экономистов и политологов, проскальзывающими в печати.

Заключение

В этой главе разработана модель пенсионной системы, превосходящей все существующие в мире системы. Перечислим её важнейшие аспекты:

— Система проста и универсальна — для всех граждан, включая президента страны, пенсия определяется только по суммарному отчислению в Пенсионный фонд за всю жизнь. Исчезают отраслевые пенсии, понятие трудового стажа и многое другое. При этом пенсии для чиновников будут достаточно высокими. Но назначаться они будут на общих основаниях. Это уменьшит негативное отношение широких слоёв населения к чиновникам и снизит социальную напряжённость.

— Размер пенсионных отчислений заносится на лицевой счёт работника не в рублях, а в долях средней зарплаты по стране за тот же месяц. В таких же единицах устанавливается размер пенсии. Это автоматически индексирует на инфляцию как прошлый трудовой вклад, так и пенсию, и обеспечивает стабильность системы в течение многих десятков лет, даже в условиях сильной инфляции и финансовых обвалов.

— Система имеет налогово-распределительный тип, что надёжно предохраняет её от финансовых мошенничеств и казнокрадства, что также способствует стабильности.

— Расчёты показывают, что даже при нынешнем состоянии экономики эта система может обеспечить всем категориям населения пенсии в 1,5–2 раза выше нынешних.

— Система обеспечивает социальную справедливость размеров пенсий. Она не уравнительна и поощряет легализацию доходов, укрепление семьи, рождение детей.

— Деятельность Пенсионного фонда становится гораздо более простой и прозрачной, а управление всей системой производится единственным параметром — процентом пенсионных отчислений. Это позволит уменьшить чиновничий аппарат, снизить накладные расходы и соответственно увеличить пенсии или снизить пенсионные отчисления.

Ряд идей данной модели можно использовать для других целей, например — построения системы медицинского страхования.

Список литературы

1. Закон Российской Федерации № 340–1 от 20.11.1990 «О государственных пенсиях в Российской Федерации».
2. Изменения и дополнения к Закону «О государственных пенсиях в РФ» с Закона РФ № 1030–1 от 19.04.1991 по Федеральный закон № 77-ФЗ от 05.05.1997 (см. напр. Пенсионное законодательство, ЗАО «Бизнес-школа Интел-Синтез». — Москва, 1998).
3. Федеральный закон «О порядке исчисления и увеличения государственных пенсий» № 113-ФЗ от 21.07.1997 (см. там же).
4. Федеральный закон «О государственном обеспечении в Российской Федерации» № 166-ФЗ от 15.12.2001.
5. Федеральный закон «Об обязательном пенсионном страховании в Российской Федерации» № 167-ФЗ от 15.12.2001.
6. Федеральный закон «О трудовых пенсиях в Российской Федерации» № 173-ФЗ от 17.12.2001.
7. Итог Всероссийской переписи населения 2002 года. Российская газета, 2004, № 65 (3442) от 31 марта.

Глава V **МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СЛОЖНЫХ ТРУДНОФОРМАЛИЗУЕМЫХ ПРОБЛЕМ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ РЕГИОНА**

Введение

В настоящее время в России идёт переходный процесс ломки административно-командной системы хозяйствования и становления рыночных отношений в экономике, а в политическом устройстве — замены унитаризма федерализмом. Несовершенство политического и экономического устройства как следствие данного переходного процесса развития сказывается на всех важнейших направлениях деятельности регионов России. В этих условиях важнейшей задачей становится стабилизация социально-экономической и общественно-политической ситуации и снижение социальной напряжённости.

О влиянии социальных и экономических факторов на состояние и развитие социальных явлений в обществе, таких как преступность, бедность, социальное неравенство и др., говорится во множестве научных и публицистических работ. Однако в подавляющем большинстве случаев имеет место только констатация фактов, не приводится строгих подтверждений причинно-следственных связей, основанных на статистических данных. Настоящее исследование посвящено методологии выявления основных зависимостей между социальным явлением и характеризующими его социально-экономическими показателями как на межрегиональном уровне, так и для отдельных регионов.

Любой социальный процесс определяется значительным числом факторов, включающим в себя состояние экономики, социальное положение населения (жизненный уровень, занятость и др.), демографическую обстановку, этно- и геополитические особенности регионов и др. В данном исследовании особое внимание уделяется вопросам оценки степени влияния факторов на состояние социального явления [1], [2], определяемого только числовыми показателями, при этом само социальное явление может и не быть определено численно.

Исследование состояния социального явления в отдельном регионе начинается с установления наборов характеризующих его признаков (показателей). В этом случае регион можно рассматривать как точку

многомерного пространства, размерность которого определяется числом признаков. Расстояния между точками пространства, исчисляемые по правилам аналитической геометрии, позволяют определить положение каждой точки относительно остальных точек и, следовательно, определить место этой точки во всей совокупности, что делает возможным их упорядочение и классификацию. В основе построения индикатора состояния и показателей уровня развития лежит таксономический подход, предложенный в [3].

Актуальной на сегодняшний день проблемой является краткосрочное прогнозирование развития социально-экономических явлений, когда ситуацию отражает не какой-либо отдельный показатель, а целые группы взаимосвязанных признаков, что затрудняет применение теории динамических рядов. Поэтому чтобы оценить и спрогнозировать развитие социально-экономического явления, необходимо уметь точно предсказывать значения этих параметров. Для решения этой задачи предлагается использовать модифицированный обучающийся генетический алгоритм LGAP (Learning Genetic Algorithm for Prognosis).

Другим актуальным вопросом межрегионального сравнительного анализа является проблема многокритериального выбора. Как правило, возникает некая задача и несколько вариантов её решения, из которых необходимо выбрать наиболее оптимальное. Аналогичная, по сути, задача появляется при выборе наиболее предпочтительных объектов. Сложность её решения заключается в том, что объекты (альтернативы), из которых приходится выбирать «оптимальный», имеют большое количество показателей. Решение задачи выбора очевидно, если среди объектов существует такой, который превосходит все остальные по всем показателям. Но на практике такая ситуация встречается крайне редко. Как правило, среди претендентов на оптимальность одни объекты лучше по одним параметрам, но хуже по другим, и точно сказать, какой из них лучше, сразу не представляется возможным. Предлагаемый алгоритм многокритериального оптимального по Парето выбора позволяет ответить на поставленные вопросы.

§ 1. Исследование динамики развития социального явления

1.1. Анализ исходной информации. При формировании системы показателей необходимо выполнение условия их полноты, репрезентативности и достоверности.

Показатели, включаемые в исследование, должны отвечать следующим требованиям.

Они должны быть некоррелированы и не обладать эффектом мультиколлинеарности.

Показатели должны быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количествен-

ного измерения, то ему нужно придать количественную определённую.

1.2. Эталон развития. Конструирование динамического показателя уровня развития социального явления предполагает формирование индивидуальной матрицы наблюдений X , элементами которой x_{is} являются значения признака s в момент времени i ($s = 1, 2, \dots, n$, $i = 1, 2, \dots, t$, n — число признаков, описывающих социальное явление) и определение координат x_{0s} эталона развития P_0 , число которых равно числу признаков. Для каждого s -го признака вводится эталонное значение x_{0s} .

Координатами эталона служат нормативные либо определяемые экспертно оптимальные значения признаков. В задачу настоящего исследования не входит обсуждение самих нормативных значений показателей — они принимаются за рациональные, хотя будут постоянно пересматриваться и совершенствоваться.

Координаты эталона $P_0(x_{01}, x_{02}, \dots, x_{0n})$ можно рассчитать также на основе матрицы наблюдений. Все признаки делятся на стимуляторы и дестимуляторы. Основанием разделения признаков на две группы служит характер влияния каждого из них на уровень развития изучаемого социального явления. Признаки, оказывающие положительное, стимулирующее влияние на уровень развития, называются стимуляторами, в отличие от признаков-дестимуляторов, которые оказывают тормозящее влияние. Тогда координаты эталона развития определяются следующим образом

$$x_{0s} = \max_i x_{is}, \quad \text{если } s \in I, \quad (1)$$

$$x_{0s} = \min_i x_{is}, \quad \text{если } s \notin I \quad (s = 1, \dots, n), \quad (2)$$

I — множество стимуляторов, x_{is} — значение признака s в период i .

Численный способ определения координат эталона развития по фактическим наблюдениям нарушает свойство его инвариантности и, как следствие, затрудняет или делает невозможным сравнительный анализ. Принятие в качестве координат эталона оптимальных значений признаков приводит к введению точки утопии в многомерном пространстве признаков, практически недостижимой в реальных условиях и времени.

Таким образом, в качестве эталона целесообразно принять фиксированные на некоторый период значения признаков, учитывая как нормативные, так и расчётные показатели, рассматривая их не только как цель, но и как условия развития.

1.3. Преобразование данных. Признаки, включённые в матрицу наблюдений, неоднородны, поскольку описывают различные особенности социального явления, кроме того, имеют различные единицы измерения.

Для исключения влияния размерности показателей на результаты исследований производится стандартизация матрицы наблюдений

и эталона развития. Значения приводятся к масштабу $(0, 1)$ либо $(-3, 3)$.

1.4. Определение значимости признаков. Стандартизация данных приводит к тому, что каждый признак оказывает в среднем одинаковое влияние на расстояние между изучаемыми точками-регионами. Смягчения этого нежелательного явления можно добиться введением коэффициентов иерархии λ_s (весов), разделяющих признаки по их важности.

Весовые коэффициенты можно задать экспертно, при этом они должны удовлетворять условию нормировки

$$\sum_{s=1}^n \lambda_s = 1,$$

либо рассчитать на основе матрицы наблюдений.

Предварительная процедура определения коэффициентов иерархии — получение матрицы расстояний C

$$C = [c_{is}], \quad i = 1, 2, \dots, t, \quad s = 1, 2, \dots, n,$$

элементы которой равны

$$c_{is} = \sqrt{\sum_{k=1}^t (z_{ki} - z_{ks})^2}. \quad (3)$$

Непосредственный расчёт коэффициентов иерархии λ_s начинается с определения критического расстояния, равного наибольшему расстоянию между признаками, которое определяется следующим образом

$$k = \max_s \min_i c_{is}. \quad (4)$$

Затем для каждого s -го признака находятся и суммируются все расстояния, не превышающие критического,

$$\omega_s = \sum_{c_{is} \leq k} c_{is}, \quad (5)$$

выбирается признак, для которого исчисленная сумма расстояний наибольшая

$$\omega_M = \max_s \omega_s, \quad (6)$$

рассчитываются коэффициенты иерархии

$$\lambda_s = \frac{\omega_s}{\omega_M}. \quad (7)$$

В данном способе определения коэффициентов иерархии фигурируют все расстояния, которые меньше критического. Вследствие этого выбранными оказываются расстояния между признаками, ближе всего лежащими друг к другу, и, следовательно, эти расстояния отражают наиболее сильные связи между признаками.

1.5. Динамический показатель уровня развития социального явления. Конструирование показателя начинается с расчёта отклонений до эталона развития

$$c_{i0} = \left[\sum_{s=1}^n (z_{is} - z_{0s})^2 \right]^{1/2}, \quad i = 1, 2, \dots, t,$$

где n — число признаков, характеризующих социальное явление, z_{is} — стандартизованные значения признака s в период i , z_{0s} — стандартизованные координаты эталона развития.

Находятся: среднее расстояние до эталона развития

$$\bar{c}_0 = \frac{1}{t} \sum_{i=1}^t c_{i0};$$

стандартное отклонение

$$s_0 = \left[\frac{1}{t-1} \sum_{i=1}^t (c_{i0} - \bar{c}_0)^2 \right]^{1/2};$$

максимально возможное отклонение от сводного эталона

$$c_0 = \bar{c}_0 + 2s_0.$$

Показатель уровня социально-экономического развития региона определяется как нормированное расстояние до эталона

$$d_i^* = \frac{c_{i0}}{c_0}. \quad (8)$$

Показатель уровня развития характеризуется тем, что является величиной положительной и лишь с вероятностью, близкой к нулю, может оказаться больше единицы. Интерпретация его следующая: данное социальное явление находится на тем более высоком уровне развития (в смысле приближения к эталону), чем ближе значение показателя уровня развития к нулю.

На практике удобнее использовать показатель развития в виде

$$d_i = 1 - \frac{c_{i0}}{c_0}. \quad (9)$$

1.6. Определение степени воздействия признаков на показатель уровня развития. Построенный показатель отражает изучаемое социальное явление, характеризуя все совокупные изменения, произошедшие в значениях рассматриваемых признаков. Его недостатком является то, что в нём не учитываются изменения, происходящие с отдельными признаками, влияние этих изменений на общую величину показателя. Не известны ни направления изменений отдельных составляющих, ни масштабы этих изменений.

По этим причинам необходимо произвести дальнейшую модификацию предлагаемого показателя, а именно, определить зависимость, характеризующую влияние изменений индивидуальных показателей

уровня развития каждого признака на сводный показатель уровня развития, относящийся ко всем признакам.

Обозначим $c_{i0,s}$ расстояние значения s -го признака в момент времени i до его индивидуального эталона P_s .

$$c_{i0,s} = abs(z_{is} - z_{0s}), \quad i = 1, 2, \dots, t, \quad s = 1, 2, \dots, n.$$

Тогда квадрат расстояния до эталона развития

$$c_{i0}^2 = \sum_{s=1}^n (z_{is} - z_{0s})^2 = \sum_{s=1}^n c_{i0,s}^2, \quad i = 1, 2, \dots, t. \quad (10)$$

Из формулы (3) следует

$$c_{i0} = c_0 d_i^*. \quad (11)$$

Обозначим

$$c_{0,s} = \bar{c}_{0,s} + 2 \cdot s_{0,s}$$

максимально возможное отклонение от эталона для s -го признака. Здесь

$$\bar{c}_{0,s} = \frac{1}{t} \sum_{i=1}^t c_{i0,s},$$

$$s_{0,s} = \left[\frac{1}{t-1} \sum_{i=1}^t (c_{i0,s} - \bar{c}_{0,s})^2 \right]^{1/2}.$$

Индивидуальный показатель уровня развития для s -го признака определяется формулой

$$d_i^* = \frac{c_{i0,s}}{c_{0,s}},$$

откуда следует

$$c_{i0,s} = c_{0,s} d_{i,s}^*. \quad (12)$$

После преобразований получим

$$c_0^2 (d_i^*)^2 = \sum_{s=1}^n c_{0,s}^2 (d_{i,s}^*)^2$$

или

$$(d_i^*)^2 = \frac{\sum_{s=1}^n c_{0,s}^2 (d_{i,s}^*)^2}{c_0^2}. \quad (13)$$

Обозначив нормирующие множители

$$w_s = \frac{c_{0,s}}{c_0},$$

получим зависимость сводного динамического показателя уровня развития от показателей уровня развития отдельных признаков

$$(d_i^*)^2 = \sum_{s=1}^n (w_s d_{i,s}^*)^2. \quad (14)$$

Зависимость (14) позволяет определить влияние показателей уровня развития отдельных признаков на общую величину сводного показателя уровня развития, достигнутыми всеми признаками вместе взятыми.

Обозначим

$$\Delta d_{i+1}^{*2} = (d_{i+1}^*)^2 - (d_i^*)^2,$$

тогда

$$\begin{aligned} \Delta d_{i+1}^{*2} &= \sum_{s=1}^n (w_s d_{i+1,s}^*)^2 - \sum_{s=1}^n (w_s d_{i,s}^*)^2 = \\ &= \sum_{s=1}^n w_s^2 \left((d_{i+1,s}^*)^2 - (d_{i,s}^*)^2 \right) = \sum_{s=1}^n w_s^2 \Delta d_{i+1,s}^{*2}, \quad (15) \end{aligned}$$

где

$$\Delta d_{i+1,s}^{*2} = (d_{i+1,s}^*)^2 - (d_{i,s}^*)^2.$$

Примем за базовый период $(i = 1, 2, \dots, t)$

$$w_s^2 = \text{const}.$$

Формулу (15) можно использовать для оценки динамики изменения сводного показателя уровня развития, вызванной различиями в динамике отдельных показателей.

О положительной ситуации в исследуемом социальном явлении можно говорить тогда, когда величина Δd_{i+1}^{*2} будет отрицательной, поскольку тогда последующий $(i + 1)$ -й период характеризуется в целом лучшими значениями показателей (более близкими к эталонным), что означает принятие показателем d_{i+1}^* меньшего значения по сравнению с его значением d_i^* в предыдущий период.

1.7. Оценочная шкала. Для качественной оценки текущего и прогнозируемого состояния социального явления необходимо идентифицировать показатель уровня развития на вербально-числовой шкале.

Конструирование шкалы начинается с определения числа градаций (интервалов значений). Для каждого показателя границы интервалов задаются делением области допустимых значений на интервалы. Вербально-числовая шкала с пятью градациями для показателей представлена табл. 1.

Таблица 1. Вербально-числовая шкала оценки значений признака

Показатель	Градации	Интервал значений
X_s	Очень низкий уровень значений	$a_{0s} - a_{1s}$
	Низкий уровень значений	$a_{1s} - a_{2s}$
	Средний уровень значений	$a_{2s} - a_{3s}$
	Высокий уровень значений	$a_{3s} - a_{4s}$
	Очень высокий уровень значений	$a_{4s} - a_{5s}$

У признаков-стимуляторов правая граница последнего интервала совпадает с соответствующей координатой эталона развития $a_{5s} = x_{0s}$, у признаков-дестимуляторов эталону развития соответствует левая граница первого интервала $a_{0s} = x_{0s}$.

Далее вводятся фиктивные регионы R_j , состояние которых описываются граничными значениями показателей (см. табл. 2.), и для них рассчитывается показатель уровня развития D_j .

Таблица 2. Матрица координат фиктивных объектов

Фиктивный объект	Значения показателей						Значение уровня развития
	X_1	X_2		X_s		X_n	
R_0	a_{01}	a_{02}	...	a_{0s}	...	a_{0n}	D_0
R_1	a_{11}	a_{12}	...	a_{1s}	...	a_{1n}	D_1
R_2	a_{21}	a_{22}	...	a_{2s}	...	a_{2n}	D_2
R_3	a_{31}	a_{32}	...	a_{3s}	...	a_{3n}	D_3
R_4	a_{41}	a_{42}	...	a_{4s}	...	a_{4n}	D_4
R_5	a_{51}	a_{52}	...	a_{5s}	...	a_{5n}	D_5

Таблица 3. Вербально-числовая шкала оценки уровня развития

Градации	Интервал значений
Очень низкий уровень развития	$D_0 - D_1$
Низкий уровень развития	$D_1 - D_2$
Средний уровень развития	$D_2 - D_3$
Высокий уровень развития	$D_3 - D_4$
Очень высокий уровень развития	$D_4 - D_5$

Табл. 3 представляет вербально-числовую шкалу для качественной оценки социально-экономического развития региона, в которой границами интервалов являются исчисленные фиктивные показатели уровня развития.

1.8. Пропорциональность развития. Если рассматривать эталон развития как вектор целей

$$\overline{P}_0(x_{01}, x_{02}, \dots, x_{0n}),$$

а совокупность фактических показателей для отдельного региона как вектор достигнутых результатов

$$\overline{s}_i(x_{i1}, x_{i2}, \dots, x_{in}),$$

тогда мерой степени приближения социально-экономического состояния региона в i -й момент времени к целевому (нормативному) состоя-

нию можно считать

$$\sigma_{\varphi i} = \frac{|\bar{s}_i|}{|\lambda|} \cos \varphi_i.$$

Расчёт величин $\sigma_{\phi j}$ для фиктивных регионов позволяет построить вербально-числовую шкалу для оценки степени пропорциональности достигнутого состояния.

Таблица 4. Вербально-числовая шкала оценки пропорциональности развития

Градации	Интервал значений
Очень низкая пропорциональность развития	$\sigma_{\phi 0} - \sigma_{\phi 1}$
Низкая пропорциональность развития	$\sigma_{\phi 1} - \sigma_{\phi 2}$
Средняя пропорциональность развития	$\sigma_{\phi 2} - \sigma_{\phi 3}$
Высокая пропорциональность развития	$\sigma_{\phi 3} - \sigma_{\phi 4}$
Очень высокая пропорциональность развития	$\sigma_{\phi 4} - \sigma_{\phi 5}$

1.9. Прогнозирование на основе динамического показателя уровня развития. Сконструированный показатель уровня развития представляет собой динамическую характеристику изменения состояния социального явления. Совокупность исчисленных значений d_i ($i = 1, 2, \dots, t$) можно рассматривать как временной ряд

$$d_i = y_i + \varepsilon_i,$$

где y_i — детерминированная неслучайная компонента динамического ряда, ε_i — стохастическая случайная компонента, и использовать для прогнозирования модели и методы анализа динамических рядов.

Исследования показали, что в условиях отсутствия форсмажорных обстоятельств наилучшей аппроксимирующей функцией является линейная

$$y_i = b_0 + b_1 i, \quad (16)$$

Качество модели прогнозирования (16) определяется на основе свойств ряда остатков, т.е. величины расхождений на участке аппроксимации между фактическим уровнями и их расчётными значениями.

Подчинение ряда остатков нормальному закону позволит построить доверительный интервал прогнозируемого значения показателя, т.е. определить допустимые отклонения прогноза от основной тенденции.

В ряде случаев возникает ситуация, когда не подтверждается подчинение ряда остатков нормальному закону либо совокупность значений показателя уровня развития d_i в строгом смысле не является динамическим рядом. В таких случаях прогнозное значение уровня развития получается подстановкой в формулу (9) спрогнозированных значений признаков.

§ 2. Прогнозирование значений показателей обучающимся генетическим алгоритмом прогнозирования LGAP

2.1. Проблема прогнозирования взаимосвязанных признаков.

