

На правах рукописи



Кондратьев Михаил Александрович

РАЗРАБОТКА МОДЕЛИ РАСПРОСТРАНЕНИЯ ИНФЕКЦИОННЫХ
ЗАБОЛЕВАНИЙ НА ОСНОВЕ АГЕНТНОГО ПОДХОДА

Специальность 05.13.18 — Математическое моделирование, численные
методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург
2012

Работа выполнена на кафедре «Распределенные вычисления и компьютерные сети» Санкт-Петербургского государственного политехнического университета.

Научный руководитель	доктор технических наук, профессор Ивановский Ростислав Игоревич
Официальные оппоненты	доктор физико-математических наук, профессор Нарбут Михаил Александрович кандидат технических наук, доцент Ключарев Александр Анатольевич
Ведущая организация	Санкт-Петербургский институт информатики и автоматизации РАН

Защита состоится 17 апреля 2012 г. в 14:00 часов на заседании диссертационного совета Д 212.233.02 при Федеральном государственном автономном образовательном учреждении высшего профессионального образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» по адресу: 190000, г. Санкт-Петербург, ул. Большая Морская, д. 67.

С диссертацией можно ознакомиться в библиотеке Санкт-Петербургского государственного университета аэрокосмического приборостроения, с авторефератом на сайте <http://guar.ru/>.

Автореферат разослан «02» _____ марта _____ 2012 г.

Ученый секретарь
диссертационного совета
доктор технических наук, профессор



Осипов Л. А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Сохранение и укрепление здоровья населения — важная социально-экономическая проблема, неотъемлемым аспектом которой является снижение инфекционной заболеваемости. В решении этой задачи предупредительные меры имеют ключевое значение. Прогнозирование динамики распространения заболевания позволяет разработать и применить адекватные меры противодействия, обеспечить рациональное использование материальных и людских ресурсов.

Информационной основой прогноза в рассматриваемой области служат статистические данные, регулярно получаемые соответствующими структурами. В Санкт-Петербурге (и многих других регионах России) надзор за распространением гриппа и острых респираторных вирусных инфекций осуществляет центр, организованный на базе НИИ гриппа Минздравсоцразвития России. Специалисты НИИ гриппа инициировали настоящее исследование, а также выступили в качестве научных консультантов и экспертов в ходе его проведения. Исследование направлено на разработку методов и моделей, позволяющих получать прогноз развития эпидемий и оценку эффективности различных мер противодействия их развитию с приемлемой для решения практических задач точностью.

Качественный прогноз распространения заболевания достижим только на основе адекватных математических моделей. К сожалению, традиционные модели распространения заболеваний используют допущения (например, однотипность индивидуумов, их непрерывное равномерное перемешивание на моделируемой территории), делающие результаты прогнозирования с их помощью недостаточно точными. Учитывая последние достижения в области математического и имитационного моделирования, формирование моделей, не обладающих такими ограничениями, является в настоящее время вполне реализуемой задачей.

Целью исследования является создание модели распространения инфекционных заболеваний на примере гриппа А, позволяющей выполнить прогноз изменения уровня заболеваемости на основе текущих статистических данных. Для этого необходимо решить следующие **задачи**:

- 1) анализ существующих моделей распространения заболеваний и методов прогнозирования распространения инфекции, выявление их основных ограничений;
- 2) выбор современного подхода к моделированию распространения заболеваний, позволяющего преодолеть ограничения традиционных моделей;

- 3) создание математической модели территориального распространения заболевания на основе выбранного подхода;
- 4) программная реализация компьютерной модели распространения заболевания для получения прогноза в автоматическом режиме;
- 5) анализ созданной модели распространения заболевания: калибровка параметров модели, проверка корректности модели, оценка чувствительности модели;
- 6) оценка точности прогнозирования уровня заболеваемости на основе построенной модели.

Методы исследования. Для решения поставленных задач использовались методы теории вероятностей и математической статистики: теория проверки гипотез, теория случайных процессов, корреляционный анализ, регрессионный анализ и методы теории оценивания. Разработанная модель основана на методах имитационного моделирования, в частности, агентного и дискретно-событийного моделирования, ее реализация — на принципах объектно-ориентированного программирования.