Традиционно алгоритмы самообучения анализируют данные, в которых имеются описания входных воздействий и реакций изучаемой системы на эти воздействия. Как правило, в кибернетической литературе задачи установления зависимостей между характеристиками входа и выхода называются задачами анализа «чёрного ящика». Однако при прогнозировании развития социально-экономических явлений возникает более сложная ситуация, когда многие входные воздействия сами по себе остаются неизвестными и отражаются в наблюдениях лишь косвенно, через изменения выходных характеристик изучаемых объектов. Ситуация также усугубляется осложнениями, причина которых — наличие скрытых внутренних влияний характеристик одних явлений на характеристики других явлений.

В таких ситуациях задача оценки и краткосрочного прогнозирования развития социально-экономических явлений или объектов (всюду далее под объектом будем понимать не только объект, но и любое явление, характеризующее набором показателей, изменяющимися во времени) может быть решена лишь на базе двух гипотез: гипотезы повторяемости (в прошлом встречались такие же или аналогичные внешние воздействия) и гипотезы адекватности реакции (похожие воздействия вызывают похожие реакции), являющейся очевидной модификацией гипотезы локальной компактности.

Для решения данной задачи предлагается использовать обучающийся генетический алгоритм LGAP (Learning Genetic Algorithm for Prognosis), наиболее подробно описанный в работе [4], который извлекает закономерности (знания) из специфичного рода данных и использует эти знания для краткосрочного прогнозирования. Применение данного метода обусловлено тем, что большинство классических алгоритмов прогнозирования временных рядов не учитывают специфику и сложность задачи многофакторного прогнозирования развития социально-экономических объектов. В основном они рассчитаны на прогноз в одномерном случае и не учитывают взаимосвязей, которые существуют между различными показателями, а тем более между объектами и поэтому носят сугубо ограниченный, локальный характер и не позволяют давать достоверный прогноз. Те же методы, которые способны учитывать многомерный случай имеют слишком большое время работы и их работа практически невозможна без серьёзного предварительного анализа данных. Поэтому они не могут применяться в силу ограниченности специалистов математиков-аналитиков и вычислительных ресурсов. Рассматриваемый же алгоритм LPGA предназначен для исследования многомерных временных рядов, что позволяет

ему избавиться от многих недостатков присущих традиционным методам в качественном смысле, а применение такого новейшего аппарата, как генетические алгоритмы даёт ему возможность во много раз уменьшить временную сложность работы. В алгоритме LGAP выделяется четыре этапа работы:

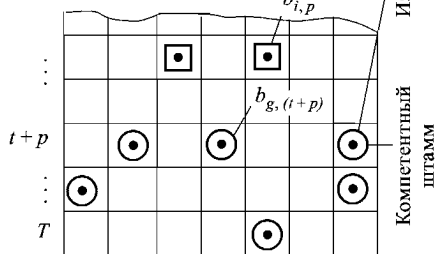
- 1) формирование базовых элементов (базовых штаммов);
- 2) отбор компетентных штаммов;
- 3) выработка частных вариантов прогноза;
- 4) получение окончательного прогноза.

Пояснение алгоритма будет происходить с помощью рис. 1.

2.2. Формирование базовых штаммов. Вместо трёхвходовой таблицы «объект-свойство-время», обычно рассматриваемой в задачах прогноза социально-экономического развития, алгоритм будет описан на примере двухвходовой таблицы «объект-время». Каждый элемент $b_{i,t}$ таблицы (протокола событий за T прошедших дней, месяцев и т.д.) отражает значение одной характеристики i -го объекта ($i = 1, 2, \dots, Y$) в t -й момент времени. Моменты времени ($t = 1, 2, \dots, T$) упорядочены в таблице по «возрасту»: самые свежие данные имеют индекс $t = 1$, данные за предшествующий день $t = 2$ и т.д. до дня с индексом $t = T$.

Обычно в алгоритмах данного уровня компетентность элементов оценивается по их похожести на некоторое базовое подмножество элементов, которые находятся в одной и той же строке или в одном и том же столбце. В данном случае целесообразно рассматривать более общий случай, когда в состав базового множества могут входить G любых элементов таблицы, взятых из протокола за τ последних дней. Такие множества называются базовыми штаммами мощности G . В протоколе за τ дней имеется N элементов ($N = Y \times \tau$). Мощность базовых штам-

Рис. 1. Принцип работы алгоритма



мов называются базовыми штаммами мощности G . В протоколе за τ дней имеется N элементов ($N = Y \times \tau$). Мощность базовых штам-

мов можно изменять в пределах от 1 до N . При каждой заданной мощности G штаммы могут иметь разную «архитектуру», вследствие чего количество вариантов базовых штаммов мощности G равно числу сочетаний из N по G , а общее количество разных базовых штаммов выражается равенством

$$M = \sum_{G=1}^N C_N^G.$$

Нетрудно видеть, что испытание всех базовых штаммов может превысить возможности даже самого мощного суперкомпьютера. Для сокращения перебора предлагается использовать комбинацию метода последовательного наращивания числа элементов с методами генетического программирования.

Метод последовательного наращивания выглядит так. Вначале оценивается ожидаемая ошибка прогноза элемента $b_{i,0}$ при использовании в качестве базовых всех N элементов по одному ($G = 1$). Выбирается подмножество из n наилучших вариантов ($n < N$). Затем просматриваются базовые штаммы из пар, образованных каждым из n элементов, отобранных на первом шаге, и всеми остальными ($N - 1$) элементами. Таких вариантов имеется $n \times (N - 1)$. Из них выбираются n лучших пар, к которым по очереди добавляем по одному все остальные ($N - 2$) элемента (здесь число вариантов около $n \times (N - 2)$). Тем же путём отбираются лучшие тройки и т.д. до заданного количества элементов G в базовом штамме. При постоянной величине n общее число вариантов имеет значение порядка $G \times n \times N$, что, например, при $G = 12$, $n = 12$, $n = 2$ и $N = 36$ меньше объёма полного перебора в миллион раз. Описанный процесс можно считать некоторой моделью, сочетающей мутации (в виде появления новых элементов у потомков) с естественным отбором лучших из них для использования в качестве «родительских» штаммов в следующем поколении [5].

Однако известно, что этот метод обеспечивает получение лишь локально-оптимального решения. Чтобы попытаться улучшить найденные решения, нужно применить на каждом шаге (т.е. при каждом новом значении G) процедуру «скрещивания». Для этого следует упорядочить элементы отобранных базовых штаммов по индексам времени. Если два элемента имеют одинаковый индекс t (т.е. находятся в одной строке таблицы), то упорядочивание этих элементов происходит по номеру объекта i (по номеру столбца). В соответствии с этим порядком присваивается каждому элементу в базовом штамме номер от единицы до G .

Из каждой пары родителей формируется определённое количество потомков путём замены k элементов первого родителя на k элементов второго родителя с одинаковыми порядковыми номерами этих элементов. При каждом значении k число новых потомков будет равно C_G^k . Можно ограничиться меньшим числом потомков, например двумя от каждой родительской пары, путём соединения первой половины одного

родителя со второй половиной второго (и наоборот). Но наиболее перспективный вариант, когда выборка k элементов происходит случайно, при этом начинается своеобразный стохастический направленный поиск нужного штамма. Качество таких гибридных штаммов оценивается наряду с качеством породивших их родителей и из этого набора отбирается n наилучших штаммов для последующего участия в мутациях и естественном отборе. Графически процесс скрещивания можно представить следующим образом:

$$\begin{array}{ccc} (a_1, a_2, a_3, a_4, \dots, a_i, \dots, a_G) & \text{скрещивание} & (a_1, b_2^1, b_3^2, a_4, \dots, b_i^k, \dots, a_G) \\ & \Rightarrow & \\ (b_1, b_2^1, b_3^2, b_4, \dots, b_i^k, \dots, b_G) & & (b_1, a_2, a_3, b_4, \dots, a_i, \dots, b_G) \end{array}$$

Таким образом, можно, формально сказать, что на данном этапе решается оптимизационная задача по максимизации функции качества (компетентности) базового штамма, а если быть точнее (забегая немного вперёд) порождённой им группы изоморфных штаммов, по которой строится прогноз. Аналитическое представление данной функции может быть различной и зависит от некой априорной информации; один из вариантов такой функции будет приведён ниже.

Итак, упоминаемая постоянно процедура оценки качества базовых штаммов включает в себя этап поиска компетентных штаммов и этап оценки ожидаемой ошибки предсказания с использованием этих штаммов. Перейдём к описанию процедуры отбора множества компетентных штаммов.

2.3. Отбор компетентных штаммов. Элементы базового штамма q , начинающегося с элементов t -й строки, в массиве исходных данных помечены их порядковыми номерами g в штамме: $(1, t)$, $(2, t)$, $\dots$, (g, t) , $\dots$, (G, t) . Набор адресов этих элементов описывает структуру (архитектуру) конкретного штамма, состоящего из G элементов. Если индекс t у всех элементов штамма увеличить на заданное число p , то мы получим штамм той же структуры, что и исходный, но только сдвинутый во времени на p шагов назад. Такой штамм называется изоморфным данному базовому штамму.

Далее следует выделить среди изоморфных штаммов группу, состоящую из k «потенциально компетентных» штаммов (ГПК). В неё будут включаться изоморфные штаммы с наибольшей похожестью на базовый штамм q . Меру похожести между штаммами можно оценивать разными способами.

Если ориентироваться на абсолютные значения характеристик их элементов, то можно использовать евклидово расстояние между штаммами:

$$R_{A,q,p} = \sqrt{\sum_{g=1}^G [b_{g,t} - b_{g,(t+p)}]^2}.$$

В ГПК отбирается k штаммов с наименьшими расстояниями $R_{A,q,p}$.

Если считать, что работа идёт с данными, измеренными в шкале отношений, то похожесть двух штаммов можно обнаружить через расстояние между соответствующими отношениями:

$$R_{O,q,p} = \sqrt{\sum_{g=1}^G \sum_{s=1}^G \left[\frac{b_{g,t}}{b_{s,t}} - \frac{b_{g,(t+p)}}{b_{s,(t+p)}} \right]^2}.$$

Инвариантами шкалы интервалов являются расстояния между отношениями интервалов:

$$R_{I,q,p} = \sqrt{\sum_{g=1}^G \sum_{s=1}^G \sum_{n=1}^G \sum_{m=1}^G \left\{ \frac{[b_{g,t} - b_{s,t}]}{[b_{n,t} - b_{m,t}]} - \frac{[b_{g,(t+p)} - b_{s,(t+p)}]}{[b_{n,(t+p)} - b_{m,(t+p)}]} \right\}^2},$$

Хорошими свойствами обладает мера Акайка [6], выражающая степень похожести P объектов через расстояние R между ними:

$$P_{q,p} = \frac{1}{1 + R_{q,p}}.$$

Для всех приведённых выше сильных шкал похожесть штаммов можно определять и по модулю коэффициента корреляции между значениями соответствующих элементов этих штаммов. Если в таблице представлены данные, измеренные в более слабых шкалах (порядка и наименований), то для измерения расстояния между штаммами можно воспользоваться любыми другими мерами.

На следующем этапе проверяется, является ли данная ГПК на самом деле компетентной. Чтобы обосновать процедуру такой проверки, достаточно вспомнить смысл гипотезы компактности H . Из неё следует, что сначала нужно удостовериться, что объекты множества (A, q) компактны в пространстве описывающих характеристик X . В данном случае имеется множество A из k штаммов, входящих в ГПК. Пространство X имеет размерность, равную G . Каждая g -я характеристика этого пространства есть номер g -го элемента базового штамма. Значение этого элемента $b_{g,t}$ и есть значение g -й. характеристики объекта (в данном случае базового штамма). Для изоморфного штамма, сдвинутого относительно базового на p моментов времени назад, значение g -й характеристики равно значению $b_{g,(t+p)}$ элемента, находящегося в клеточке с координатами $v_g(i, t+p)$. Таким образом, мы имеем дело с множеством A объектов в G -мерном пространстве X . Эти объекты были отобраны по критерию похожести на базовый штамм q . Если критерий похожести принять в качестве критерия компактности, то условие компактности $C_{A,q}^X$ для множества штаммов из ГПК выполняется автоматически.

Добавим к X ещё одну — целевую характеристику z , т.е. ещё один элемент таблицы. Пусть для базового штамма этим элементом будет тот, значение которого требуется предсказать, т.е. $b_{i,0}$. Соответственно к каждому штамму из ГПК добавим по одному известному

элементу $b_{i,p}$. Теперь появляется возможность проверить, выполняется ли условие компактности для целевой характеристики z у объектов множества A , входящих в ГПК. Условие компактности для значений одной характеристики может быть определено по-разному. Например, по величине дисперсии значений, по разности между $z_{\min}$ и $z_{\max}$ и т. д.

Если условие компактности для z на штаммах множества A не выполняется, то это означает, что по свойствам X они похожи друг на друга, а по свойству z не похожи. Следовательно, между свойствами X и z этих объектов закономерной связи нет, и нет оснований рассчитывать на успешное прогнозирование целевого свойства с опорой на штаммы из данной группы. Такая ГПК не включается в список компетентных групп. Она и породивший её базовый штамм из дальнейшего рассмотрения исключаются. Если же штаммы данной ГПК оказались компактными в пространстве свойств (X, z) , то они совместно со своим базовым штаммом образуют группу компетентных штаммов [7].

Затем вся описанная последовательность процедур повторяется для другого штамма той же мощности G , но другой архитектуры. Таким способом множество базовых штаммов порождает коллектив из W групп компетентных штаммов мощности G .

2.4. Выработка частных вариантов прогноза. Стратегии прогнозирования зависят от того, какие отношения между элементами разных штаммов считаются инвариантными при переходе от одного штамма к другому (т.е. постоянными для разных моментов времени). Если считать, что информативными являются абсолютные значения соответствующих элементов штаммов, то значение прогнозируемого элемента $b_{i,0}$ можно найти по абсолютным значениям k соответствующих элементов (предикторов), связанных с компетентными штаммами данной группы, т.е. с элементами $b_{i,p}$. Следует напомнить, что количество разных компетентных базовых штаммов мощности G равно W , так что в прогнозировании элемента $b_{i,0}$ может принять участие $W \times k$ предикторов (W групп по k предикторов).

Если предполагается линейная зависимость между соответствующими элементами разных штаммов, то можно воспользоваться прогнозированием с помощью линейной регрессии. Строится линейная регрессия между G элементами базового и p -го компетентного штаммов. Подстановка элемента $b_{i,p}$ в уравнение этой регрессии позволяет получить величину $b_{p,i,0}$ в качестве p -го варианта прогноза для элемента $b_{i,0}$. Аналогично получают и все другие k вариантов прогноза для штаммов этой группы: $b_{1,i,0}, b_{2,i,0}, \dots, b_{p,i,0}, \dots, b_{k,i,0}$. Здесь, как и в предыдущем случае, общее число предикторов (и вырабатываемых ими частных прогнозов) равно $W \times k$.

Кроме этого, если предположить, что отношения между элементом $b_{i,p}$ и всеми g -ми элементами данного штамма $b_{g,(t+p)}$ такие же, как и отношения между элементом $b_{i,0}$ базового штамма и всеми его элементами $b_{g,t}$. Тогда можно получить G вариантов значения прогнози-

руемого элемента с опорой на каждый g -й элемент p -го компетентного штамма:

$$\begin{aligned} b_{1,i,0} &= (b_{1,t} \times b_{i,p})/b_{1,p}, \\ b_{2,i,0} &= (b_{2,t} \times b_{i,p})/b_{2,p}, \\ &\dots\dots\dots \\ b_{g,i,0} &= (b_{g,t} \times b_{i,p})/b_{g,p}, \\ &\dots\dots\dots \\ b_{G,i,0} &= (b_{G,t} \times b_{i,p})/b_{G,p}. \end{aligned}$$

В этом случае роль предиктора играет каждый элемент штамма в отдельности, так что каждый p -й штамм представляет собой набор из G предикторов, а все k штаммов из компетентной группы образуют коллектив из k групп по G предикторов. Общее число предикторов (частных прогнозов) для этого случая будет равным $W \times k \times G$ (W коллективов по k групп из G предикторов).

Возможны и другие способы получения вариантов прогноза для данных, измеренных как в количественных, так и в любых других шкалах. Эти способы в основном зависят от располагаемой априорной информации о характере исходных данных.

2.5. Получение окончательного прогноза. Частные прогнозы получены с помощью предикторов, организованных в иерархическую структуру: отдельные предикторы объединяются в группы, которые объединяются в коллективы и т. д. Фактически это коллективно-групповой метод прогнозирования (КГМ).

При свёртке частных решений в общее имеет смысл обратить внимание на тот факт, что компетентность всех предикторов, породивших эти решения, превышает некоторый порог, но может быть неодинаковой. В результате разные группы предикторов вырабатывают разные наборы частных решений. Можно рассчитывать на то, что более компетентные группы будут выдавать более точные прогнозы,

Но компетентность групп проверялась по критерию компактности на множестве известных элементов A и отражала ситуацию в прошлом. Так делается обычно при ретроспективном выборе стратегий прогнозирования или распознавания: лучшей считается та стратегия, применение которой приводило к наименьшей сумме ошибок прогноза известных элементов. Однако экспериментально нередко обнаруживается, что в ряде случаев лучшие результаты дают некоторые другие стратегии, а не та, которая была лучшей «в среднем». По-видимому, они более точно соответствуют тем закономерностям процесса, которые проявляют себя в данный конкретный момент.

Соответственно возникает вопрос: какую же стратегию использовать для реального прогнозирования неизвестных величин? Где гарантия, что стратегия, чаще других в прошлом приводившая к успеху, будет лучшей и в этом конкретном случае?

Гипотеза компактности позволяет ответить на этот вопрос. Из неё следует, что истинная компетентность группы предикторов должна

проявляться не только в компактности объектов (A, q) в пространстве X и компактности объектов A в пространстве (X, z) , но и в компактности вариантов прогноза характеристики z объекта q . Теперь, когда эти варианты найдены, можно оценить компактность прогнозов, полученных от группы предикторов f , например, через дисперсию их значений D_f . Как показывают эксперименты, дисперсионный критерий оказывается в большинстве случаев очень информативным: корреляция между D_f и ошибкой прогнозирования достигает величины $+0,7$. Опираясь на это, в качестве меры компетентности группы f целесообразно принимать величину $L_f = 1/(1 + D_f)$.

Введение описанного выше дисперсионного критерия позволяет сформулировать следующий принцип выбора стратегий (в нашем случае — групповых предикторов): следует применять все имеющиеся в распоряжении стратегии и отдавать предпочтение той из них, которая даёт варианты прогноза, обладающие наименьшей дисперсией.

Эта эмпирическая гипотеза отражает закономерность более высокого уровня (метазакономерность) по сравнению с традиционно используемой гипотезой о том, что хорошие результаты в прошлом обеспечивают хорошие результаты и в будущем.

С учётом сказанного выше процедура получения обобщённого решения состоит в следующем. Вначале на базе частных решений для каждой группы f вырабатывается групповое решение B_f . При прогнозировании в качестве группового решения может быть использовано среднееарифметическое значение частных прогнозов или их медиана. Вычисляется также характеристика компетентности группы L_f . Эти характеристики находятся для всех k групп предикторов из коллектива групп, участвовавших в прогнозировании.

Обобщённое решение (B) на следующем иерархическом уровне (на уровне коллектива) может быть получено с использованием параметрического семейства функций взвешенного усреднения групповых решений:

$$B = \sum_{f=1}^k \frac{(B_f \times L_f^\alpha)}{\sum L_f^\alpha}.$$

Здесь величина показателя степени α отражает стратегию учёта влияния компетентности. Если $\alpha = 0$, то решения всех групп учитываются с равными весами. С ростом α растёт влияние более компетентных групп. При очень больших значениях α в усреднении участвуют одна или несколько самых компетентных групп.

Если в прогнозировании участвовал один коллектив групповых предикторов, то величина коллективного решения (B) принимается в качестве окончательного значения прогнозируемой величины $b_{i,0}$. Если определить дисперсию групповых прогнозов, то можно получить представление о компетентности коллектива в целом и о величине ожидаемой ошибки данного прогноза.

Метод допускает использование не одного коллектива, а множества из W коллективов. Новые коллективы могут порождаться базовыми штаммами, мощность которых отличается от G . В этом случае по формуле, аналогичной вышеприведённой, делается взвешенное усреднение коллективных решений.

На рис. 2 представлена схема многоуровневого алгоритма, коллективно-группового прогнозирования. На первом уровне используется базовый штамм мощности G_1 и архитектуры F_1 . Находятся компетентные штаммы $C_1, C_1, \dots, C_k$. Они образуют первую группу предикторов. С их использованием получаются частные прогнозы $b_1, b_2, \dots, b_k$, вычисляются их дисперсия D_1 и групповой прогноз b'_1 .

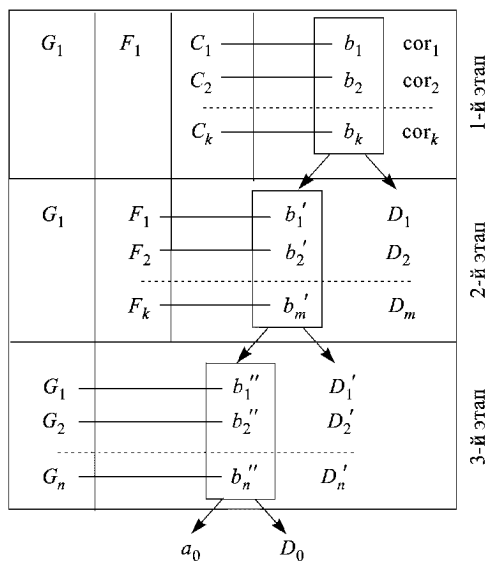


Рис. 2. Формирование прогнозируемого значения

На втором уровне эти процедуры повторяются для штаммов той же мощности G_1 , но для других архитектур: $F_1, F_2, \dots, F_m$. В результате получаются групповые решения $b'_1, b'_2, \dots, b'_m$ и по ним вычисляется коллективное решение a'_1 , и дисперсия D'_1 .

На третьем уровне изменяется мощность базовых штаммов, и для каждой мощности ($G_2, G_3, \dots, G_n$) повторяются все предыдущие процедуры. В итоге получается множество из n коллективных решений: $b''_1, b''_2, \dots, b''_n$ со своими значениями групповых дисперсий $D'_1, D'_2, \dots, D'_n$. По ним находится окончательный прогноз b_0 и дисперсия коллективных прогнозов D_0 , по которой можно судить об ожидаемой ошибке полученного прогноза.

2.6. Критерии оценки точности прогноза. Теперь, когда идея алгоритма полностью описана, нужно ответить на сложный вопрос, который, как правило, обычно возникает в процессе анализа результатов прогнозирования: как оценивать то, что мы называем точностью прогноза?

Часто берётся абсолютное отклонение прогноза $b'_{i,0}$ от истинного значения $b_{i,0}$, делённое на истинное значение:

$$d' = \frac{|b'_{i,0} - b_{i,0}|}{b_{i,0}}.$$

Такая относительная величина мало чувствительна к ошибкам прогноза больших значений и чрезмерно чувствительна к ошибкам прогноза величин, близких к нулю. Кроме того, разница dd_i , между минимальным и максимальным значениями может быть различной у разных наблюдаемых характеристик и одинаковая относительная ошибка d' будет приемлемой для принятия решений в одних случаях и не приемлемой в других. Тем не менее данная относительная ошибка, при всех её недостатках, остаётся основным средством оценивания точности прогноза.

Кроме этого, судить о точности прогноза i -й характеристики можно судить по величине ошибки, нормированной по разнице dd_i :

$$d = \frac{|b'_{i,0} - b_{i,0}|}{dd_{i,0}}.$$

Такая мера обладает одинаковой чувствительностью к ошибкам прогноза для разных значений прогнозируемой характеристики. Её чувствительность к ошибкам тем выше, чем в меньших пределах колеблется прогнозируемая характеристика, что представляется вполне логичным.

Иногда важно знать не абсолютную величину характеристики $b_{i,0}$ в будущем, а лишь то, будет ли она больше или меньше значения $b_{i,t}$ в данный момент времени. В таких случаях применима мера точности прогноза, учитывающая лишь совпадения знаков:

$$d^* = \begin{cases} 0, & \text{если } ((b_{i,0} > b_{i,t}) \text{ и } (b_{i,0}^! > b_{i,t}^!)) \\ & \text{или } ((b_{i,0} < b_{i,t}) \text{ и } (b_{i,0}^! < b_{i,t}^!)); \\ 0,5, & \text{если } ((b_{i,0} = b_{i,t}) \text{ и } (b_{i,0}^! \neq b_{i,t}^!)); \\ 1, & \text{в других случаях.} \end{cases}$$

2.7. Результаты работы алгоритма. Для оценки качества работы алгоритма был проведён ряд экспериментов. Производился прогноз значений первичных показателей [8], [9], [10], характеризующих регионы Уральского федерального округа с позиции такого социального явления как незаконный оборот наркотиков. Результирующие данные эксперимента приведены в табл. 5.

Таким образом, средняя величина ошибки составила 8,32 %. Данную ошибку можно объяснить в основном двумя причинами. Во-пер-

Таблица 5. Величина ошибки прогноза для различных показателей

Показатель	Средняя (по регионам) величина ошибки, %
1. Продажа водки и ликёро-водочных изделий	7,09
2. Число разводов на 1000 человек населения	4,16
3. Население моложе трудоспособного возраста	8,73
4. Численность безработных	11,82
5. Численность экономически активного населения	11,59
6. Заявленная потребность в работниках	13,19
7. Потребительские расходы на душу населения	5,68
8. Денежные доходы на душу населения	7,29
9. Стоимость минимального набора продуктов питания	11,28
10. Беженцы и вынужденные переселенцы	4,73
11. Миграционный прирост населения	5,92

вых, неполным объёмом показателей, среди которых происходит поиск взаимосвязей и за счёт которых, в свою очередь, строится прогноз. Вторых, недостаточной длиной временного ряда.

Однако, несмотря на все плюсы рассмотренного алгоритма, хотелось бы затронуть и минусы, которые были выявлены при его использовании. Итак, основным недостатком данного метода является большое количество параметров, которые необходимо установить экспертно. Причём, при некорректной установке этих параметров работа алгоритма неудовлетворительна. К таким параметрам, в первую очередь, относится метрика при определении компетентности штамма. Её грамотное определение представляет собой отдельную и весьма серьёзную задачу. Также определение метода построения локального и результирующего прогнозного значения — задача не тривиального характера и требует определённой проработки и дополнительных исследований.

Но в целом, особенно при оптимально выбранных параметрах, использование данного алгоритма для краткосрочного прогнозирования показателей, характеризующих развитие социально-экономических явлений, оправданно, особенно в случаях, когда применение других методов затруднительно. Таким образом, данный метод может быть применён, как при некоем предварительном анализе, так и при построении результирующих прогнозных значений.

§ 3. Исследование динамики развития социального явления при большом числе признаков

3.1. Выделение групп сходных признаков. В случае, когда изучаемое социальное явление описывается достаточно большим числом признаков, предполагается разделение их на группы. Разделение производится либо экспертно, либо методами факторного анализа, корреляционных плеед и др. [11]. Каждую такую группу показателей будем называть фактором.

Наибольшими достоинствами при решении задачи разбиения имеют методы главных компонент или главных факторов, позволяющие минимизировать характерность, распределить по факторам до 95 % общности и рассчитать факторные нагрузки таким образом, чтобы исключить влияние специфических факторов на суммарную дисперсию.

С помощью этих методов можно решать задачи отыскания скрытых, но объективно существующих закономерностей, определяемых воздействием внутренних и внешних причин, описания изучаемого процесса числом факторов, значительно меньшим, чем число первоначально взятых признаков. Выявление признаков, наиболее тесно связанных с главными факторами, позволяет выработать и принять научно обоснованное управляющее воздействие, способствующее повышению эффективности функционирования изучаемого процесса.

Главные компоненты — некоррелированные обобщённые показатели, построенные на основе измеренных признаков, могут быть использованы в качестве исходных данных кластерного анализа. Практика показывает, что такая классификация объектов исследования часто оказывается более объективной, чем разделение при помощи отдельных исходных признаков.

При решении практических задач часто используют не все, а только первые главные компоненты. Величина изменения относительной доли накопленной дисперсии, вносимой первыми p главными компонентами, в зависимости от числа этих компонент, даёт основу для решения вопроса о том, сколько главных компонент можно исключить из рассмотрения.

Рассмотрим проблему преступлений, связанных с незаконным оборотом наркотических средств, психотропных и сильнодействующих веществ, обострившуюся за последние годы, которая вносит заметную долю в ухудшение криминогенной ситуации в России. Следует отметить, что динамика этого вида преступлений не имеет ярко выраженной зависимости от определённого социально-экономического показателя. Криминальные процессы, связанные с наркотизацией населения, являются следствием воздействия множества социально негативных факторов, таких как безработица, вынужденная миграция населения, низкий жизненный уровень значительной части населения, детская безнадзорность и пр. Перечень показателей, выбранных для проведения исследования зависимости динамики преступлений, связанных

с незаконным оборотом наркотиков, от социально-экономических процессов в регионах представлен табл. 5.

Для непосредственных расчётов использовались относительные показатели (см. табл. 6.), рассчитанные по фактическим значениям исходных показателей [8], [9], [10].

Таблица 6. Перечень относительных показателей

№ п/п	Относительный показатель
1	Продажа алкоголя на душу населения
2	Отношение числа разводов к населению ниже трудоспособного возраста
3	Отношение числа безработных (МОТ) к экономически активному населению
4	Число вакансий на одного безработного (МОТ)
5	Отношение расходов к доходам
6	Отношение расходов к стоимости минимального набора продуктов питания
7	Отношение числа беженцев и вынужденных переселенцев к численности населения
8	Отношение миграционного прироста к численности населения

Результаты расчётов показывают, что все рассматриваемые относительные показатели имеют значимые ортогональные факторные нагрузки в разложении первых четырёх главных компонент, которые объясняют 80,1 %, 79,1 % и 76,7 % (в среднем — 78,6 %) изменчивости результирующего показателя «Незаконный оборот наркотиков» в 1999, 2000, 2001 гг. соответственно.

В состав первой главной компоненты устойчиво входят (по результатам 1999 г – 2001 гг.) относительные показатели:

- отношение числа разводов к населению ниже трудоспособного возраста;
- отношение числа безработных (МОТ) к экономически активному населению;
- число вакансий на одного безработного (МОТ);
- отношение расходов к стоимости минимального набора продуктов питания.

Показатель «Продажа алкоголя на душу населения» по результатам 1999, 2001 гг. составляет отдельный фактор. Во вторую главную компоненту устойчиво входит единственный показатель «Отношение миграционного прироста к численности населения», объясняющий в среднем 18,1 % изменчивости результирующего показателя. Показатель «Отношение числа беженцев и вынужденных переселенцев к численности

населения» также является отдельным фактором и объясняет в среднем 11,9% изменчивости показателя «Незаконный оборот наркотиков».

Таким образом, все рассматриваемые относительные показатели следует принять в качестве значимых. Исключение любого из них повлечёт снижение накопленного процента относительной доли дисперсии главных компонент.

3.2. Сводный динамический показатель уровня развития социального явления. Проведение исследования матрицы наблюдений методами факторного анализа приводит к выделению групп признаков, сильно коррелированных между собой, при этом признаки каждой группы характеризуют один и тот же аспект. Если на основе выделенных таким образом групп признаков рассчитать показатель уровня развития, то получатся синтетические характеристики, отображающие отдельные аспекты исследуемого явления, соответствующие выделенным факторам.

Одновременное исследование нескольких факторов начинается с установления наборов характеризующих их признаков.

Для каждого отдельно взятого j -го фактора формируется индивидуальная матрица наблюдений X_j ($j = 1, 2, \dots, r$, r — число факторов), элементами которой x_{is} являются значения признака s в момент времени i ($s = 1, 2, \dots, n_j$, $i = 1, 2, \dots, t$, n_j — число признаков, описывающих j -й фактор). Индивидуальные матрицы наблюдений имеют одинаковое число строк. Число столбцов может быть различным и определяется числом признаков.

Сводная матрица наблюдений X_0 имеет блочный вид

$$X_0 = [X_1, X_2, \dots, X_j, \dots, X_r]$$

и характеризует всю совокупность факторов в целом.

Для каждого j -го фактора определяются координаты индивидуального эталона развития P_j , их число равно числу признаков, характеризующих фактор.

Сводный эталон развития P_0 также можно представить в виде блочного вектора, элементами которого являются индивидуальные эталоны развития отдельных факторов

$$P_0^T = [P_1, P_2, \dots, P_j, \dots, P_r].$$

Сводные матрица наблюдений и эталон развития стандартизируются и рассчитываются отклонения до сводного эталона развития

$$c_{i0} = \left[\sum_{j=1}^r \sum_{s=1}^{n_j} \left(z_{is}^j - z_{0s}^j \right)^2 \right]^{1/2}, \quad i = 1, 2, \dots, t,$$

где n_j — число признаков, характеризующих j -й фактор, z_{is}^j — стандартизованные значения признака s в период i для j -го фактора, z_{0s}^j — координаты эталона развития j -го фактора.

Находятся: среднее расстояние до сводного эталона развития

$$\bar{c}_0 = \frac{1}{t} \sum_{i=1}^t c_{i0};$$

стандартное отклонение

$$s_0 = \left[\frac{1}{t-1} \sum_{i=1}^t (c_{i0} - \bar{c}_0)^2 \right]^{1/2};$$

максимально возможное отклонение от сводного эталона

$$c_0 = \bar{c}_0 + 2s_0.$$

Сводный показатель уровня развития определяется формулой

$$d_i^* = \frac{c_{i0}}{c_0}. \quad (17)$$

Сводный показатель уровня развития характеризуется тем, что является величиной положительной и лишь с вероятностью, близкой к нулю, может оказаться больше единицы. Интерпретация его следующая: данное социальное явление находится на тем более высоком уровне развития (в смысле приближения к эталону), чем ближе значение сводного показателя уровня развития к нулю.

На практике удобнее использовать сводный показатель уровня развития в виде

$$d_i = 1 - \frac{c_{i0}}{c_0}, \quad (18)$$

который интерпретируется следующим образом: данное социальное явление имеет тем лучшее состояние, чем ближе значение показателя к единице.

3.3. Определение степени воздействия факторов на сводный показатель. Сводный показатель уровня развития отражает изучаемое социальное явление, характеризуя все совокупные изменений, произошедшие в значениях рассматриваемых факторов, в нём не учитываются изменения, происходящие с отдельными факторами, влияние этих изменений на общую величину показателя.

Для определения направления изменений отдельных факторов и масштабов этих изменений обозначим $c_{i0,j}$ расстояние до индивидуального эталона P_j j -го фактора, исчисленное по стандартизированной индивидуальной матрице наблюдений Z_j

$$c_{i0,j} = \left[\sum_{s=1}^{n_j} (z_{is}^j - z_{0s}^j)^2 \right]^{1/2}, \quad i = 1, 2, \dots, t, \quad j = 1, 2, \dots, r,$$

где $z_{0s,j}$ — координаты эталона развития j -го фактора, z_{is} — стандартизированные значения индивидуальной матрицы наблюдений j -го фактора, n_j — число признаков, характеризующих j -й фактор.

Выражение

$$\begin{aligned} c_{i0}^2 &= \sum_{j=1}^r \sum_{s=1}^{nj} (z_{is} - z_{0s})^2 = \\ &= \sum_{s=1}^{n1} (z_{is} - z_{0s,1})^2 + \dots + \sum_{s=1}^{nj} (z_{is} - z_{0s,j})^2 + \dots + \sum_{s=1}^{nr} (z_{is} - z_{0s,r})^2 \end{aligned}$$

можно записать в виде

$$c_{i0}^2 = \sum_{j=1}^r c_{i0,j}^2. \quad (19)$$

Учитывая, что

$$c_{i0} = c_{0i}^*, \quad (20)$$

полагаем индивидуальный эталон развития j -го фактора равным

$$d_{i,j}^* = \frac{c_{i0,j}}{c_{0,j}}, \quad (21)$$

где $c_{0,j}$ — максимально возможное отклонение от индивидуального эталона развития j -го фактора.

После преобразований получим

$$(d_i^*)^2 = \frac{\sum_{j=1}^r c_{0,j}^2 \cdot (d_{i,j}^*)^2}{c_0^2}.$$

Введём нормирующие множители, постоянные для базового периода ($i = 1, 2, \dots, t$),

$$w_j = \frac{c_{0,j}}{c_0} = \text{const},$$

тогда

$$(d_i^*)^2 = \sum_{j=1}^r w_j^2 \cdot (d_{i,j}^*)^2.$$

Обозначим

$$\Delta d_{i+1}^{*2} = (d_{i+1}^*)^2 - (d_i^*)^2 = \sum_{j=1}^r w_j^2 \cdot \Delta d_{i+1,j}^{*2}, \quad (22)$$

здесь

$$\Delta d_{i+1,j}^{*2} = (d_{i+1,j}^*)^2 - (d_{i,j}^*)^2.$$

Формулу (22) можно использовать для оценки динамики изменений сводного показателя уровня развития социального явления, вызванных различиями в динамике отдельных факторов.

О положительной ситуации в развитии исследуемого социального явления можно говорить тогда, когда величина Δd_{i+1}^{*2} будет отрицательной, поскольку в этом случае последующий $(i+1)$ -й период характеризуется в целом лучшими значениями показателей отдельных

факторов (более близкими к эталонным), что означает принятие показателем d_{i+1}^* меньшего значения по сравнению с его значением d_i^* в предыдущий период.

§ 4. Сравнительный межрегиональный анализ социального явления

4.1. Классификация регионов. В рассматриваемой проблеме незаконного оборота наркотиков не все относительные показатели устойчиво распределяются во времени по главным факторам. Коэффициенты множественной детерминации в среднем принимают значения 0,23.

Таким образом, в целом по всем субъектам Российской Федерации анализ относительных показателей методом главных компонент не даёт устойчивых во времени результатов, а корреляционный анализ показывает отсутствие линейной парной и множественной регрессионной зависимости между показателем «Незаконный оборот наркотиков» и относительными показателями.

Подобных проблем удаётся избежать с помощью предварительного разбиения регионов на группы, имеющие сходные совокупные значения показателей, и последующего исследования внутри каждой группы отдельно. Такой подход, с одной стороны, позволяет выявить наиболее типичный регион и распространить результаты его исследований на другие регионы группы, с другой — анализировать всю группу регионов в целом, а затем проводить межгрупповой сравнительный анализ. Кроме того, распределение показателей по факторам внутри кластеров позволяет выявить региональные особенности.

Практика исследований показывает, что наилучшие разбиения на основе фактических значений показателей получаются методами кластерного анализа [12].

Для 1999–2001 гг. была проведена кластеризация субъектов РФ по стандартизованным значениям 8 относительных показателей (см. табл. 6) при пошаговом увеличении числа кластеров $k = 2, 3, \dots, 10$. Анализ изменения критерия качества разбиения позволил сделать вывод о том, что оптимальное число классов $k^* = 6$.

Устойчивость кластеров устанавливалась на основе критерия близости региона к центру класса. Вопрос о вхождении в кластер «граничного» региона, переходящего из класса в класс, при ретроспективном исследовании решался по результатам 2001 г. Число граничных регионов в группе в среднем составляло 16 %.

Состав групп регионов

1 группа в составе 20 регионов: Московская область, Орловская область, Смоленская область, Тамбовская область, Тверская область, Ярославская область, Республика Карелия, Архангельская область, Вологодская область, Ленинградская область, Республика Башкортостан.

стан, Республика Татарстан, Пермская область, Челябинская область, Красноярский край, Иркутская область, Кемеровская область, Томская область, Республика Саха (Якутия), Амурская область.

2 группа в составе 17 регионов: Брянская область, Воронежская область, Ивановская область, Костромская область, Курская область, Липецкая область, Рязанская область, Тульская область, Псковская область, Республика Адыгея, Краснодарский край, Астраханская область, Волгоградская область, Нижегородская область, Оренбургская область, Пензенская область, Саратовская область.

3 группа в составе 14 регионов: Республика Дагестан, Кабардино-Балкарская Республика, Республика Калмыкия, Карачаево-Черкесская Республика, Республика Марий Эл, Республика Мордовия, Удмуртская Республика, Чувашская Республика, Курганская область, Республика Алтай, Республика Бурятия, Республика Тыва, Читинская область, Еврейская автономная область.

4 группа в составе 14 регионов: Калининградская область, Новгородская область, Ставропольский край, Ростовская область, Кировская область, Самарская область, Ульяновская область, Свердловская область, Республика Хакасия, Алтайский край, Новосибирская область, Омская область, Приморский край, Хабаровский край.

5 группа в составе 4 регионов: Белгородская область, Владимирская область, Калужская область, г. Санкт-Петербург.

6 группа в составе 5 регионов: Республика Коми, Мурманская область, Камчатская область, Сахалинская область, Магаданская область.

Республики Северная Осетия-Алания и Ингушетия, а также Тюменская область и г. Москва при кластеризации не вошли ни в один из классов, являются обособленными регионами.

Для каждой представительной группы регионов (число регионов больше числа относительных показателей) проводился факторный и регрессионный анализ.

Методом главных компонент для каждой группы регионов произведён анализ относительных показателей.

Результаты расчётов показали следующее.

1 группа. В первую главную компоненту вошли относительные показатели:

- отношение числа разводов к населению ниже трудоспособного возраста;
- отношение числа безработных (МОТ) к экономически активному населению;
- число вакансий на одного безработного (МОТ);
- отношение расходов к доходам;
- отношение миграционного прироста к численности населения.

Показатели, вошедшие в первую главную компоненту, объясняют в среднем 39,5 % изменчивости результирующего показателя.

Во вторую главную компоненту вошёл относительный показатель «Отношение числа беженцев и вынужденных переселенцев к численности населения». Он объясняет в среднем 18,0 % изменчивости результирующего показателя.

Третий и четвёртый факторы составляют показатели «Продажа алкоголя на душу населения» и «Отношение расходов к стоимости минимального набора продуктов питания» (13,6 % и 10,5 % объяснения изменчивости результирующего показателя соответственно).

Вся совокупность относительных показателей объясняет в среднем 81,6 % изменчивости результирующего показателя «Незаконный оборот наркотиков — число зарегистрированных преступлений на 100 тыс. населения региона».

2 группа. В первую главную компоненту вошли относительные показатели:

- отношение числа безработных (МОТ) к экономически активному населению;
- число вакансий на одного безработного (МОТ);
- отношение расходов к доходам;
- отношение расходов к стоимости минимального набора продуктов питания;
- отношение миграционного прироста к численности населения.

Они объясняют в среднем 33,1 % изменчивости результирующего показателя.

Вторую главную компоненту составляют показатели «Продажа алкоголя на душу населения» и «Отношение числа беженцев и вынужденных переселенцев к численности населения» (21,2 % изменчивости).

В третий фактор вошёл показатель «Отношение числа разводов к населению ниже трудоспособного возраста» (16,1 % изменчивости).

Вся совокупность относительных показателей объясняет в среднем 70,4 % изменчивости результирующего показателя «Незаконный оборот наркотиков — число зарегистрированных преступлений на 100 тыс. населения региона».