Научная новизна работы заключается в следующем:

- 1) агентный подход к построению моделей сложных процессов и систем развит для создания имитационной модели распространения заболеваний, выделены агенты-люди, значимые особенности их поведения, окружающее пространство и его свойства;
- 2) построена многоагентная модель распространения гриппа нового типа, не связанная ограничениями, присущими традиционным подходам к моделированию распространения заболеваний и позволяющая учитывать привязку агента к определенному месту в окружающем его пространстве, особенности поведения агента в зависимости от его возраста, вероятностную природу эпидемических процессов;
- 3) разработанная модель, а также группа вычислительных экспериментов с ней, обеспечивающих различные режимы работы модели, реализованы в виде единой программной системы, поддерживающей принятие административных и санитарно-эпидемиологических решений для борьбы с инфекционными заболеваниями;
- 4) предложены алгоритмы прогнозирования уровня заболеваемости на основе имитационной модели, обеспечивающие корректировку ее поведения в соответствии с текущей статистикой заболеваемости за счет динамической подстройки параметров модели.

Практическая ценность результатов работы. Созданный программный комплекс позволяет построить краткосрочный прогноз развития эпидемической обстановки в городе. За счет использования

агентного подхода, в предложенной модели без затруднений могут быть учтены любые меры борьбы с заболеваемостью (вакцинация, карантин и проч.). Разработанная система позволяет количественно оценить эффективность тех или иных мер и выбрать наиболее подходящие, исходя из текущего уровня заболеваемости.

Исследование проводилось при финансовой поддержке в форме грантов Правительства Санкт-Петербурга 2009 г. и 2010 г. для студентов, аспирантов вузов и академических институтов, расположенных на территории Санкт-Петербурга.

Апробация работы. Основные результаты диссертационного исследования докладывались и обсуждались на следующих конференциях:

- международной научно-практической конференции «XXXVIII Неделя науки СПбГПУ» (Санкт-Петербург, 2009);
- международной студенческой олимпиаде по автоматическому управлению «Балтийская олимпиада» (Санкт-Петербург, 2010);
- XIV международной научно-практической конференции «Системный анализ в проектировании и управлении» (Санкт-Петербург, 2010);
- научно-технической конференции «Проблема комплексного обеспечения информационной безопасности и совершенствование образовательных технологий подготовки специалистов силовых структур» (Санкт-Петербург, 2010).

Внедрение результатов. Разработанный на основе имитационной модели программный комплекс и прочие результаты работы переданы специалистам ФГБУ «НИИ гриппа» Минздравсоцразвития России и используются на практике при краткосрочном прогнозировании уровня заболеваемости гриппом и острыми респираторными вирусными инфекциями в Санкт-Петербурге. Результаты работы применены в ЗАО «МБНПК «Цитомед» в ходе статистического анализа и прогнозирования заболеваемости гриппом в регионах РФ.

Публикации. Материалы, отражающие основное содержание и результаты диссертации, опубликованы в 7 печатных работах. Из них 2 работы опубликованы в рецензируемых научных журналах, утвержденных в перечне ВАК.

Основные положения, выносимые на защиту диссертации:

- 1) агентный подход к имитационному моделированию, развитый для решения медико-социальных проблем;
- 2) модель распространения заболевания нового типа, построенная на основе агентного подхода;

- 3) прототип программной системы поддержки принятия решений по управлению силами и средствами противодействия развитию эпидемий, основанной на разработанной модели;
- 4) процедуры прогнозирования с помощью имитационного моделирования, основанные на адаптации параметров модели.

Структура и объем работы. Диссертационная работа состоит из введения, пяти разделов, заключения, списка использованных источников и 20 приложений. Основной текст работы содержит 135 страниц машинописного текста, 38 рисунков и 4 таблицы. Объем приложений — 57 страниц, включая 10 рисунков и 20 таблиц. Список использованной литературы содержит 82 наименования.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы работы, представлена научная новизна и практическая ценность работы, а также приведено краткое содержание диссертационной работы по разделам.

В **первом разделе** рассмотрены существующие модели распространения заболеваний (далее МРЗ), возможные подходы к их построению и их историческое развитие. Начало применению математических методов при изучении эпидемий было положено в XVII веке. За прошедшее время методы моделирования заболеваний многократно усовершенствовались, однако сделать вывод, что поставленная задача решена, пока нельзя.