3 группа. В первую главную компоненту вошли относительные показатели:

- отношение числа разводов к населению ниже трудоспособного возраста;
- отношение числа безработных (МОТ) к экономически активному населению;
- число вакансий на одного безработного (МОТ);
- отношение расходов к доходам;
- отношение расходов к стоимости минимального набора продуктов питания.

Они объясняют в среднем 44,7 % изменчивости результирующего показателя.

Вторую главную компоненту (22,7 % изменчивости) составляют показатели:

- продажа алкоголя на душу населения;
- отношение миграционного прироста к численности населения.

В третий фактор вошёл показатель «Отношение числа беженцев и вынужденных переселенцев к численности населения» (13,5 % изменчивости).

Вся совокупность относительных показателей объясняет в среднем 80,9 % изменчивости результирующего показателя «Незаконный оборот наркотиков — число зарегистрированных преступлений на 100 тыс. населения региона».

4 группа. В первую главную компоненту вошли относительные показатели:

- отношение числа разводов к населению ниже трудоспособного возраста;
- отношение числа безработных (МОТ) к экономически активному населению;
- число вакансий на одного безработного (МОТ);
- отношение миграционного прироста к численности населения.

Они объясняют в среднем 32,7 % изменчивости результирующего показателя.

Вторую главную компоненту (21,7 % изменчивости) составляют показатели:

- отношение расходов к доходам;
- отношение числа беженцев и вынужденных переселенцев к численности населения.

В третий фактор (16,1 % изменчивости) вошли показатели:

- продажа алкоголя на душу населения;
- отношение расходов к стоимости минимального набора продуктов питания.

Вся совокупность относительных показателей объясняет в среднем 70,5 % изменчивости результирующего показателя «Незаконный оборот наркотиков — число зарегистрированных преступлений на 100 тыс. населения региона».

Результаты корреляционного анализа, проведённого по каждому представительному кластеру, позволяют сделать следующие выводы:

- между результирующим показателем «Преступления связанные с незаконным оборотом наркотиков» и относительными показателями либо нет парной линейной корреляционной зависимости, либо эта связь слабая;
- между результирующим показателем «Преступления связанные с незаконным оборотом наркотиков» и относительными показателями существует достаточно сильная множественная линейная корреляционная зависимость (коэффициент множественной детерминации в среднем равен 0,8);
- между относительными показателями отсутствует эффект мультиколлинеарности.

Для каждой представительной группы регионов построены уравнения множественной линейной регрессии.

4.2. Сводный показатель уровня развития социального явления. Другим подходом к проведению сравнительного анализа социального явления в региональном разрезе является построение некоторого индикатора состояния, на основе которого производится последующий сравнительный анализ. В качестве индикатора можно рассматривать некоторую свёртку показателей (линейную, мультипликативную и т. д.), первую главную компоненту, и др. [11].

В настоящей работе в качестве индикатора рассматривается показатель уровня развития, построенный по матрице наблюдений X , элементами которой x_{is} являются значения признака s для i -го региона ($s = 1, 2, \dots, n$, $i = 1, 2, \dots, t$, n — число признаков, описывающих социальное явление, t — число регионов).

После стандартизации, введения весовых коэффициентов, эталона развития и возможной группировки факторов показатель уровня развития для каждого региона рассчитывается по формуле (9) либо (18) и имеет ту же интерпретацию.

По изложенной выше методике возможно построение вербально-числовой оценочной шкалы, а также проведение исследования пропорциональности развития социального явления в регионах.

По исчисленным значениям показателя возможно линейное упорядочение и кластеризация регионов.

4.3. Выбор наилучшего региона. С помощью описанного подхода строится эталон «наилучшего» региона, который должен служить примером для данного региона. Роль эталонного объекта в этом случае может выполнять один или несколько регионов, находящихся на более высоком уровне развития.

Расчёт такого эталона некоторого региона P_i производится следующим образом.

Находятся все регионы P_j , для которых:

— расстояния от точки P_i меньше критического значения c_0

$$c_{ij} < c_0,$$

где c_{ij} — расстояние от точки P_i до точки P_j ;

— значения показателя уровня развития больше его значения для рассматриваемого региона

$$d_j > d_i.$$

Построение эталона «наилучшего» региона сводится к расчёту среднеарифметических значений признаков, взятых у выбранных регионов P_j . Этот абстрактный регион представляет точку $P_i^g(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ многомерного пространства.

4.4. Проблема многокритериального выбора. В науке и на практике данная проблема встала достаточно давно. Было разработано

большое количество различных методов и методик, в основе которых, как правило, лежит следующая схема [12]:

1. Сначала осуществляется построение иерархической структуры, необходимой для вычисления интегрального показателя. На этом этапе происходит выбор показателей (параметров), характеризующих объекты в рамках решения конкретной проблемы. Хочется отметить, что одни и те же объекты, но в разных проблемных ситуациях, могут характеризоваться совершенно различными списками показателей. Выбор обычно производится следующим образом: определяются наиболее значимые группы показателей или так называемые обобщённые показатели, важные для решения рассматриваемой проблемы, но непосредственно не измеряемые в силу абстрактного характера. Далее, каждый такой абстрактный показатель детализируется до тех пор, пока не будут получены конкретные параметры, характеризующие объект. В результате этой операции происходит построение иерархического дерева, концевыми элементами которого являются показатели объекта.

2. Далее происходит присвоение весов каждой вершине построенного дерева. Вес фактически характеризует относительную важность показателя.

3. Затем для каждого объекта значения всех показателей каким-либо образом приводятся к одной единой шкале и вычисляется интегральный показатель, который, как правило, является линейной свёрткой выбранных и нормированных параметров.

4. И, наконец, тот объект, у которого значение интегрального показателя выше, признается оптимальным.

Данная схема имеет ряд недостатков.

Во-первых, приведение значений к единой шкале само по себе составляет большую проблему. Существующие методики нормировок и преобразования шкал весьма специфичны и ответить на вопрос, какую из них необходимо использовать в конкретном случае, без предварительного исследования затруднительно. А если учесть, что ряд показателей может носить качественный характер, то корректное проведение этой процедуры аналитиком, не достаточно хорошо разбирающимся в математических методах, практически невозможно.

Во-вторых, линейный характер интегрального показателя можно поставить под сомнение, и целый ряд примеров подтверждает неправомерность такого суждения.

Кроме того, существуют другие недостатки описанного подхода, специфичные для решения конкретных прикладных задач.

Чтобы сгладить эти недочёты, разработаны методики, позволяющие учесть взаимосвязи между показателями, в том числе ориентированные на качественные шкалы. Однако при большом количестве показателей применение таких методик затруднительно вследствие значительной трудоёмкости решения задачи. Так, при N показателях, измеряемых в качественных m -бальных шкалах, число предварительных экспертных оценок, которые необходимы для расчёта интегрального показате-

ля, будет составлять $N(N-1)/(2m^2)$. Например, при 10 параметрах, измеряемых в 5-бальных шкалах, это число превышает 1000. При этом следует отметить, что речь идёт всего лишь о предварительной настройке механизма расчёта интегрального показателя.

Поэтому для решения задач данного класса предлагается использовать специальный подход, основанный на многокритериальном выборе предпочтительных объектов с использованием формальной процедуры выделения оптимальных по Парето альтернатив.

4.5. Многокритериальный выбор, оптимальный по Парето.

Пусть имеется N объектов. Каждый объект характеризуется n -мерным вектором $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$, где n — количество показателей, x_{ij} — значение j -го показателя для i -го объекта, $i \in \overline{1, N}$. Какое количество и какие конкретно показатели необходимо использовать, решает непосредственно эксперт-аналитик. Причём, если имеется иерархическая структура (полученная с использованием описанной выше стандартной схемы), то вектор, характеризующий объект, будет состоять из набора показателей концевых вершин.

Учтём значимость отдельных показателей во всей группе. Для этого введём общее понятие вектора значимости $W = (\omega_1, \omega_2, \dots, \omega_n)$, где $\omega_i > 0$, $i \in \overline{1, n}$, ω_i — числовое значение, характеризующее значимость i -го показателя.

В случае, когда каждому показателю можно приписать некоторый вес, будем считать, что ω_i есть вес i -го показателя, который характеризует его относительную важность, при этом $\sum_{i=1}^n \omega_i = 1$.

Однако достаточно часто аналитику трудно определить веса, то есть он не может ответить на вопрос, на сколько или во сколько один показатель важнее другого. В то же время упорядочить показатели по степени важности — ранжировать эти показатели, то есть ответить на вопрос, какой показатель важнее (или сказать, что они равнозначные), эксперт может практически всегда. В этом случае будем считать, что ω_i — это ранг i -го показателя, где $i \in \overline{1, n}$.

Наконец, в случае, когда все показатели равнозначны или если не удаётся определить даже ранги показателей, будем считать, что $\omega_i = 1$, где $i \in \overline{1, n}$, то есть важность всех показателей одинакова.

Таким образом, вектор значимости показателей определён всегда.

Будем говорить, что показатель имеет прямую степень ранжирования, если большее значение по этому показателю предпочтительнее, и обратную степень ранжирования, если большее значение показателя менее предпочтительно. Примером показателей с прямой степенью ранжирования может быть показатель «Доходы населения», а с обратной — «Количество преступлений». В первом случае, чем выше доходы населения, тем лучше; во втором, чем выше преступность, тем хуже.

Примем, что i -й объект доминирует все объекты по j -му показателю, если $x_{ij} \geq x_{kj}$, $\forall k \in \overline{1, N}$, при условии, что показатель имеет

прямую степень ранжирования и $x_{ij} \leq x_{kj}$, $\forall k \in \overline{1, N}$, когда степень ранжирования показателя обратная. Кроме этого, в общем случае, можно сказать, что i -й объект доминирует все объекты по j -му показателю, если $x_{ij} \geq x_{kj}$, $\forall k \in \overline{1, N}$, где $\geq$ — отношение превосходства (а не обычное числовое отношение «больше»). Последнее замечание позволяет учесть случаи, когда трудно сказать, какое значение показателя лучше — большое или маленькое (пример — показатель «температура человеческого тела»), а также позволяет не привязываться к количественной шкале (однако данное отношение всё-таки должно удовлетворять свойству транзитивности).

Будем говорить, что i -й объект доминирует j -й объект, если i -й объект доминирует j -й объект по всем показателям ($X_i \geq X_j$).

Оптимальными по Парето являются те объекты, которые не доминируются другими объектами, т.е. X_i — оптимально по Парето, если $\exists j \in \overline{1, n} \setminus \{i\}: X_j \geq X_i$. Множество таких оптимальных объектов будем называть ядром Парето.

Проблема оптимальности по Парето заключается в том, что на практике (особенно при большом количестве показателей) число формально выбранных оптимальных объектов может быть достаточно большим, а в предельном случае — равным числу исходных объектов. В этой ситуации никакой пользы от их поиска нет. Поэтому необходимо ввести процедуру, позволяющую сужать данное множество оптимальных объектов до выделения из них одного-двух.

Сущность разработанной процедуры выбора на содержательном уровне заключается в следующем — необходимо найти объект, доминирующий остальные объекты по показателям, имеющим максимальную суммарную значимость. Если таких объектов несколько, а это вполне возможно, хотя и маловероятно, то считаем, что они равнозначны и вся эта группа оптимальна. В случае равнозначности показателей ($\omega_i = 1, \forall i \in \overline{1, n}$) это означает, что «оптимальный» (по данному выбору) объект доминирует, то есть превосходит все объекты, по наибольшему количеству показателей. Очевидно, что выбранные таким образом объекты лежат в ядре Парето, то есть являются его подмножеством и соответственно оптимальны по Парето. Причём мощность (количество элементов) найденного множества значительно меньше мощности ядра и, как правило, равна единице.

Формально данный критерий выбора можно записать следующим образом:

$$X^{\text{опт.}} = \arg \max_{i \in \overline{1, N}} (F(X_i)),$$

где $X^{\text{опт.}}$ — искомый оптимальный объект, $F(u)$ — функция полезности, определённая как

$$F(X_i) = \sum_{j=1}^n \omega_j I_j(X_i),$$

$I_j(X_i)$ — индикатор того, что объект X_i доминирует все объекты по показателю j ,

$$I_j(X_i) = \begin{cases} 1 & \text{если } X_i \text{ доминирует все объекты по } j\text{-ому показателю,} \\ 0 & \text{в противном случае.} \end{cases}$$

Итак, разработанная процедура является формальной процедурой, основанной на сужении ядра оптимальных по Парето объектов. К основному достоинству данной процедуры следует отнести её достаточно универсальный и в то же время формальный характер. Всё, что необходимо знать для применения процедуры, — это ответ на вопрос «Какое из двух значений отдельного показателя лучше?».

С целью апробирования данной методики был проведён ряд практических экспериментов. Суть исследования заключалась в том, чтобы упорядочить регионы уральского федерального округа по месту, занимаемому в незаконном обороте наркотиков, т.е. на первом месте должен оказаться регион, в котором незаконный оборот наркотиков минимален. Стоит отметить, что задача упорядочения объектов фактически идентична задаче выбора «оптимального». Действительно, первый объект (или последний, в зависимости от типа упорядочения) в упорядоченном списке будет «наилучшим». С другой стороны, если в группе объектов находить оптимальный, изымать его из списка и снова повторять процедуру поиска, то в результате получим упорядоченную последовательность.

Итак, для решения поставленной задачи экспертно был определён набор показателей, который характеризует каждый регион в рамках рассматриваемой предметной области. Перечень этих показателей приведён в табл. 6., а значения, уже непосредственно описывающие объект, в табл. 7. Как уже отмечалось, список параметров был определён экспертно и, безусловно, может быть изменён для получения более точных (или, наоборот, для более округлённых) результатов.

Таблица 7. Значения показателей, характеризующих незаконный оборот наркотиков в регионах

Регион	Относительный показатель							
	1	2	3	4	5	6	7	8
1	1,50	0,02	0,26	0,02	0,82	1,58	0,40	-5,13
2	1,56	0,03	0,23	0,04	1,02	2,53	0,11	0,93
3	2,12	0,03	0,25	0,02	0,49	3,07	0,26	3,86
4	2,24	0,03	0,24	0,02	0,41	3,29	0,26	10,51
5	3,38	0,03	0,24	0,03	0,42	3,29	0,08	0,45
6	1,86	0,03	0,24	0,01	0,74	1,97	0,16	0,11

В результате работы описанной выше процедуры все регионы Уральского федерального округа были упорядочены и результаты были сравнены с показателем — «Незаконный оборот наркотиков — число зарегистрированных преступлений на 100 тыс. населения». Результаты сравнительного анализа приведены в табл. 8.

Таблица 8. Сравнительные данные по вычисленным и реальным значениям

Регион	Место региона (по данным МВД)	Место региона, определённое процедурой
1	1	1
2	3	4
3	5	5
4	6	6
5	4	3
6	2	2

Результаты показывают, что разработанная процедура достаточно хорошо описывает ситуацию.

Особое внимание необходимо уделить вопросу вычисления ошибки упорядочения регионов. Предлагается воспользоваться следующей формулой $err_{абс} = \sum_{i=1}^N \left| x_i^{факт} - x_i^{прогн} \right|$, где N — количество объектов,

$x_i^{факт}$ — фактическое место i -го объекта, $x_i^{прогн}$ — спрогнозированное место i -го объекта. Фактически эта формула характеризует абсолютную ошибку. Но абсолютная ошибка плохо отражает реальную действительность, так как сразу трудно определить диапазон изменения ошибки (эта проблема присуща всем абсолютным величинам). Поэтому целесообразно воспользоваться нормированной (относительной) величиной ошибки. В качестве нормировочного коэффициента предлагается взять максимальное значение абсолютной ошибки. Это удобно тем, что тогда относительная ошибка будет лежать в отрезке от 0 до 1 (или от 0 до 100 %).

$$err_{отн} = \frac{err_{абс}}{err_{max}} \cdot 100\%$$

Теперь для вычисления относительной ошибки необходимо определить максимальное значение абсолютной ошибки. Поэтому надо проанализировать ситуацию и ответить на вопрос: «Когда она возникает?». Не вызывает сомнения, что максимальная ошибка получается, когда произошло диаметрально противоположное ранжирование, т. е. худший объект стал лучшим, а лучший худшим и т. д. Несложными математи-

ческими вычислениями получаем следующее её значение:

$$\text{err}_{\text{отн}} = \begin{cases} N^2/2 & \text{при } N = 2k, \text{ где } k \in \mathbb{N}, \\ 2[N/2]([N/2] + 1) & \text{при } N = 2k + 1, \text{ где } k \in \mathbb{N}, \end{cases}$$

где N — количество объектов.

С учётом введённых выше понятий величина ошибки составила $\text{err}_{\text{отн}} \approx 11\%$. Данную ошибку можно объяснить не полным перечнем показателей, характеризующим регионы в рамках данной социальной проблемы. Но с другой стороны, её величина достаточно невысокая — не совпадают всего лишь два места. Таким образом, процедура многокритериального выбора доказала свою работоспособность и может быть рекомендована для применения при анализе социально-экономических явлений.

Заключение

Построен комплекс математических моделей анализа, оценки и прогнозирования состояния сложных, неопределяемых численно социальных явлений, таких как преступность, коррупция, бедность, социальное неравенство и др.

Показатель уровня развития социального явления в регионе, идентифицированный по вербально-числовой шкале, позволяет качественно и количественно оценить динамику изменения изучаемого процесса, пропорциональность развития и близость к эталону как цели и условию успешного развития.

Модификация показателя уровня развития позволяет учесть направления и масштабы изменений, происходящих с отдельными признаками и факторами, выявить влияние этих изменений на общую величину показателя.

Проблема прогнозирования развития социально-экономических явлений, описываемых набором взаимосвязанных признаков, при отсутствии информации об их взаимозависимости, решена с помощью обучающегося генетического алгоритма.

Группировка регионов со сходными социально-экономическими показателями, позволяет выявлять особенности и скрытые объективно существующие закономерности развития социальных процессов. Выделение факторов позволяет сформировать группы признаков, сильно коррелированных между собой и отображающих отдельные аспекты исследуемого явления.

Построенный индикатор, идентифицированный по оценочной шкале, позволяет проводить сравнительный анализ состояния изучаемого процесса для группы регионов. Метод многокритериального выбора позволяет, кроме упорядочения, определять усреднённые показатели наилучшего региона как промежуточной цели развития.

Список литературы

1. Карпенко Н.В. Оценка степени влияния факторов на изменение социально-экономической ситуации в регионе / Труды юбилейной научно-технической конференции. Часть II. — М.: Изд. ЦНИИРЭС, 2001.
2. Карпенко Н.В. Влияние факторов на социально-экономическую ситуацию в регионе / Математическое моделирование социальных процессов. Вып. 4. — М.: Изд. МГУ, 2002.
3. Плюта В. Сравнительный многомерный анализ в экономических исследованиях. Пер. с польского. — М.: Статистика, 1980.
4. Goldberg D.E. Genetic Algorithms in search, optimization and machint lean- ing. Adison Wesley, Reading, MA, 1989.
5. Кендалл М., Стъарт А. Статистические выводы и связи. М: Наука, 1973. — 900 с.
6. Данчул А.Н., Корнееенко В.П. Системный анализ управления экономиче- скими процессами. — М.: Изд. РАГС, 2001. — 140 с.
7. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. — Новоси- бирск: Изд. Института математики, 1999. — 270 с.
8. Российский статистический ежегодник. 1999.: Стат. сб / Госкомстат Рос- сии. — М., 1999.
9. Российский статистический ежегодник. 2000.: Стат. сб / Госкомстат Рос- сии. — М., 2000.
10. Российский статистический ежегодник. 2001.: Стат. сб / Госкомстат Рос- сии. — М., 2001.
11. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы экономет-рики. — М.: ЮНИТИ, 1998. — 1022 с.
12. Айвазян С.А., Бежиева З.И., Староверов О.В. Классификация многомер-ных наблюдений. — М.: Статистика, 1974. — 240с.

Глава VI

МОДЕЛИРОВАНИЕ ВЗАИМОСВЯЗЕЙ ПРОБЛЕМ ПРИ ОБРАБОТКЕ ТЕКСТОВ

Введение

Важным источником информации о социально-экономическом и общественно-политическом развитии страны в целом и отдельных субъектов Российской Федерации являются средства массовой информации (СМИ). Несмотря на порой ангажированный характер информации, содержащийся в отдельных изданиях и электронных СМИ, существующие сегодня средства её обработки, основанные на использовании закона больших чисел, позволяют исследователям и аналитикам достаточно точно оценивать процессы, протекающие в обществе. При обработке больших объёмов текстовой информации наибольшую сложность у аналитиков вызывает поиск и анализ возможных взаимосвязей между изучаемыми проблемами в различных предметных областях. Анализ такого рода взаимосвязей позволяет исследователям воссоздать общую картину социально-экономических отношений как единой системы в целом по стране, региону и тем самым способствует выработке рекомендаций по эффективному решению рассматриваемых проблем.

С этой точки зрения представляет интерес определение эффективной стратегии поиска исходной информации для анализа взаимосвязей проблем P_i , P_j , произвольно фиксируемых в составе некоторого множества проблем $M_p = \{P_1, \dots, P_i, \dots, P_j, \dots, P_h\}$. Данная задача рассматривается здесь в рамках определённых ограничений, в значительной степени детерминирующих, с одной стороны, действия эксперта-аналитика, с другой, — структуру и функцию программных средств, ориентированных на поддержку указанных действий при поиске требуемой информации. При этом, в частности, предполагается что исходная совокупность сообщений СМИ, служащая источником выбора необходимых аналитику данных, представляет собой массив рубрицированных документов M_d , документы которого полностью или частично относящиеся к проблеме $P_i \in M_p$, образуют класс $r_i \in M_d$. Выраженное в естественном языке имя данного класса, совпадающее с наименованием (именем) проблемы P_i фиксируется в качестве i -й статьи рубрикатора R . Соответственно, выражение « i -я рубрика» в зависимости от контекста далее означает или класс r_i документов, относящихся к проблеме P_i , или название проблемы P_i , фиксированное в качестве статьи используемого рубрикатора.

Следующее важное допущение касается способа формализованного представления отдельно взятой проблемы P_i из множества M_p . Предполагается, что в общем случае такое представление является результатом определённым образом организованного просмотра некоторой совокупности документов M_o^i экспертом-аналитиком и последующей редукции сформированного им концепта проблемы к некоторой когнитивной структуре Ks_i , которая и фиксируется в формализованном языке модели поиска. Здесь под когнитивной структурой Ks_i проблемы P_i понимается множество слов T_i (терминологический портрет P_i) и пар слов из T_i , играющих ключевую роль в раскрытии содержания проблемы P_i .

В § 1 рассматривается процедура построения когнитивной структуры проблемы P_i в контексте решения ещё одной, вспомогательной, задачи: формирования выборки относящихся к P_i документов с целью выделения в M_d рубрики $r_i \in M_d$, соответствующей данной проблеме, в случае, если эта рубрика заранее отсутствует. Тем самым предусматривается возможность оперативного включения возникающих новых проблем в сферу применения рассматриваемой модели поиска. Целесообразность такого подхода обусловлена также и тем обстоятельством, что, в рамках принятой концепции, совместно с построением когнитивной структуры проблемы завершается и основной этап решения данной вспомогательной задачи.

Важным этапом анализа взаимосвязи проблем является выявление причинно-следственных зависимостей. Главная роль здесь отводится выдвижению и проверке гипотез о возможных причинах тех или иных явлений. Такие предположения позволяют наметить границы, в рамках которых действуют, либо прекращают действовать причинные (казуальные) закономерности и, тем самым, сузить область поиска причины того или иного явления. В результате, применение известных методов установления причинной связи, таких, например, как метод сходства, метод различия [1] становится более результативным.

В § 2 применительно к задаче установления области поиска возможных причин значимых изменений частоты рубрик строится модель взаимосвязей рубрик. Основу этой модели составляет множество рубрик $\{r_i\}$ с заданным на нём отношением, связывающим рубрики r_i, r_j при наличии в них общих документов, количество которых определяет вес связи этих рубрик a_{ij} . Исходную информацию модели составляет матрица $\|a_{ij}\|$, применение которой для выявления причинно-следственных связей достаточно ограничено. В этой связи в § 3 рассматривается модель, обеспечивающая возможность более гибкого описания условий, ограничивающих область поиска возможных причин исследуемых событий.

Наконец, в § 4 приводится описание модели поиска, ориентированной на поддержку действий исследователя при подборке исходной информации для анализа взаимосвязей произвольной пары проблем P_i, P_j из множества M_p .

§ 1. Проблемы и их представление в моделях поиска; выделение актуальных проблем для анализа взаимосвязей

1.1. Представление процесса формирования системы знаний по проблеме как итеративной процедуры. При поиске по рубриктору R отбор документов, относящихся к некоторой проблеме (теме) P , осуществляется экспертом-аналитиком непосредственно, исходя из заданного распределения документов по рубрикам. Если соответствующая проблеме P рубрика r отсутствует, она, при необходимости, может быть создана, например, с применением средств авторубрицирования. В этом случае, однако, требуется решение вспомогательной задачи — формирования обучающей совокупности $M_{об}$ на базе некоторого множества документов M_0 . Ниже рассматривается ряд вопросов, связанных с её постановкой и решением, приводится описание логико-лингвистической модели, ориентированной на поддержку действий эксперта-аналитика при решении этой задачи.

Сходство задач рубрицирования множества документов M_d и формирования обучающей совокупности $M_{об}$ очевидно: как одна, так и другая, вне зависимости от подхода к её решению, включает в качестве основной компоненты подзадачу отнесения документа из заданного множества документов к некоторой рубрике. Вместе с тем, имеют место и определённые различия. В задаче рубрицирования документов множества M_d исходными являются требования полноты и точности выделения рубрик. При формировании обучающей совокупности $M_{об}$ требуется лишь точность выделения документов. Обязательным является также условие, согласно которому совокупность $M_{об}$ должна отражать всё разнообразие представленных в документах M_0 аспектов проблемы P , значимых для проводимого пользователем исследования. Ещё одно требование состоит в том, что документы $M_{об}$ не должны содержать «лишней» информации, и, следовательно, являясь документами рубрики r , одновременно принадлежать каким-либо другим рубрикам рубриктора R . Таким образом, полноты поиска документов, принадлежащих r , в этом случае не требуется: необходимо включение в состав $M_{об}$ лишь некоторой части таких документов.

Ясно, что поиск документов, включаемых в состав $M_{об}$, с необходимостью базируется на сопоставлении наличного знания о проблеме P с содержанием каждого из документов M_0 . При этом важно, что само знание о проблеме P в ходе поиска не остаётся неизменным: возможно его пополнение, корректировка; соответственно, подлежат изменению и критерии отбора документов множества M_0 в состав $M_{об}$. Наличие тесной связи процедур формирования обучающей совокупности $M_{об}$ и построения системы знаний о проблеме P свидетельствует о том, что решение рассматриваемой задачи формирования $M_{об}$ в общем случае может быть достигнуто лишь в ходе итеративного процесса. Согласно сказанному, каждый шаг этого процесса вплоть до шага, завершающе-

го построение системы знаний о проблеме P , может быть представлен в виде следующей последовательности подлежащих выполнению процедур:

1) исходя из наличного знания о проблеме P , поиск документов $d_q \in M_0$ ($q = 1, 2, \dots, Q$), имеющих отношение к данной проблеме;

2) просмотр документов, относящихся к проблеме P ; выделение документов, включаемых в состав рубрики r ; в ходе просмотра — фиксация и накопление элементов нового знания о данной проблеме;

3) объединение имевшегося ранее знания о проблеме (п. 1) с новым в единую систему;

4) корректировка критериев соотнесения документов M_0 рубрики r в соответствии с полученной системой знаний о проблеме P .

Начало данного процесса определяется заданием исходной системы знаний о проблеме, а также согласованных с ней критериев отнесения документов множества M_0 рубрики r . На последнем, заключительном этапе решения задачи просмотр документов $d_q \in M_0$, относящихся к r , сопровождается выделением тех из них, которые не содержат «лишней» информации, и включением последних в состав обучающей совокупности M_{06} ¹⁾.

Уже одно перечисление подлежащих выполнению подзадач свидетельствует о том, что в общем случае алгоритмизация процесса решения рассматриваемой задачи требует привлечения весьма сложных средств анализа содержания как самой проблемы, так и документов исходной совокупности M_0 . Использование таких средств, однако, выходит за рамки принятого здесь подхода. Далее речь идёт лишь о достаточно простом инструментарии, ориентированном на поддержку действий пользователя, направленных на формирование обучающих совокупностей M_{06} .

Концептуальной основой инструментария служит модель, элементами которой являются проблема P и документы (сообщения СМИ), образующие множество $M_0 = \{d_q : q = 1, 2, \dots, Q\}$. Проблема представлена в модели своей когнитивной структурой Ks . В составе Ks выделяется также некоторая подструктура $\bar{K}s$, отдельные элементы — слова или пары слов которой в контексте содержания документов M_0 устойчиво ассоциируются с рассматриваемой проблемой. Вхождения элементов $\bar{K}s$ в тексты документов M_0 называются далее потенциальными ссылками на проблему P .

Характеристиками произвольного $d_q \in M_0$ являются объём $V(d_q)$ и степень релевантности d_q проблеме P — $\text{Rel}(d_q)$. Объём $V(d_q)$ измеряется числом символов в составе текста документа; степень ре-

¹⁾ Случай, когда за недостатком имеющихся в распоряжении материалов при формировании обучающих совокупностей используются отдельные фрагменты документов M_0 , удовлетворяющие указанным условиям, принципиальных отличий не имеет и, поэтому, здесь не рассматривается.

релевантности документа d_q проблеме P — количеством потенциальных ссылок на P , имеющих вхождение в текст данного документа. Таким образом, степень релевантности проблеме P документов, не содержащих ссылок на данную проблему, равняется 0: считается, что, в отличие от остальных, такие документы отношения к проблеме P не имеют.

Вводится в рассмотрение и ещё один показатель — «качество» документа d_q , характеризующий на содержательном уровне реальные шансы d_q на включение аналитиком в состав обучающей совокупности $M_{об}$. Данный показатель, хотя в явном виде и не фиксируется в формализме модели, однако, неявно, в силу ожидаемого характера его зависимости от характеристик $V(d_q)$ и $Rel(d_q)$, в значительной степени определяет состав функций, выполняемых инструментарием.

Действительно, весьма правдоподобным представляется допущение о том, что поведение «качества» документа d_q как функции двух аргументов — объёма d_q и степени релевантности d_q проблеме P , имеет следующую особенность: при фиксированном значении $V(d_q)$ с увеличением степени релевантности d_q проблеме P «качество» документа возрастает; напротив, при $Rel(d_q) = \text{const}$ с ростом объёма документа d_q его «качество» снижается. Поэтому, если упорядочить документы M_0 по степени релевантности проблеме, разбив их на группы, соответствующие последовательным значениям степени релевантности — 1, 2, 3, ... и затем в каждой из полученных групп упорядочить документы по объёму, то, фиксируя группу с достаточно большим значением степени релевантности и извлекая из неё документы в порядке увеличения их объёма, начиная с минимального, можно ожидать появления в первую очередь «достаточно качественных» документов.

Нетрудно заметить, что без развитых средств лингвистического анализа текстов, полнота поиска документов M_0 , относящихся к рубрике r , недостижима. То, что данное обстоятельство приводит к определённым ограничениям возможности выбора документов $M_{об}$ из множества M_0 на самом деле не столь существенно. Важно лишь, чтобы и при таких ограничениях в документах $M_{об}$ были затронуты все наиболее значимые, актуальные аспекты проблемы P , получившие своё освещение в заданном массиве сообщений СМИ — M_0 . Однако, согласно принятому здесь подходу — это задача аналитика, которую необходимо иметь в виду при каждой из выполняемых им процедур: и при выявлении, накоплении новых элементов знания о проблеме P , и при синтезе полученного знания в единую систему, и, наконец, при редукции системы знаний о проблеме к когнитивным структурам $K's$ и $\bar{K}'s$, представляющим эту систему в формализованной модели. В последнем случае в качестве методической основы формирования $K's$, $\bar{K}'s$ целесообразно исходить из положений, развитых в работах лингвистов и изложенных, например, в [2], [3].

Согласно этим положениям, основными смыслообразующими компонентами системы знаний о проблеме являются аргументно-предикатные структуры, представляющие в естественном языке ситуации, связанные с событиями, процессами, действиями, характеризующими эту проблему. Центральное место в таких структурах занимает предикатор (глагол, отглагольное существительное и т. д.), представляющий само действие. Сюда же входят имена или именные группы (на аргументных местах предикатора), называющие участников ситуации, выполняющих в ней определённые роли: например, действующего лица, объекта, на который направлено действие, инструмента, которым данное действие совершается и т. д. При этом количество и характер ролей всецело определяется семантикой предикатора. Таким образом, если при решении поставленной задачи ориентироваться на поиск документов по описаниям ситуаций, в состав K_s следует включать пары, образованные именами основных участников ситуаций и соответствующими предикаторами: чем более точно будут подобраны указанные элементы пар для включения в когнитивную структуру проблемы, тем применение инструментария окажется более эффективным.

1.2. Выявление значимых рубрик. Пусть в течение некоторого месяца поступает N статей, каждая из которых относится к одной или нескольким рубрикам. Некоторой k -й рубрике ($k = 1, 2, \dots, n$) ставится в соответствие число статей N_k , попавших в эту рубрику. При этом

$$\sum_{k=1}^n N_k = N_c \geq N. \quad (1)$$

При ежемесячном рубрицировании поступающих статей для k -й рубрики определяется относительная частота $n_k(t)$

$$n_k(t) = \frac{N_k}{N_c}, \quad (2)$$

где t — число наблюдаемых временных тактов, ($t = 1, 2, \dots, p$). Совокупность наблюдений $n_k(t)$ (далее обозначенных n_t) образует временной ряд:

$$n_t = y_t + \varepsilon_t, \quad (3)$$

где y_t — детерминированная неслучайная компонента динамического ряда, ε_t — стохастическая случайная компонента.

Поскольку построение прогнозного значения относительной частоты производится только для ответа на вопрос о возможном изменении тенденции в первом месяце после наблюдаемого периода (прогноз на один шаг вперёд по времени) случайной компонентой пренебрегаем, а в качестве аппроксимирующей функции выбираем линейную:

$$y_t = a_0 + a_1 \cdot t, \quad (4)$$

где a_0, a_1 — параметры, определяемые методом наименьших квадратов, t — время ($t = 1, 2, \dots, p$), p — число временных тактов.

Качество модели прогнозирования (3) определяется на основе свойств ряда остатков:

$$y_t - y_{\tau t}, \quad t = 1, 2, \dots, p, \quad (5)$$

т.е. величины расхождений на участке аппроксимации между фактическим уровнями y_t и их расчётными значениями $y_{\tau t}$.

Соответствие ряда остатков нормальному закону распределения важно с точки зрения правомерности построения интервала прогноза. Основными свойствами ряда остатков является их симметричность относительно тренда и преобладание малых по абсолютной величине ошибок над большими. В этой связи определяется близость к соответствующим параметрам нормального закона распределения коэффициентов асимметрии — A_c и эксцесса Θ_k наблюдений около модели, т.е.

$$A_c = \frac{\frac{1}{p} \sum_{i=1}^p (y_i - y_{\tau i})^3}{\left(\frac{1}{p} \sum_{i=1}^p (y_i - y_{\tau i})^2 \right)^{3/2}}, \quad (6)$$

$$\Theta_k = \frac{\frac{1}{p} \sum_{i=1}^p (y_i - y_{\tau i})^4}{\left(\frac{1}{p} \sum_{i=1}^p (y_i - y_{\tau i})^2 \right)^2} - 3 \quad (7)$$

Если эти коэффициенты близки к нулю или равны нулю, то ряд остатков распределён в соответствии с нормальным законом. Для оценки степени их близости к нулю вычисляют средние квадратические отклонения:

$$S_a = \frac{6(p-2)}{(p+1)(p+3)}, \quad (8)$$

$$S_s = \frac{24p(p-2)(p-3)}{(p+1)^2(p+3)(p+5)} \quad (9)$$

Если выполняются соотношения:

$$A_c \leq 1,5 \cdot S_a, \quad (10)$$

$$\Theta_k \leq 1,5 \cdot S_s, \quad (11)$$

то считается, что закон распределения ряда остатков не противоречит нормальному закону. В случае когда $A_c > 2 \cdot S_a$ или $\Theta_k > 2 \cdot S_s$ распределение ряда не соответствует нормальному закону, и построение доверительных интервалов прогноза неправомерно. В случае попадания A_c и Θ_k в зону неопределённости (между полутора и двумя среднеквадратическими отклонениями) может быть использован RS-критерий:

$$RS = \frac{(E_{\max} - E_{\min})}{S}, \quad (12)$$

где $E_{\max}$ — максимальный уровень ряда остатков, $E_{\min}$ — минимальный уровень ряда остатков, S — среднее квадратическое отклонение остатков.

Если значение этого критерия попадает между табулированными границами с заданным уровнем значимости, то гипотеза о нормальном распределении ряда остатков принимается.

Подчинение ряда остатков нормальному закону позволяет построить доверительный интервал прогнозируемого значения относительной частоты, т. е. определить допустимые отклонения прогноза от основной тенденции. Для решения этой задачи построим интервальные оценки параметров регрессии a_0 и a_1 (2) в формах:

$$a_0 = a'_0 \pm t_\nu \cdot \sigma_{a'0}, \quad (13)$$

$$a_1 = a'_1 \pm t_\nu \cdot \sigma_{a'1}. \quad (14)$$

Здесь серединами интервалов являются точечные оценки a'_0 и a'_1 , рассчитанные с помощью метода наименьших квадратов. Величина t_ν — теоретическое значение критерия Стьюдента при уровне значимости, равном 5%, и числе степеней свободы $\nu = p - 2$.

Стандартные ошибки коэффициентов регрессии вычисляются по следующим формулам

$$\sigma_{a'0} = \sqrt{\frac{\sigma^2}{p}}, \quad (15)$$

$$\sigma_{a'1} = \sqrt{\frac{\sigma^2}{\sum_{i=1}^p (t_i - t_{cp})^2}} \quad (16)$$

Несмещённая оценка дисперсии σ^2 случайной составляющей

$$\sigma'^2 = \frac{1}{p-1} \sum_{i=1}^p (y_i - y_{\tau i})^2, \quad (17)$$

где y_i — фактические значения динамического ряда, $y_{\tau i}$ — теоретические значения, рассчитанные по уравнению регрессии, t_i , t_{cp} — фактические значения и среднее значение фактора времени.

Верхняя и нижняя границы доверительного интервала в точке прогноза на один такт вперёд будут равны

$$y_{\text{в}} = a'_0 + a'_1 + t_p (\sigma_{a'0} + \sigma_{a'1} \cdot t_{p+1}), \quad (18)$$

$$y_{\text{н}} = a'_0 + a'_1 - t_p (\sigma_{a'0} + \sigma_{a'1} \cdot t_{p+1}).$$

Непопадание фактического значения относительной частоты в доверительный интервал ($y_{\text{н}}$, $y_{\text{в}}$) служит индикатором нарушения тенденции либо проявлением сезонности. В этом случае рубрику можно отнести к группе значимых рубрик.

Если подчинение ряда остатков нормальному закону не подтверждается, также возможно использование границ интервала прогноза в форме (18), хотя в этом случае на результаты анализа не распространяется свойство надёжности.

§ 2. Установление связей между значимыми рубриками

2.1. Уровни связей. После рубрицирования статей, поступивших за фиксированный период времени, составляется матрица связей A , элементами которой a_{ij} ($i, j = 1, 2, \dots, n$, n — число рубрик) являются частоты принадлежности статей одновременно к i -й и j -й рубрикам, a_{ii} — число статей, принадлежащих только i -й рубрике. Элементы матрицы A обладают свойством симметричности относительно главной диагонали $a_{ij} = a_{ji}$ ($i \neq j$).

Матрице A соответствует неориентированный граф, вершинами которого являются рубрики. Наличие ребра (связи) между i -й и j -й рубриками определяется ненулевым значением элемента матрицы связей $a_{ij} \neq 0$ (при $a_{ij} = 0$ ребро отсутствует). Под весом связи будем понимать значение a_{ij} .

По матрице A составляется матрица B , элементы которой $b_{ij} = 1$, если $a_{ij} \neq 0$, и $b_{ij} = 0$, если $a_{ij} = 0$.

Среди всей совокупности рубрик выделены m ($m \leq n$) рубрик, внутри которых произошли большие изменения, такие рубрики назовём значимыми.

Возникают следующие задачи.

Задача 1. Среди полного набора n рубрик выделить рубрику (несколько рубрик) имеющих существенную связь с отдельно взятой значимой рубрикой.

Задача 2. Среди полного набора n рубрик выделить рубрику (несколько рубрик), связывающую две проблемные рубрики.

Для выявления и упорядочения рубрик, имеющих связь с отдельно взятой значимой рубрикой обозначим: k — номер значимой рубрики, j — номер текущей рубрики.

Процедура решения первой задачи состоит из следующих этапов.

На первом этапе устанавливается наличие связи первого уровня, т.е. связи между значимой и текущей рубриками (рис. 1).

В случае существования такой связи ($a_{kj} \neq 0$) вводится величина

$$S_j^{1k} = b_{kj} = 1, \quad (k \neq j, \quad j = 1, 2, \dots, n) \quad (19)$$

и фиксируется вес установленной связи

$$S_j^{2k} = a_{kj}. \quad (20)$$

При отсутствии связи

$$S_j^{1k} = 0, \quad S_j^{2k} = 0. \quad (21)$$

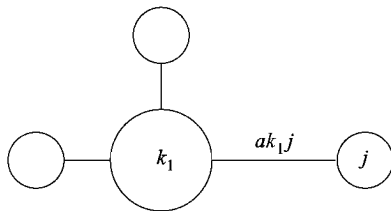


Рис. 1. Связь первого уровня между значимой и текущей рубриками

Каждой j -й рубрике ставится в соответствие пара чисел

$$j \rightarrow (S_j^{1k}, S_j^{2k}), \quad (k \neq j). \quad (22)$$

Пары упорядочиваются в порядке возрастания значений элементов.

Таким образом, устанавливается номер рубрики, имеющей наиболее тесную связь (наибольший вес связи S_j^{2k}) с значимой рубрикой.

На втором этапе устанавливается наличие связей второго уровня между текущей j -й ($j = 1, 2, \dots, n, j \neq k$) рубрикой и рубриками, имеющими связь первого уровня ($S_i^{1k} \neq 0, i = 1, 2, \dots, n, i \neq k, j$), (рис. 2).

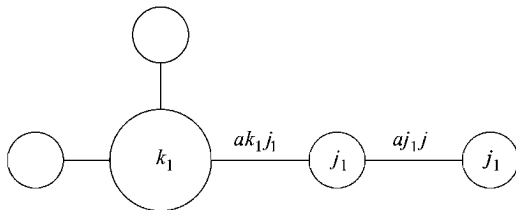


Рис. 2. Связь второго уровня между значимой и текущей рубриками

Для каждой j -й рубрики подсчитывается число связей второго уровня

$$S_j^{3k} = \sum_i b_{ij} \quad (i \neq j, j \neq k), \quad (23)$$

если $S_i^{1k} = 1$ и

$$S_j^{3k} = 0 \quad (i \neq j, j \neq k), \quad (24)$$

если $S_i^{1k} = 0$.