Большая часть МРЗ основана на использовании систем алгебро-дифференциальных уравнений, решением которых является уровень инфекционной заболеваемости в каждый момент модельного времени. Наиболее известной среди таких МРЗ является модель SIR, предложенная В. О. Кермаком (W. O. Kermack) и А. Г. Мак-Кендриком (A. G. McKendrick) в 1927 г. В этой модели, как и в прочих моделях такого типа, вся популяция моделируемой территории делится на группы:

- «Susceptible» — здоровые люди, восприимчивые к заболеванию (обозначим их количество S);
- «Infectious» — инфекционные больные (обозначим их количество I);
- «Recovered» — переболевшие моделируемым заболеванием люди, более к нему не восприимчивые (обозначим их количество R).

Учитывая, что общее число индивидов в популяции остается неизменным, приращение числа людей в каждой из выделенных групп можно описать с помощью следующей системы уравнений:

$$\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma I,$$

где β и γ — интенсивности контактов и выздоровления индивидов.

Существует множество модификаций модели SIR, выполненных, в основном, для учета особенностей того или иного моделируемого заболевания. Например, обычно вводят группу индивидов «Exposed» (модель SEIR), к которой относятся люди с заболеванием в инкубационном периоде. Назовем такой подход к построению МРЗ классическим, традиционным. К этому же классу МРЗ можно отнести и модели, использующие системы уравнений более сложной структуры. Например, в широко известной в России модели О. В. Барояна и Л. А. Рвачева решается система нелинейных интегро-дифференциальных уравнений в частных производных.

Классические МРЗ продолжают развиваться и сегодня, в частности, в работах Б. В. Боева. Однако для применения математического аппарата дифференциальных уравнений в таких МРЗ предполагается, что все индивиды однотипны и непрерывно и равномерно перемешиваются на моделируемой территории. В действительности же большинство реальных популяций имеют крайне сложную структуру, обусловленную социальным расслоением, разнообразием географических условий и сложными временными и пространственными схемами перемещения. Кроме того, классические МРЗ являются полностью детерминированными и подходят только для оценки поведения очень больших популяций. Природа эпидемических процессов носит вероятностный характер, пренебрегая случайными факторами, можно получить грубые или ошибочные результаты моделирования.

Для преодоления ограничений традиционных моделей сегодня применяются дискретно-событийные модели. Такие модели на этапе исполнения представляют собой некоторое множество переменных состояния, характеризующих модель, а также список запланированных событий, которые должны произойти в определенной последовательности в будущем. Каждое событие в модели — некоторый набор атомарных действий, переопределяющих значения переменных состояния.

В последние годы наибольший вклад в моделирование распространения заболеваний внесли так называемые «популяционные» (population-based) модели, одна из первых которых была разработана А. М. Лонджини (I. M. Longini). Это дискретно-событийные модели, в которых отражена простейшая структура общества: в популяции явно выделены индивиды различного возраста, распространение заболевания между которыми может произойти только в рамках одной «контактной» группы (contact group). Контактные группы определяются характерной структурой общества, которая будет зависеть от моделируемой

территории. Например, в одну контактную группу попадут одноклассники, сослуживцы, члены семьи, дети и проч.

Стадии протекания заболевания в популяционных моделях обычно соответствуют выделенным в модели SEIR. Время в популяционных моделях движется с фиксированным шагом дискретизации, обычно достаточно большим — в 12 или 24 часа. На каждом шаге анализируется, какие локации посетил каждый индивид и по сложной формуле, учитывающей различные факторы, рассчитывается вероятность его заражения за прошедший промежуток времени (достаточного контакта). Фиксированный шаг дискретизации и формула расчета достаточного контакта — основные источники погрешностей в работе популяционных моделей. Кроме того, коэффициенты, используемые в этой формуле, требуют калибровки, проведение которой может быть затруднительно в случае необходимости учета в модели административных мер борьбы с заболеваемостью.

Справиться с этими сложностями позволяет агентный подход к имитационному моделированию. Этот подход, развитый в последнее время, основан на учете множества параллельно протекающих элементов исследуемого процесса, каждый из которых при моделировании описывается совокупностью детерминированных и случайных параметров, определяющих особенности «жизненного цикла» элемента. Многоагентная модель представляет реальный мир в виде многих отдельно специфицируемых активных подсистем, называемых агентами. Каждый из агентов взаимодействует с другими агентами и в процессе функционирования может менять свое поведение и учитывать изменение внешней среды.