Подсчитывается сумма весов связей второго уровня

$$S_j^{4k} = \sum_i a_{ij} \quad (i \neq j, j \neq k), \quad (25)$$

если $S_i^{1k} = 1$,

$$S_j^{4k} = 0, \quad (26)$$

если $S_i^{1k} = 0$.

и средний вес связи второго уровня

$$P_j^{1k} = \frac{S_j^{4k}}{S_j^{3k}} \quad (27)$$

$$P_j^{1k} = 0, \quad (28)$$

если $S_i^{3k} = 0$.

На третьем этапе каждой рубрике ставится в соответствие набор из четырёх чисел

$$j \rightarrow (S_j^{1k}, S_j^{2k}, S_j^{3k}, P_j^{1k}), \quad (k \neq j, j = 1, 2, \dots, n). \quad (29)$$

Данные наборы упорядочиваются в порядке убывания числа связей S_i^{3k} , далее в порядке убывания среднего веса связи P_j^{1k} .

По данным наборам выбирается рубрика J_1 , имеющая наибольшее число связей с проблемно-ориентированной рубрикой,

$$S_{J_1} = \max_j (S_j^{1k} + S_j^{3k}) \quad (30)$$

и рубрика J_2 , имеющая наибольший средний вес связи,

$$P_{J_2} = \max_j \left(\frac{1}{2} (S_j^{2k} + P_j^{1k}) \right). \quad (31)$$

Среди рубрик, имеющих с значимой рубрикой связи первого либо второго уровня, выбирается пара рубрик J_3, J_4 с наибольшим весом связи, т. е. находится

$$a_{J_3 J_4} = \max_j \max_i a_{ij} \quad (32)$$

при выполнении одного из условий

$$\begin{aligned} S_j^{3k} S_i^{3k} &\neq 0, & S_j^{1k} S_i^{3k} &\neq 0, \\ S_j^{3k} S_i^{1k} &\neq 0, & S_j^{1k} S_i^{1k} &\neq 0. \end{aligned}$$

Таким образом, после проведения расчётных этапов для k -й значимой рубрики устанавливаются:

1. Рубрики, прямо связанные с значимой рубрикой, расположенные в порядке убывания веса связи.

2. Рубрики, непосредственно не связанные с значимой рубрикой (имеющие с ней опосредованную связь через прямо связанные рубрики), расположенные в порядке убывания числа связей и среднего веса связи.

3. Рубрики с наибольшим числом связей и наибольшим средним весом связи.

4. Пара (несколько пар) рубрик с наибольшим весом внутренней связи.

Для выявления и упорядочивания рубрик, связывающих две проблемные рубрики, обозначим: k_1, k_2 – номера значимых рубрик, j – номер текущей рубрики.

На первом этапе устанавливаются связи первого уровня текущей рубрики с каждой значимой рубрикой в отдельности.

Введём величины: $S_j^{1k_1}$ — наличие связи и $S_j^{2k_1}$ — вес связи первого уровня.

Прямым перебором по всем $j = 1, 2, \dots, n$, $j \neq k_1$ проверяется условие

$$b_{k_1 j} = 1. \quad (33)$$

Если условие (32) выполняется, между рубриками k_1 и j имеет место связь первого уровня и

$$S_j^{1k_1} = b_{k_1 j} = 1, \quad (34)$$

$$S_j^{2k_1} = a_{k_1 j}. \quad (35)$$

В случае невыполнения условия (32) связи первого уровня нет и

$$S_j^{1k_1} = 0, \quad S_j^{2k_1} = 0.$$

Таким образом, на первом этапе каждой рубрике ставится в соответствие пара чисел

$$j \rightarrow (S_j^{1k_1}, S_j^{2k_1}), \quad (k_1 \neq j). \quad (36)$$

Рубрики поочерёдно упорядочиваются в порядке убывания значений элементов пар. Первое место занимает рубрика, имеющая наибольший вес связи первого уровня.

Аналогично для значимой рубрики k_2 формируются и упорядочиваются пары

$$j \rightarrow (S_j^{1k_2}, S_j^{2k_2}), \quad (k_2 \neq j). \quad (37)$$

На втором этапе для текущей рубрики устанавливаются связи второго уровня. Под связью второго уровня здесь понимается связь текущей рубрики с рубриками, имеющими связь первого уровня. При этом должно выполняться хотя бы одно условие

$$S_i^{1k_1} = 1, \quad S_i^{1k_2} = 1 \quad (38)$$

Введём величины: $S_j^{3k_1}$ — число связей и $S_j^{4k_1}$ — вес связей второго уровня текущей рубрики с k_1 -й значимой рубрикой, $S_j^{3k_2}$ — число связей и $S_j^{4k_2}$ — вес связей второго уровня с k_2 -й значимой рубрикой.

Для каждого j ($j \neq k_1, k_2$, $j = 1, 2, \dots, n$) прямым перебором по $i = 1, 2, \dots, n$ ($i \neq j$) определяются значения

$$S_j^{3k_1} = \sum_i b_{ij}, \quad (39)$$

$$S_j^{4k_1} = \sum_i a_{ij} \quad (40)$$

при выполнении требования $S_i^{1k_1} = 1$, если это требование не выполняется

$$S_j^{3k_1} = 0, \quad S_j^{4k_1} = 0.$$

Определяется средний вес связи

$$P_j^{1k_1} = \frac{S_j^{4k_1}}{S_j^{3k_1}}, \quad \text{если } S_j^{3k_1} \neq 0 \quad (41)$$

и

$$P_j^{1k_1} = 0, \quad \text{если } S_j^{3k_1} = 0. \quad (42)$$

Аналогично для значимой рубрики k_2

$$S_j^{3k_2} = \sum_i b_{ij}, \quad (43)$$

$$S_j^{4k_2} = \sum_i a_{ij} \quad (44)$$

при выполнении требования $S_i^{1k_2} = 1$, если это требование не выполняется

$$S_j^{3k_2} = 0, \quad (45)$$

$$S_j^{4k_2} = 0. \quad (46)$$

Определяется средний вес связи

$$P_j^{1k_2} = \frac{S_j^{4k_2}}{S_j^{3k_2}}, \quad \text{если } S_j^{3k_2} \neq 0 \quad (47)$$

$$P_j^{1k_2} = 0, \quad \text{если } S_j^{3k_2} = 0. \quad (48)$$

Для текущей рубрики рассчитываются средний вес связей второго уровня

$$P_j^1 = \frac{1}{2} (P_j^{1k_1} + P_j^{1k_2}), \quad (49)$$

общее число связей второго уровня

$$S_j^3 = S_j^{3k_1} + S_j^{3k_2}, \quad (50)$$

общий вес связей второго уровня

$$S_j^4 = S_j^{4k_1} + S_j^{4k_2}. \quad (51)$$

На третьем этапе каждой текущей рубрике ставится в соответствие набор чисел

$$j \rightarrow (S_j^3, S_j^4, P_j^1), \quad (j \neq k_1, k_2, j = 1, 2, \dots, n). \quad (52)$$

Рубрики поочерёдно упорядочиваются в порядке убывания значений элементов пар. На первом месте находится рубрика, имеющая наибольшее число связей второго уровня. Если $\max(S_j^3) = 0$, связь между проблемными рубриками k_1, k_2 отсутствует.

По полученным наборам определяется рубрика J_1 с максимальной суммой весов

$$S_{J_1}^4 = \max_j \left(S_j^4 + S_j^{1k_1} + S_j^{1k_2} \right), \quad (j \neq k_1, k_2, j = 1, 2, \dots, n) \quad (53)$$

Определяется рубрика J_2 с максимальным средним весом связи

$$P_{J_2} = \max_j \left(\frac{1}{l} \left(P_j^1 + S_j^{2k_1} + S_j^{2k_2} \right) \right), \quad (j \neq k_1, k_2, j = 1, 2, \dots, n), \quad (54)$$

где l — число ненулевых слагаемых.

Определяется пара рубрик I и J , имеющих наибольший вес связи, следующим образом

$$a_{IJ} = \max_i a_{ij}. \quad (55)$$

При выполнении хотя бы одного из следующих условий

$$S_i^{1k_1} \neq 0, \quad S_i^{1k_2} \neq 0, \quad S_i^{3k_1} \neq 0, \quad S_i^{3k_2} \neq 0, \quad (56)$$

$$a_{IJ} = \max_j a_{ij}. \quad (57)$$

При выполнении хотя бы одного из следующих условий

$$S_j^{1k_1} \neq 0, \quad S_j^{1k_2} \neq 0, \quad S_j^{3k_1} \neq 0, \quad S_j^{3k_2} \neq 0. \quad (58)$$

После проведения расчётных этапов для пары значимых рубрик установлены:

1. Рубрики, прямо связанные с каждой значимой рубрикой отдельно, расположенные в порядке убывания веса связи.

2. Рубрики, непосредственно не связанные с проблемными (имеющие с ними опосредованную связь через прямо связанные рубрики), расположенные в порядке убывания числа, веса и среднего веса связи.

3. Рубрики с наибольшей суммой весов, средним весом связи.

4. Пара (несколько пар) рубрик с наибольшим весом внутренней связи.

2.2. Типы связей. Между проблемными рубриками k_1, k_2 существует связь первого типа (рис. 3), если для них выполняется условие

$$b_{k_1 k_2} \neq 0. \quad (59)$$

Вес такой связи равен $a_{k_1 k_2}$.

Межу проблемными рубриками существует связь второго типа (рис. 4), если существует рубрика (несколько рубрик), для которой выполняется условие

$$S_j^{1k_1} \cdot S_j^{1k_2} \neq 0, \quad (j \neq k_1, k_2, j = 1, 2, \dots, n). \quad (60)$$

Вес такой связи равен сумме весов $a_{k_1 j} + a_{k_2 j}$

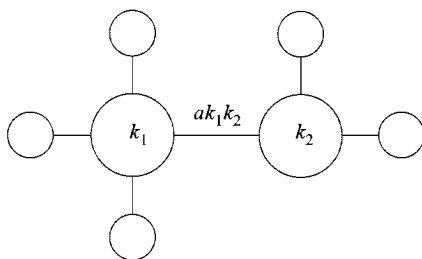


Рис. 3. Связь первого типа между проблемными рубриками

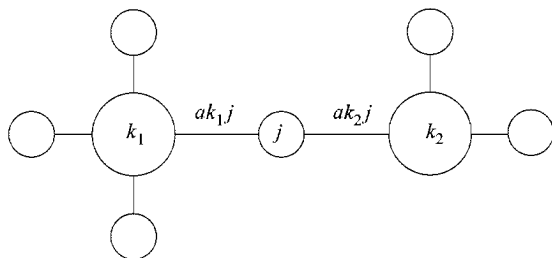


Рис. 4. Связь второго типа между проблемными рубриками

Между проблемными рубриками существует связь третьего типа (рис. 5), если существует пара рубрик с номерами j_1, j_2 (несколько пар рубрик), для которых выполняется условие

$$S_{j_1}^{1k_1} \cdot S_{j_2}^{1k_2} \cdot b_{j_1 j_2} \neq 0, \quad (61)$$

$$(j_1 \neq k_1, k_2, j_2 \neq k_1, k_2, j_1 \neq j_2, j_1, j_2 = 1, 2, \dots, n).$$

Между проблемными рубриками существует связь четвёртого типа (рис. 6), если существует тройка рубрик с номерами j_1, j_2, j (несколько троек), для которых выполняется условие

$$S_{j_1}^{1k_1} \cdot S_{j_2}^{1k_2} \cdot b_{j_1 j} \cdot b_{j j_2} \neq 0, \quad (62)$$

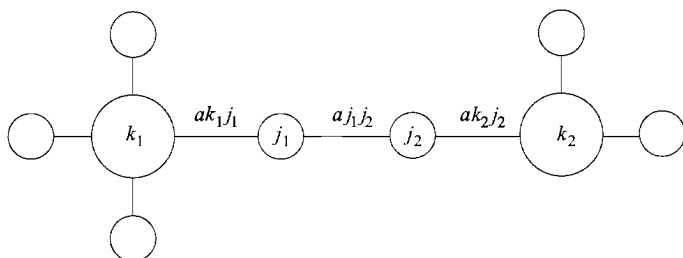


Рис. 5. Связь третьего типа между проблемными рубриками

$$(j_1 \neq k_1, k_2, j_2 \neq k_1, k_2, j \neq k_1, k_2, j_1 \neq j_2 \neq j, j_1, j_2, j = 1, 2, \dots, n) \quad (63)$$

Вес такой связи равен сумме весов

$$a_{k_1 j_1} + a_{j_1 j} + a_{j_2 j} + a_{k_2 j_2}. \quad (64)$$

Если между проблемными рубриками не установлено связи хотя бы одного типа проблемные рубрики считаются несвязанными.

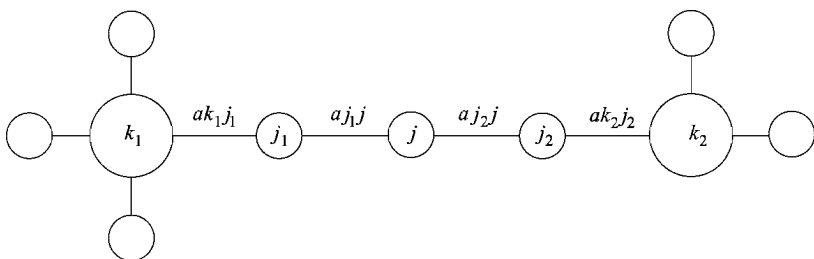


Рис. 6. Связь четвёртого типа между проблемными рубриками

Пусть k_1, k_2, k_m ($m \leq n$) — номера значимых рубрик. Для пар рубрик k_1 и k_3 , k_1 и k_4 и т. д. до k_{m-1} и k_m аналогично паре рубрик k_1 и k_2 выявляются все типы связей.

§ 3. Вероятностная модель определения области действия причинно-следственных связей

Существует целый ряд различных трактовок понятия причины. В одних случаях о событии А говорят как о причине события В, если осуществление А является достаточным условием последующего осуществления В. В других случаях причина рассматривается как необходимое условие наступления следствия. Ещё одна трактовка связана с пониманием причины как такого события, которое одновременно удовлетворяет как условию достаточности, так и условию необходимости. Такого понимания причины придерживался Дж. С. Милль, впервые чётко сформулировавший основные методы установления причинной зависимости в известном труде «Система логики» (1843 г.). Традиционно предполагалось, что причинная зависимость носит жёстко детерминированный характер. Однако, она может быть и не столь однозначной, например, когда связь между причиной А и следствием В осуществляется вероятностным образом. Через понятие вероятности, согласно Суппесу [2], причинная связь определяется следующим образом: событие А является причиной события В только если

1. $P(A)$, вероятность события А, больше 0;
2. А происходит раньше В;

3. $P(B|A)$, вероятность наступления B при наличии A , больше вероятности B ,

то есть

$$\frac{P(B|A)}{P(B)} > 1. \quad (65)$$

Каждое из перечисленных условий необходимо для осуществления причинной (каузальной) зависимости B от A . Взятые вместе они достаточны для того, чтобы считать событие A причиной B в силу приведённого здесь определения. Принимая данное определение за основу дальнейшего анализа причинно-следственных отношений, следует в первую очередь подвергнуть детальному разбору соотношение (65), играющее важную роль уже на предварительной стадии установления причинно-следственных связей, а именно, при выделении в пространстве признаков пары событий (A, B) области поиска возможных причин того или иного конкретного события. Решение задачи определения границ данной области (области Ω), начиная с которых действуют, либо прекращают своё действие каузальные закономерности, имеет немаловажное значение: в частности, сужение Ω способствует более эффективному применению известных методов или критериев установления причинной связи B с другими событиями $A_1, A_2, \dots, A_k$. Существенную роль при этом играет возможность характеризовать события A, B с различных точек зрения, что позволяет сформулировать ряд вариантов представления граничных условий, выделяющих Ω в различных пространствах признаков данных событий. Каждому из таких вариантов соответствуют свои характеристики (показатели) случайных событий A, B , рассматриваемые в качестве переменных либо параметров в формулировках условий, ограничивающих эту область. В результате реализуется возможность выявления различных факторов, от которых зависит установление каузальных отношений, определяются взаимосвязи между ними, что и позволяет учитывать характер воздействия этих факторов на конфигурацию границ рассматриваемой области.

Относительно A и B будем предполагать далее, что они образуют систему упорядоченных во времени случайных событий, каждое из которых может произойти, либо не произойти в интервале времени T наблюдения за развитием некоторого процесса, протекающего при заданном комплексе условий. Принимая такое наблюдение за единичное испытание, в качестве его возможных исходов будем иметь следующее множество взаимно-несовместимых ситуаций:

1. имеет место событие AB : событие A произошло и в пределах определённого промежутка времени τ ($\tau < T$) за ним последовало событие B ;
2. имеет место событие $A\bar{B}$: событие A произошло, однако, за время τ после его осуществления события B не последовало;

3. имеет место событие $\bar{A}B$: событие B произошло, однако, за период времени τ , предшествующий данному событию, событие A отсутствовало;

4. имеет место событие $\bar{A}\bar{B}$: как событие A , так и событие B за период наблюдения T отсутствовали.

В этом случае все вероятности, входящие в выражение (65), определяются по отношению к полному множеству несовместимых событий $U = \{AB, \bar{A}B, A\bar{B}, \bar{A}\bar{B}\}$ с заданной на нём вероятностной мерой. Нетрудно убедиться также, что и вероятности событий — элементов множества U могут быть однозначно определены как только $P(A)$, $P(B)$ и $P(B|A)$ становятся известными.

Следуя терминологии Бригга [3], будем называть далее отношение между событиями A , B , возникающее при выполнении условия (65), отношением R каузальной релевантности событий A , B , а само соотношение (65) условием каузальной релевантности. Отметим сначала одно свойство данного отношения, непосредственно вытекающее из (65). Ясно, что в силу определения условной вероятности, или теоремы умножения, согласно которой

$$P(B|A) = \frac{P(AB)}{P(A)}, \quad (66)$$

условие (65) может быть представлено в виде:

$$P(AB) > P(A) \cdot P(B) \quad (67)$$

События, для которых $P(AB) \neq P(A) \cdot P(B)$ по определению являются зависимыми и, таким образом, соотношение (67) свидетельствует о наличии частного вида зависимости между A и B . Так как из (67) в силу (66) при $P(AB) \neq 0$ тривиально следует (65), а также и

$$\frac{P(A|B)}{P(A)} > 1, \quad (68)$$

то отношение R является симметричным, и условие его установления может быть записано в любой из трёх эквивалентных форм — (65), (67) или (68). Ниже преимущественно используется симметричная форма данного условия (67).

Представим теперь каждое из событий A и B в виде суммы несовместимых событий из U , а именно:

$$A = AB \cup \bar{A}B, \quad \bar{A} = \bar{A}B \cup \bar{A}\bar{B} \quad (69)$$

и таким образом

$$P(A) = P(AB) + P(\bar{A}B) \quad \text{и} \quad P(B) = P(AB) + P(\bar{A}\bar{B}). \quad (70)$$

Учитывая (69) и (70), исходя из (67) получим:

$$P(AB) > (P(AB) + P(\bar{A}B)) \cdot (P(AB) + P(\bar{A}\bar{B})) =$$

$$= (P(AB))^2 + (P(\bar{A}B) + P(A\bar{B})) \cdot P(AB) + P(A\bar{B}) \cdot P(\bar{A}B). \quad (71)$$

Но $P(A\bar{B}) + P(\bar{A}B) = 1 - P(AB) - P(\bar{A}\bar{B})$, откуда следует, что

$$P(AB) > (P(AB))^2 + (1 - P(AB) - P(\bar{A}\bar{B})) \cdot P(AB) + P(A\bar{B}) \cdot P(\bar{A}B), \quad (72)$$

и в результате:

$$\frac{P(AB)}{P(A\bar{B})} > \frac{P(\bar{A}B)}{P(\bar{A}\bar{B})} \quad (73)$$

или, что то же:

$$P(AB) > \frac{P(A\bar{B}) \cdot P(\bar{A}B)}{1 - P(A \cup B)} \quad (74)$$

Из (74) сразу следует, что для любых событий A и B в сумме составляющих достоверное событие, при $P(A), P(B) < 1$ условие каузальной релевантности невыполнимо. Другие простейшие выводы из (73), (74) можно сделать, обратившись к графической интерпретации этих соотношений. На рис. 7 множеству событий U соответствует квадрат $KLQS$ со стороной, равной 1.

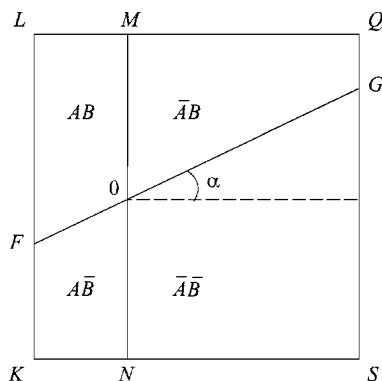


Рис. 7. Полная группа несовместимых событий $AB, A\bar{B}, \bar{A}B, \bar{A}\bar{B}$: распределение вероятностей при выполнении условия (67)

Событиям A и B отвечают фигуры, площади которых равны вероятностям $P(A)$ и $P(B)$: это, соответственно, прямоугольник $KLMN$ с основанием $P(A)$ и трапеция $FLQG$. Ясно, что если прямая FG параллельна основанию квадрата KS , то $FL = P(B)$, $P(AB) = P(A) \cdot P(B)$ и, следовательно, события A и B являются независимыми. В этом случае условия (73), (74) не выполняются. Ясно также, что данные соотношения справедливы лишь тогда, когда прямая FG образует положительный угол α с направлением KS . Изменяя вероятности событий A и B в интервале $0 < P(A), P(B) < 1$ или, в графической интерпретации, изменяя положение вертикали NM , позицию точки O в интервале (N, M) , а также величину угла α в пределах

$0 < \alpha \leq \frac{\pi}{2}$, нетрудно убедиться, что соотношения (73), (74) имеют место для любых значений $P(A)$ и $P(B)$, но только при определённых ограничениях на соответствующие величины $P(AB)$ и $P(A \cup B)$. В совместном изменении последних при варьировании значениями $P(A)$ и $P(B)$ прослеживается некоторая закономерность. Чтобы сделать её более прозрачной, количественно оценить зависимость, существующую между указанными величинами, достаточно перейти от переменных $P(AB)$, $P(A\bar{B})$, $P(\bar{A}B)$, $P(\bar{A}\bar{B})$ к другим, подходящим для данной цели характеристикам рассматриваемых событий.

Обращаясь к (67) и учитывая, что в силу известного соотношения между средним арифметическим и средним геометрическим:

$$P(A) \cdot P(B) \leq \left(\frac{P(A) + P(B)}{2} \right)^2, \quad (75)$$

получим неравенство:

$$P(AB) > \left(\frac{P(A) + P(B)}{2} \right)^2, \quad (76)$$

служащее достаточным условием выполнения (67). Так как (76) можно представить также в виде:

$$P(AB) > P(A) \cdot P(B) + \left(\frac{P(A) - P(B)}{2} \right)^2, \quad (77)$$

то для любых равновероятных событий, условия (76) и (67) эквивалентны.

Учитывая, что $P(A) + P(B) = P(A \cup B) + P(AB)$, получим из (76):

$$P(AB) > \frac{(P(A \cup B) + P(AB))^2}{4} \quad (78)$$

или, деля обе части (78) на $P(A \cup B) \neq 0$,

$$\frac{4P(AB)}{P(A \cup B)} > \left(1 + \frac{P(AB)}{P(A \cup B)} \right)^2 \cdot P(A \cup B) \quad (79)$$

Вводя показатель совместности событий A, B :

$$\nu = \frac{P(AB)}{P(A \cup B)}, \quad (80)$$

будем иметь, наконец, что

$$\frac{4\nu}{(1 + \nu)^2} > P(A \cup B) \quad (81)$$

является достаточным условием установления отношения R между событиями A и B .

Как видно из данного соотношения, начиная с любого произвольно малого $\nu > 0$, события A и B будут связаны отношением каузальной ре-

левантности если только вероятности этих событий достаточно малы. Напротив, события, в сумме близкие к достоверному, удовлетворяют соотношению (81) лишь при значениях ν весьма близких к 1. Являясь достаточным для установления отношения R , условие (81) равносильно (67) лишь при $P(A) = P(B)$, и не даёт отвечающей (67) картины соответствия между ν и $P(A \cup B)$ для неравновероятных событий. Различие состоит хотя бы в том, что область возможных значений ν задаётся соотношением $0 < \nu \leq 1$ лишь для событий A, B , таких, что $P(A) = P(B)$. Нетрудно видеть, однако, что при $P(A) \neq P(B)$ верхний предел области изменения ν в любом случае меньше 1, и его значение существенно зависит от величины $P(A)/P(B)$. Действительно, вводя в рассмотрение показатель ω :

$$\omega = \frac{P(A)}{P(B)}, \quad (82)$$

из определения ν найдём, что

$$\nu = \frac{P(AB)}{(1 + \omega) \cdot P(B) - P(AB)} \quad (83)$$

или

$$\frac{1}{\nu} = \frac{P(B)}{P(AB)} \cdot (1 + \omega) - 1. \quad (84)$$

Т.к. при произвольно фиксированном значении $\omega > 1$ величина $P(AB)/P(B)$ изменяется в пределах:

$$0 < \frac{P(AB)}{P(B)} \leq 1, \quad (85)$$

то минимальное значение $P(B)/P(AB)$ равно 1 и, в силу (84)

$$\frac{1}{\nu} \geq (1 + \omega) - 1 = \omega. \quad (86)$$

При $\omega < 1$ $P(AB)/P(B)$ изменяется лишь в пределах:

$$0 < \frac{P(AB)}{P(B)} \leq \frac{P(A)}{P(B)} = \omega, \quad (87)$$

минимальное значение $P(B)/P(A)$ равно $\frac{1}{\omega}$ и, следовательно,

$$\frac{1}{\nu} \geq \frac{1 + \omega}{\omega} - 1 = \frac{1}{\omega}. \quad (88)$$

Таким образом, область Ω' возможного изменения показателей ν, ω задаётся следующей системой неравенств:

$$\Omega' = \begin{cases} 0 < \nu \leq \frac{1}{\omega} & \text{при } \omega > 1; \\ 0 < \nu \leq 1 & \text{при } \omega = 1; \\ 0 < \nu \leq \omega & \text{при } 0 < \omega < 1. \end{cases} \quad (89)$$

Рассматривая ω в качестве независимого параметра, получим из (84) и соотношения $P(A \cup B) = P(A) + P(B) - P(AB)$, что

$$P(B) = \frac{1}{1 + \omega} (P(A \cup B) + P(AB)). \quad (90)$$

После преобразований получим

$$P(AB) > \frac{\omega}{(1 + \omega)^2} (P(AB) + P(A \cup B))^2, \quad (91)$$

в результате чего (67) сводится к следующему неравенству:

$$F(\nu, \omega) = \frac{\nu}{(1 + \nu)^2} \cdot \frac{(1 + \omega)^2}{\omega} > P(A \cup B), \quad (92)$$

где $(\nu, \omega) \in \Omega'$. При $\omega = 1$, т. е. в случае равновероятных A и B данное условие, очевидно, переходит в условие (81).

Здесь важно отметить, что при перестановке A и B правая часть (92) остаётся неизменной. По определению ν то же верно и относительно первого, зависящего от ν , сомножителя в левой части. Однако, в соответствии с (82), ω не обладает таким свойством: при $(P(A)/P(B)) = \omega$ $(P(B)/P(A)) = \omega' = 1/\omega$. Тем не менее, в силу симметрии отношения R , условие (92) должно быть инвариантным относительно перестановки A и B . На самом деле так и есть, т. к. в данном случае имеет место функциональное равенство:

$$f(\omega) = \frac{(\omega + 1)^2}{\omega} = \frac{(1/\omega + 1)^2}{1/\omega} = f(1/\omega), \quad (93)$$

и асимметрия, заложенная в определении ω компенсируется указанным свойством $f(\omega)$. Из вида функции $f(\omega)$ следует также, что граница области Ω' в части, не совпадающей с прямой $\nu = 0$, является линией уровня $F(\nu, \omega) = 1$ функции $F(\nu, \omega) = f(\omega)/f(\nu)$, которая при любом фиксированном значении $\omega > 0$ монотонно возрастает, изменяясь от 0 до 1 с увеличением ν в пределах, определяемых соотношениями (89).

Простое вычисление показывает, что при значениях ω , фиксируемых в интервале $1 < \omega < 1,5$, степень приближения (92) неравенством (81) достаточно высока: относительная погрешность приближения Σ не превышает здесь 0,042; при $\omega = 1,6$ $\Sigma = 0,056$; при $\omega = 2$, $\Sigma = 0,125$, а при $\omega = 4$ возрастает до 0,56. Таким образом, для произвольного ω_0 , такого, что $0,125 \cong 1/1,6 < \omega_0 < 1,6$ в интервале значений ν , определяемом согласно (91), использование (81) в качестве условия каузальной релевантности событий A и B представляется вполне оправданным.

Суммируя сказанное выше, можно сделать вывод, что введение показателей ν , ω позволило связать отношение R с другими наглядно интерпретируемыми отношениями между событиями A и B и на этой основе воссоздать достаточно полную картину такой взаимосвязи. В частности, была выявлена определённая зависимость между совме-

стимостью и масштабом вероятности событий A и B , для которых $R(A, B)$ имеет место. Незатронутым, однако, остался один вопрос, непосредственно примыкающий к данной теме: о возможности выделения в Ω событий A, B степень зависимости которых регламентирована определёнными условиями: например, не ниже (или не выше) некоторого заданного уровня. Решение этого вопроса требует введения ещё одного параметра — k , ограничивающего степень взаимозависимости A и B таким образом, что:

$$\frac{P(B|A)}{P(B)} > k (k \geq 1), \quad (94)$$

$$P(AB) > \frac{P(A\bar{B}) \cdot P(\bar{A}B)}{\frac{1}{k} - P(A \cup B)} \quad (95)$$

и

$$\frac{v}{k(v+1)^2} \cdot \frac{(1+\omega)^2}{\omega} > P(A \cup B). \quad (96)$$

Таким образом, в области Ω сильная зависимость ($k \gg 1$) имеет место лишь для маловероятных событий A и B . На рис. 8 представлена Ω в случае равновероятных событий: область ограничивают ось v и кривая L_1 , отвечающая значению $k = 1$.

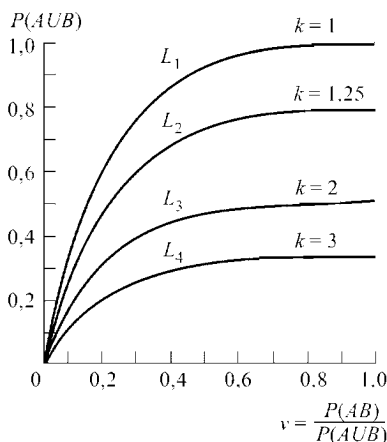


Рис. 8. Области, степень взаимосвязи равновероятных событий в которых удовлетворяет заданным ограничениям: $k \geq 1$; $k \geq 1.25$; $k \geq 2$; $k \geq 3$

Линии L_2, L_3, L_4 выделяют подобласти Ω , в которых степень зависимости между A и B превышает, соответственно 1,25; 2 и 3. Используя аналогичные графики, где семейство кривых, параметрически зависящих от k и ω представлено более полно, можно легко определять границы, начиная с которых степень взаимозависимости событий A и B не меньше (либо не больше) некоторого заданного уровня. Так,

например, в представленном на рис. 8 случае ($\omega = 1$), условие (96) даёт, что для любых A и B при $P(A \cup B) \leq 0,1$ и $\nu \geq 0,1$ степень зависимости не меньше 3. При $P(A \cup B) > 0,7$ для $\nu < 0,3$ зависимость между A и B вообще отсутствует, а при $P(A \cup B) > 0,9$ хотя и может иметь место, но лишь при значениях ν , достаточно близких к 1, да и то слабая, с $k < 1,25$.

Для произвольных значений $\omega > 0$ конфигурацию области Ω_k , выделяемой из Ω при любом фиксированном $k > 1$, проще всего представить соотношением (94), рассматривая Ω_k в пространстве признаков $P(A)$, $P(B)$ и $P(AB)$. Предварительно следует, однако, определить область допустимых значений $P(A)$, $P(B)$, $P(AB)$, выделяемую в данном пространстве системой ограничений, обусловленных общими свойствами вероятной меры. Применительно к вероятностям несовместимых событий из U эти ограничения выглядят элементарно:

$$\begin{aligned} \text{а) } & 0 \leq P(AB), P(\overline{A\overline{B}}), P(\overline{A}B), P(\overline{A}\overline{B}), \\ \text{б) } & P(AB) + P(\overline{A\overline{B}}) + P(\overline{A}B) + P(\overline{A}\overline{B}) = 1. \end{aligned} \quad (97)$$

Переформулируя (97 а, б) в терминах вероятностей событий A , B и AB , получим, что

$$0 \leq P(\overline{A\overline{B}}) = (P(\overline{A\overline{B}}) + P(AB)) - P(AB) = P(A) - P(AB)$$

и, значит,

$$0 \leq (P(AB) \leq P(A).$$

Аналогичным образом имеем

$$0 \leq (P(AB) \leq P(B).$$

Наконец, в соответствии с (97 б)

$$\begin{aligned} P(\overline{A\overline{B}}) + P(AB) + P(\overline{A}B) &= (P(\overline{A\overline{B}}) + P(AB)) + (P(\overline{A}B) + P(AB)) - \\ &- P(AB) = P(A) + P(B) - P(AB) \leq 1. \end{aligned} \quad (98)$$

В результате, обозначив для упрощения дальнейших выкладок $P(A)$, $P(B)$, $P(AB)$ соответственно через x , y , z , полученную систему общих ограничений можно записать как

$$\begin{cases} 0 \leq z \leq x, \\ 0 \leq z \leq y, \\ x + y - z \leq 1. \end{cases} \quad (99)$$

Совместно с условием (99), которое в новых обозначениях имеет вид

$$z > kxy, \quad (100)$$

система (99) определяет область Ω_k в трёхмерном пространстве координат x, y, z .

координатную плоскость Oxy гиперболу

$$\frac{(x')^2}{\left(\sqrt{\frac{2h}{k}}\right)^2} - \frac{(y')^2}{\left(\sqrt{\frac{2h}{k}}\right)^2} = 1 \quad (104)$$

с действительной полуосью, равной $\sqrt{\frac{2h}{k}}$ и асимптотами $y' = \pm x'$, совпадающими с координатными осями x, y старой системы. В той же проекции сечение пирамиды $CAOM$ даёт равнобедренный прямоугольный треугольник FGH с координатами вершин $F(h, 1)$, $G(h, h)$ и $H(1, h)$. Таким образом, множество точек Ω_k , принадлежащих секущей плоскости $z = h$, имеет своей проекцией на плоскость Oxy область $\Omega_k^{(h)}$, ограниченную сторонами прямого угла FGH и отрезком гиперболы (104), заключённым между точками K и L (рис. 10).

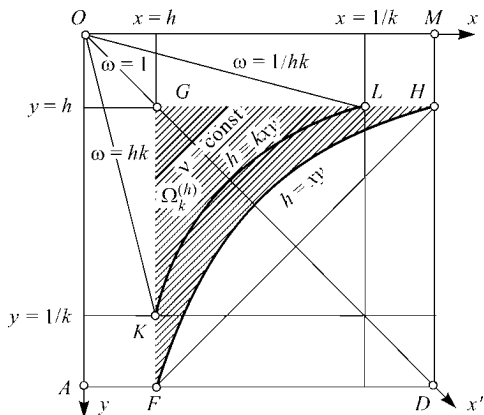


Рис. 10. Проекция сечения области Ω плоскостью $z = h$ на координатную плоскость Oxy . $h = 0,2$; $k = 1,3$.

Ясно, что $\Omega_k^{(h)}$ не пуста лишь в том случае, когда действительная ось гиперболы превышает расстояние точки G до начала координат, т. е. при условии, что

$$\sqrt{\frac{2h}{k}} > h\sqrt{2} \quad (105)$$

или

$$h < \frac{1}{k}. \quad (106)$$

Так как точки K и L пересечения гиперболы с прямыми $x = h$ и $y = h$ имеют, соответственно, координаты $(h, 1/k)$, и, по определению, $k \geq 1$, то при любых допустимых значениях h и k , удовлетворяющих условию (106), эти точки лежат на сторонах FG и GH треугольника

FGH . В случае, когда $h = \frac{1}{k}$, K и L сливаются с вершиной гиперболы и $\Omega_k^{(h)} = \emptyset$. Как видно из рис. 7, вероятности событий A и B могут принимать значения в промежутке $[h, 1/k]$. При этом каждому фиксированному в данном промежутке значению $P(A)$ соответствует своё множество допустимых значений $P(B) : h \leq P(B) < \frac{h}{k \cdot P(A)}$. При минимальном значении $P(A)$, равном h , это множество определяется условием $h \leq P(B) < \frac{1}{k}$, а при $P(A) = 1/k$ становится пустым. Перестановка A и B в указанных соотношениях приводит к аналогичному результату.

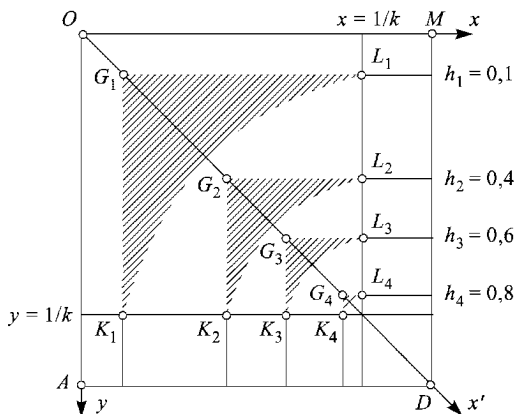


Рис. 11. Проекция сечений области Ω_k на координатную плоскость Oxy : $h = 0,1; 0,4; 0,6; 0,8$. $k = 1,2$

Рис. 11 изображает проекции на плоскость Oxy сечений области Ω параллельными плоскостями $z = h_i$ ($i = 1, 2, \dots$). Взятые вместе, они дают достаточно ясное представление о данной области трёхмерного пространства признаков $P(A)$, $P(B)$ и $P(AB)$. Характерные особенности конфигурации Ω обусловлены тем, что она является результатом пересечения двух областей, одна из которых, определяемая условием $z > xy$ ($x, y, z > 0$), расширяется с увеличением h , другая — область допустимых значений $P(A)$, $P(B)$, $P(AB)$ с увеличением h сжимается, вырождаясь в единственную точку $C(1, 1, 1)$. При этом произвольно фиксированной точке (x_0, y_0) проекции Ω на Oxy соответствует, вообще говоря, множество значений h и k , взаимное изменение которых удовлетворяет требованию $h/k = \text{const} < 1$. Это вполне естественно в «трёхмерной» модели взаимосвязи событий A и B , где в качестве начальных данных используются не только значения вероятности $P(AB)$, но и значения вероятностей самих событий $P(A)$ и $P(B)$.

Зависимость показателей ν и ω от параметра k также легко проследить, используя горизонтальные сечения. Нетрудно убедиться, что

в рассматриваемой проекции на плоскость Oxy линии уровня ν и ω имеют чрезвычайно простой вид: в случае ω — это отрезки прямых линий пучка прямых $y = Cx$, в случае ν — отрезки параллельных прямых вида $y + x = C$, где $C > 0$ (рис. 10). На каждой прямой $y = Cx$, имеющей хотя бы одну общую точку с $\Omega_k^{(h)}$, ω принимает значение, равное $1/C$. На каждой прямой семейства $y + x = C$, $\nu = h/(C - h)$. Отсюда, в частности, следует что, поскольку координаты точек K , L и G равны соответственно $(h, 1/k)$, $(1/k, h)$, (h, h) , значения ν и ω изменяются в пределах:

$$hk \leq \omega \leq \frac{1}{hk}. \quad (107)$$

и

$$hk \leq \nu \leq 1. \quad (108)$$

Завершая анализ взаимосвязи на уровне рассмотрения отдельных пар событий, следует отметить одну особенность показателя ω , определяющую его роль при поиске возможных причин заданного события B . Полагая значения h и k фиксированными, для вероятности события $B \supset A$, будем иметь:

$$\overline{P(B \supset A)} = P(A \cup \bar{B}) = P(\bar{A}B) = 1 - P(\bar{A}B) = 1 - (P(B) - h). \quad (109)$$

Тогда для произвольных пар событий (A_1, B_1) , (A_2, B_2) , вероятности которых удовлетворяют условию:

$$k_1 = \frac{h}{P(A_1)P(B_1)}, \quad k_2 = \frac{h}{P(A_2)P(B_2)}, \quad k_1 = k_2 = \bar{k} > k \quad (110)$$

выполняется равенство

$$\frac{P(B_1)}{P(B_2)} = \sqrt{\frac{\omega_2}{\omega_1}} \quad (111)$$

здесь $\omega_i = P(A_i)/P(B_i)$, $i = 1, 2$, и при $\omega_2 > \omega_1$ $P(B_1) > P(B_2)$. Отсюда и из (109) сразу следует, что на любой кривой вида

$$h = \bar{k}xy, \quad (112)$$

проходящей через $\Omega_k^{(h)}$, вероятность $P(B \supset A)$ является монотонно возрастающей функцией ω

$$P(B \supset A) = 1 - \left(\sqrt{\frac{h}{k\omega}} - h \right); \quad (113)$$

в точках пересечения K , L кривой (112) с границей области $\Omega_k^{(h)}$ эта функция принимает, соответственно, минимальное значение, равное $1 - \frac{1}{\bar{k}} + h$, и максимальное значение 1. Равенство $P(B \supset A) = 1$ означает, однако, что осуществление A является необходимым условием свершения события B . Поэтому можно сказать, что наряду с $P(B \supset A)$, показатель ω характеризует «степень необходимости» A для свершения B .

Обращаясь к рассмотрению n событий $(A, C_1, C_2, \dots, C_{n-2}, B)$ необходимо иметь в виду, что отношение каузальной релевантности R не является транзитивным, то есть из выполнения системы неравенств

$$\begin{cases} P(AC) > P(A)P(C), \\ P(CB) > P(C)P(B) \end{cases} \quad (114)$$

не следует $P(AB) > P(A)P(B)$. В справедливости данного утверждения можно убедиться хотя бы на простом примере, взяв полное множество несовместимых событий $ABC, AB\bar{C}, A\bar{B}C, A\bar{B}\bar{C}, \bar{A}BC, \bar{A}B\bar{C}, \bar{A}\bar{B}C, \bar{A}\bar{B}\bar{C}$, из которых отличную от 0 вероятность имеют лишь события $A\bar{B}C, \bar{A}BC, A\bar{B}\bar{C}, \bar{A}\bar{B}\bar{C}$. В этом случае $P(A), P(B), P(C) > 0$, однако $P(AB) = 0$. Трудность, связанная с отсутствием между A и B отношения R , устраняется в модели взаимосвязи, где каждое из событий $C_1, C_2, \dots, C_{n-2}, B$, порождается всем множеством предшествующих ему событий в целом, а именно: C_1 порождается событием A ; C_2 — совместным осуществлением A и C_1 ; C_3 порождается событиями A, C_1 и C_2 и т. д. В такой модели необходимым условием каузальной связи событий A и B , опосредованной свершением $C_1, C_2, \dots, C_{n-2}$, является выполнение системы неравенств

$$\frac{P(C_1|A)}{P(C_1)} > 1, \quad \frac{P(C_2|AC_1)}{P(C_2)} > 1, \quad \frac{P(B|AC_1C_2\dots C_{n-2})}{P(B)} > 1. \quad (115)$$

Ограничиваясь, в целях простоты изложения, рассмотрением лишь троек событий A, C, B будем иметь

$$\frac{P(C|A)}{P(C)} > 1 \quad \text{и} \quad \frac{P(B|AC)}{P(B)} > 1, \quad (116)$$

или, в силу определения условной вероятности, систему:

$$\begin{cases} P(AC) > P(A)P(C), \\ P(ACB) > P(AC)P(B). \end{cases} \quad (117)$$

Непосредственным следствием (67) является соотношение для пар событий A, B :

$$P(ABC) > P(A)P(B)P(C), \quad (118)$$

которое является лишь необходимым условием выполнения (115).

Приведённые здесь соображения позволяют сформулировать алгоритм выделения возможных причин исследуемого события B из некоего множества событий M , на фоне которых B имеет место. В основных, наиболее существенных чертах, этот алгоритм можно представить следующим образом.

Выделение из заданного множества $M = \{A_1, A_2, \dots, A_i, \dots, A_N\}$ подмножества $M_1 = \{A_{i_1}, A_{i_2}, \dots, A_{i_Q}\}$ событий, связанных с B отношением каузальной релевантности R , и последующая оценка A_{i_q} ($q = 1, 2, \dots, Q$), как возможных причин B , по совокупности показателей k, ω .

С этой целью для каждого $A_i \in M$ выполняются следующие действия:

— определение значений показателей k, ω :

$$k_i = \frac{h_i}{P(A_i)P(B)}, \quad \omega_i = \frac{P(A_i)}{P(B)}, \quad (119)$$

здесь $h_i = p(A_i B)$;

— последовательная проверка условий:

$$k_i \geq k_0, \quad k_i h_i \leq \omega_i \leq \frac{1}{k_i h_i}, \quad (120)$$

где $k_0 \geq 1$ — некоторая константа, пороговое значение показателя k . В случае выполнения указанных условий — включение A_i в состав множества M_1 ;

— оценка элементов M_1 по совокупности показателей k, ω с помощью многокритериальных методов принятия решений.

Формирование множества M_2 пар событий (A_q, A_s) , для которых событие $A_q A_s$, состоящее в совместном осуществлении A_q и A_s , связано с B отношением R . Оценка всех таких событий по совокупности показателей k, ω .

С этой целью для каждой пары элементов (A_{k_i}, A_{k_j}) множества событий $\{A_{k_1}, A_{k_2}, \dots, A_{k_s}\}$, не связанных с B отношением R , производятся следующие действия:

— определение значений показателей k, ω :

$$k_{ij} = \frac{h_{ij}}{P(A_{k_i})P(A_{k_j})}, \quad \omega_{ij} = \frac{P(A_{k_i})}{P(A_{k_j})}, \quad (121)$$

где $h_{ij} = P(A_{k_i} A_{k_j})$;

— проверка условий:

$$k_{ij} \geq k_0; \quad k_{ij} h_{ij} \leq \omega_{ij} \leq \frac{1}{k_{ij} h_{ij}}. \quad (122)$$

В случае выполнения указанных условий:

— определение значений следующей пары показателей k', ω' :

$$k'_{ij} = \frac{h'_{ij}}{P(A_{k_i} A_{k_j})P(B)}, \quad \omega'_{ij} = \frac{P(A_{k_i} A_{k_j})}{P(B)}, \quad (123)$$

где $h'_{ij} = P(A_{k_i} A_{k_j} B)$;

— проверка условий:

$$k'_{ij} \geq k_0; \quad k'_{ij} h'_{ij} \leq \omega'_{ij} \leq \frac{1}{k'_{ij} h'_{ij}} \quad (124)$$

при выполнении которых производится включение пары (A_{k_i}, A_{k_j}) во множество M_2 ;

— для каждой пары $(A_{k_i}, A_{k_j}) \in M_2$ оценка события $A_{k_i} A_{k_j}$ как возможной причины B , по совокупности показателей k, ω .

Полученные оценки могут использоваться далее как для непосредственного сужения множества возможных причин события B , так и для упорядочения элементов этого множества (по некоторому обобщённо-му критерию), предвещающего применение известных методов поиска причины B в заданном множестве событий. Последовательности, содержащие более трёх событий, рассматриваются аналогичным образом.

До сих пор условие (65) рассматривалось здесь в рамках схемы случайных событий. Покажем теперь, каким образом можно интерпретировать данное условие в терминах характеристик соответствующих случайных величин. С этой целью, возвращаясь к рассмотренному ранее вероятностному эксперименту, в качестве возможных исходов единичного испытания будем фиксировать значения случайных величин ξ_1 и ξ_2 , полагая, что:

$\xi_1 = 1, \xi_2 = 1$, если имеет место событие AB ;

$\xi_1 = 1, \xi_2 = 0$, если при наличии A событие B отсутствует, то есть при $A\bar{B}$;

$\xi_1 = 0, \xi_2 = 1$, если имеет место событие $\bar{A}B$;

$\xi_1 = 0, \xi_2 = 0$ при $\bar{A}\bar{B}$.

Зная значения $P(AB), P(A\bar{B}), P(\bar{A}B), P(\bar{A}\bar{B})$, однозначно определяющие совместное распределение вероятностей ξ_1 и ξ_2 , найдём математическое ожидание $M(\xi_1), M(\xi_2)$ этих величин и их корреляционный момент $K(\xi_1, \xi_2)$. Имеем:

$$\begin{aligned} M(\xi_1) &= p(\xi_1 = 1) \cdot 1 + p(\xi_1 = 0) \cdot 0 = \\ &= p((\xi_1 = 1)(\xi_2 = 1)) + p((\xi_1 = 1)(\xi_2 = 0)) = \\ &= P(AB) + P(A\bar{B}) = P(A), \quad (125) \end{aligned}$$

и, аналогично, $M(\xi_2) = P(B)$.

Таким образом,

$$\begin{aligned} K(\xi_1, \xi_2) &= M((\xi_1 - M(\xi_1))(\xi_2 - M(\xi_2))) = \\ &= (1 - P(A))(1 - P(B))P(\bar{A}\bar{B}) + (1 - P(A))(0 - P(B))P(\bar{A}B) + \\ &+ (0 - P(A))(1 - P(B))P(A\bar{B}) + (0 - P(A))(0 - P(B))P(AB) = \\ &= P(\bar{A}\bar{B}) - P(A)P(\bar{A}\bar{B}) - P(B)P(\bar{A}\bar{B}) + \\ &+ P(A)P(B)P(\bar{A}\bar{B}) - P(B)P(A\bar{B}) + P(A)P(B)P(A\bar{B}) - \\ &- P(A)P(\bar{A}B) + P(A)P(B)P(\bar{A}B) + P(A)P(B)P(AB) = \\ &= P(\bar{A}\bar{B}) - P(A)(P(\bar{A}\bar{B}) + P(\bar{A}B)) - P(B)(P(A\bar{B}) + P(AB)) + \\ &+ P(A)P(B) = P(\bar{A}\bar{B}) - P(A)P(B). \quad (126) \end{aligned}$$

Сопоставление полученного результата с условием (67) свидетельствует о том, что данное условие эквивалентно требованию положительности коэффициента корреляции случайных величин ξ_1 и ξ_2 .

§ 4. Модель поиска исходной информации для анализа взаимосвязей проблем

Разработка эффективной стратегии поиска информации по взаимосвязанным проблемам предполагает решение ряда вопросов, к числу которых в первую очередь относятся:

- установление области поиска информации по проблемам, взаимосвязь между которыми является предметом рассмотрения;
- определение и классификация ситуаций поиска информации по взаимосвязанным проблемам; разработка способов формализованного описания и визуального представления классов ситуаций поиска;
- организация интерфейса с пользователем; разработка способов и форм выдачи информации эксперту-аналитику.

Построение формализованной модели поиска информации о взаимосвязи проблем P_i, P_j в заданном массиве M_d рубрицированных документов базируется на определённой системе отношений, существующих между документами $d_n \in M_d$ ($n = 1, 2, \dots, N$), затронутыми в них проблемами и рубриками, которым данные документы соотносятся. В рассматриваемой модели поиска эти отношения учитываются опосредованно, через фиксацию реляционных свойств документов $d_n \in M_d$. Основой для задания реляционных свойств являются индексы (поисковые образы) документов M_d , фиксирующие вхождения элементов множества потенциальных ссылок на проблемы $P_i \in M_p$ в тексты документов M_d , а также известное распределение этих документов по рубрикам рубрикатора R .

Слово «свойство» используется здесь в более широком смысле, чем при обычном его употреблении. Так, например, к категории свойств относятся численные показатели, а также такие характеристики документов, как: «содержит вхождение какой-либо ссылки на P_i », «содержит вхождение термина „наркобизнес“» и т. п. Отличительной особенностью реляционных свойств является то, что в их языковой форме — положительном одноместном предикате, содержится хотя бы одна предметная константа (имя какого-либо объекта) или, по меньшей мере, один квантор. В рамках принятого здесь подхода такого рода константами в выражениях базовых реляционных свойств, характеризующих документы M_d в аспекте их соотнесённости проблемам из M_p , могут быть лишь ссылки (элементы ссылок) на эти проблемы, а также имена рубрик r_i ($i = 1, 2, \dots, H$) рубрикатора R .

Соответственно, любой из документов $X \in M_d$ в рассматриваемой модели характеризуется наличием либо отсутствием каждого из множества реляционных свойств:

$R_k(X) \equiv$ « X принадлежит рубрике r_k »,

$S_k(X) \equiv$ « X содержит хотя бы одну ссылку на проблему P_k »,

где $k = 1, 2, \dots, H$.

Для анализа лишь непосредственных взаимосвязей, существующих между двумя проблемами, например, P_i и P_j необходим, однако, учёт только следующих свойств:

$$R_i(X), R_j(X), S_i(X), S_j(X). \quad (127)$$

Таким образом, по значениям признаков, определяющих наличие либо отсутствие реляционных свойств (127), множество документов M_d может быть разделено на 16 непересекающихся классов, выделяемых следующими описаниями состояний образующих их документов:

1	$R_i(X) \& R_j(X) \& S_i(X) \& S_j(X)$
2	$R_i(X) \& R_j(X) \& S_i(X) \& \neg S_j(X)$
3	$R_i(X) \& R_j(X) \& \neg S_i(X) \& S_j(X)$
4	$R_i(X) \& R_j(X) \& \neg S_i(X) \& \neg S_j(X)$
5	$R_i(X) \& \neg R_j(X) \& S_i(X) \& S_j(X)$
6	$R_i(X) \& \neg R_j(X) \& S_i(X) \& \neg S_j(X)$
7	$R_i(X) \& \neg R_j(X) \& \neg S_i(X) \& S_j(X)$
8	$R_i(X) \& \neg R_j(X) \& \neg S_i(X) \& \neg S_j(X)$
9	$\neg R_i(X) \& R_j(X) \& S_i(X) \& S_j(X)$
10	$\neg R_i(X) \& R_j(X) \& S_i(X) \& \neg S_j(X)$
11	$\neg R_i(X) \& R_j(X) \& \neg S_i(X) \& S_j(X)$
12	$\neg R_i(X) \& R_j(X) \& \neg S_i(X) \& \neg S_j(X)$
13	$\neg R_i(X) \& \neg R_j(X) \& S_i(X) \& S_j(X)$
14	$\neg R_i(X) \& \neg R_j(X) \& S_i(X) \& \neg S_j(X)$
15	$\neg R_i(X) \& \neg R_j(X) \& \neg S_i(X) \& S_j(X)$
16	$\neg R_i(X) \& \neg R_j(X) \& \neg S_i(X) \& \neg S_j(X)$

(здесь $\neg$ — знак логического отрицания)

Если принять во внимание сделанное ранее предположение, согласно которому каждый из документов, относящихся к рубрике r , содержит хотя бы одну ссылку на соответствующую проблему P , документы, принадлежащие классам 2, 3, 4, 7, 8, 10, 12 из рассмотрения исключаются. В составе модели не рассматриваются также документы классов 13–16, не принадлежащие обеим рубрикам r_i и r_j одновременно. То, что документы этих классов не могут содержать непосредственных свидетельств существования (либо отсутствия) связи между проблемами P_i и P_j , достаточно очевидно. Например, в случае класса 13, для этого имеется следующее основание. Ссылки на проблемы P_i и P_j в составе документов данного класса употребляются не в контексте

обсуждения хотя бы одной из них, но при рассмотрении каких-то других проблем, и, следовательно, в непосредственной связи именно с этими другими проблемами. Таким образом, совместное вхождение ссылок на P_i и P_j в этом случае если и может свидетельствовать о наличии связи P_i и P_j , то разве что опосредованной, через иные проблемы, рассматриваемые в этих документах.

Переходя к рассмотрению остальных документов, нетрудно показать, что документы классов 1, 5 и 9 могут либо содержать свидетельства существования связи между P_i и P_j , либо свидетельствовать об отсутствии такой связи. Действительно, когда в тексте документа класса 1 существует хотя бы один фрагмент, посвящённый P_i либо P_j , и содержащий ссылку на другую проблему этой пары, этот документ свидетельствует о наличии связи P_i и P_j . В противном случае, когда ни один из фрагментов текста, посвящённых P_i либо P_j не содержит ссылки на другую проблему данной пары, документ свидетельствует об отсутствии взаимосвязи P_i и P_j . Для документов, принадлежащих классам 5 и 9 имеет место аналогичная ситуация. Что же касается документов классов 6 и 11, то они свидетельствуют лишь об отсутствии связи между этими двумя проблемами.

Таким образом, произвольная ситуация поиска, реализуемая в данной модели, может быть задана пятью параметрами $n(M_i)$, $i = 1, 2, \dots, 5$, (рис. 12) где $n(M_i)$ — число элементов множества M_i и

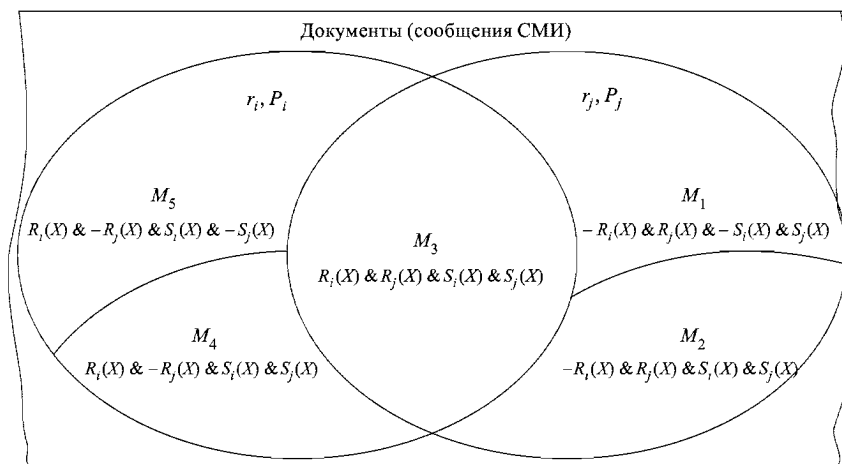
$$\begin{aligned} M_1 &= \{X \mid \neg R_i(X) \ \& \ R_j(X) \ \& \ \neg S_i(X) \ \& \ S_j(X)\}, \\ M_2 &= \{X \mid \neg R_i(X) \ \& \ R_j(X) \ \& \ S_i(X) \ \& \ S_j(X)\}, \\ M_3 &= \{X \mid R_i(X) \ \& \ R_j(X) \ \& \ S_i(X) \ \& \ S_j(X)\}, \\ M_4 &= \{X \mid R_i(X) \ \& \ \neg R_j(X) \ \& \ S_i(X) \ \& \ S_j(X)\}, \\ M_5 &= \{X \mid R_i(X) \ \& \ \neg R_j(X) \ \& \ S_i(X) \ \& \ \neg S_j(X)\}. \end{aligned}$$

При этом областью поиска возможных свидетельств существования связи P_i и P_j служит объединение множеств $M_2 \cup M_3 \cup M_4$. Одновременно объединение $M_1 \cup M_5$ является множеством документов, свидетельствующих (в рамках используемого формализованного языка) только лишь об отсутствии связи между этими проблемами. Отсюда следует, что в качестве численных характеристик ситуаций поиска целесообразно рассматривать следующие показатели:

$$\nu_{ij} = \frac{n(M_2) + n(M_3) + n(M_4)}{\sum_{i=1}^5 n(M_i)} \quad (128)$$

— верхнюю границу интервала возможных значений вероятности выбора документа, содержащего свидетельство взаимосвязи проблем P_i и P_j , из множества M_d и

$$\mu_{ij} = \frac{n(M_3)}{\sum_{i=2}^4 n(M_i)} \quad (129)$$



r_i – i -ая рубрика
 M_i – ($i = 1, \dots, 5$) подмножества документов объединения рубрик
 $n(M_i)$ – число документов в подмножестве M_i
 $R_i(X)$ – признак принадлежности документа рубрике r_i
 $S_i(X)$ – признак наличия в документе ссылки на проблему P_i

Рис. 12. Модель поиска информации по взаимосвязи проблем P_i, P_j

— долю документов, принадлежащих пересечению рубрик r_i, r_j от общего числа документов объединения этих рубрик, содержащих ссылки на каждую из проблем P_i и P_j одновременно.

Как видно из рис. 12, показатель γ_{ij} можно содержательно интерпретировать так же, как меру совместимости понятий: «документ X , принадлежащий $r_i \cup r_j$ и содержащий ссылку на проблему P_j » и «документ X , принадлежащий $r_i \cup r_j$ и содержащий ссылку на проблему P_i », объёмами которых являются объединения $(M_1 \cup M_2 \cup M_3 \cup M_4)$ и $(M_2 \cup M_3 \cup M_4 \cup M_5)$.

Аналогичную интерпретацию имеет и показатель μ_{ij} , рассматриваемый в качестве меры совместимости понятий: «документ X , принадлежащий r_i и содержащий ссылку на проблему P_j » и «документ X , принадлежащий r_j и содержащий ссылку на проблему P_i », с объёмами $M_3 \cup M_4$ и $M_2 \cup M_3$ соответственно. Так как

$$(M_1 \cup M_2 \cup M_3 \cup M_4) \cup (M_2 \cup M_3 \cup M_4 \cup M_5) = r_i \cup r_j$$

и

$$(M_1 \cup M_2 \cup M_3 \cup M_4) \cap (M_2 \cup M_3 \cup M_4 \cup M_5) = M_2 \cup M_3 \cup M_4 = (M_2 \cup M_3) \cup (M_3 \cup M_4),$$

то при условии, что $(M_2 \cup M_3 \cup M_4) \neq \emptyset$, когда наличие свидетельств существования взаимосвязи P_i и P_j не исключается, множество си-

туаций поиска можно разделить на классы по двум независимым основаниям деления: видам совместимости понятий (равнозначность, перекрещивание, подчинение) в указанных выше парах.

Нетрудно показать, что при таком делении каждому классу ситуаций поиска соответствует единственное, из общего числа 2^5 возможных, разбиение множества $\{M_1, M_2, M_3, M_4, M_5\}$ на два подмножества — пустых и непустых множеств документов M_i ($i = 1, \dots, 5$). Отнесение конкретной ситуации поиска соответствующему классу позволяет сопоставить ей обобщённый визуальный образ, который вместе с показателями γ_{ij} , μ_{ij} обеспечивает целостное представление данной ситуации и может служить ориентиром при поиске информации по взаимосвязанным проблемам.

Список литературы

1. Бочаров В.А., Маркин В.И. Основы логики. Учебник. — М.: ИНФРА-М, 2003. — 296 с.
2. Филлмор. Фреймы и семантика понимания. Сб. Новое в зарубежной лингвистике, вып. XXIII (когнитивные аспекты языка). — М.: Прогресс, 1988.
3. Падучева Е.В. Коммуникативное выделение на уровне синтаксиса и семантики // Семиотика и информатика. Вып. 36. — М.: ВИНТИ, 1998.
4. Suppes P.A. Probabilistic Theory of Causality. — Amsterdam, 1970.
5. Вригт Г.Х. фон. Логико-философские исследования. Избр. тр. — М.: Прогресс, 1986. — 600 с.

Научное издание

*КАЛИТКИН Николай Николаевич
КАРПЕНКО Надежда Викторовна
МИХАЙЛОВ Александр Петрович
ТИШКИН Владимир Федорович
ЧЕРНЕНКОВ Михаил Владимирович*

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ПРИРОДЫ И ОБЩЕСТВА

Редактор *В.Д. Боксер*
Оригинал-макет: *В.В. Худяков*

Подписано в печать 21.12.04. Формат 60×90/16. Бумага офсетная.
Печать офсетная. Усл. печ. л. 22,5. Уч.-изд. л. 25,0. Тираж 400 экз.
Заказ №

Издательская фирма «Физико-математическая литература»
МАИК «Наука/Интерпериодика»
117997, Москва, ул. Профсоюзная, 90
E-mail: fizmat@maik.ru, fmlsale@maik.ru;
<http://www.fml.ru>

Отпечатано с готовых диапозитивов
в ОАО «Московская типография № 6»
115088, Москва, Ж-88, ул. Южнопортовая, 24