Агентные МРЗ похожи на популяционные и, как правило, также дискретно-событийные. В качестве агентов в них рассматриваются индивиды, формирующие популяцию, определенным образом описывается протекание заболевания у каждого агента. На основе социальной структуры моделируемой популяции формируются контактные группы, между которыми перемещаются агенты и в которых может произойти достаточный контакт.

Основное отличие агентных МРЗ от популяционных заключается в том, что они децентрализованы. У агентных МРЗ шаг дискретизации не фиксирован, а момент времени, в который происходит достаточный контакт, определяется самим агентом. Потоки событий от всех агентов выстраиваются в один общий в хронологическом порядке и тем самым формируют интегральную динамику моделируемой системы.

В настоящее время не существует МРЗ, полностью использующей преимущества агентного подхода. Представляет интерес создание агентной МРЗ, позволяющей промоделировать распространение

различных видов инфекционных заболеваний путем параметрической настройки модели. Разработку такой МРЗ решено начать с одного заболевания — гриппа А.

Грипп выбран потому, что, с одной стороны, грипп — главная составляющая инфекционной заболеваемости и смертности по всему миру, и поиск способов борьбы с ним чрезвычайно актуален. С другой стороны, грипп А является хорошо изученным заболеванием, известны возможные формы его протекания, способы передачи вируса. Существует подробная статистика заболеваемости гриппом, которая может быть использована для настройки и анализа модели.

Второй раздел работы посвящен анализу статистических данных, описывающих процесс распространения заболевания. Для калибровки и оценки качества работы модели применялась статистика заболеваемости гриппом и острыми респираторными вирусными инфекциями (ОРВИ) в Санкт-Петербурге в 1986–2009 гг. (см. пример на рис. 1). Были рассмотрены особенности заболеваемости гриппом, а также выполнена оценка достижимой точности и срока ее прогнозирования. Для этого задача краткосрочного прогнозирования развития эпидемической обстановки в городе была решена с помощью регрессионного анализа.

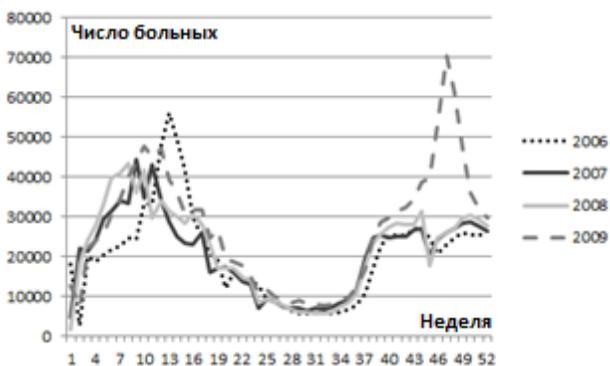


Рисунок 1 — Заболеваемость гриппом и ОРВИ в Санкт-Петербурге в 2006–2009 гг.

Процесс распространения заболевания был представлен в виде функциональной зависимости $\hat{y} = f(t, \hat{\beta})$, где \hat{y} — оценка числа больных горожан, t — время, а $\hat{\beta}$ — вектор оценок неизвестных параметров регрессии. Сформированы рекомендации по выбору регрессионной зависимости и числа наблюдений для получения наилучшего прогноза. Так, целесообразно применять регрессионные полиномы степени не выше третьей, в среднем же наиболее точные результаты прогнозирования на одну неделю вперед получены с помощью полинома первой степени. Для

данных 2006–2009 гг. в таком случае абсолютная ошибка прогнозирования составляет 5,5 % от максимального количества одновременно больных гриппом горожан за этот промежуток времени.

Предприняты попытки улучшить качество прогноза, полученного с помощью решения задачи однофакторной регрессии. Для этого в рассмотрение, помимо времени, был введен еще один фактор. Точечная оценка корреляционного отношения между заболеваемостью гриппом и температурой воздуха в городе в среднем составляет $-0,76$ (при использовании в качестве регрессионной зависимости полинома второй степени), что говорит об их заметной взаимосвязи. Для получения прогноза будем искать полиномиальную зависимость $\hat{y} = f(t, \theta, \hat{\beta})$, где θ — температура. К сожалению, таким путем не удалось получить прогноз точнее, чем в случае однофакторной регрессии. Хотя, теоретически, использование данных о погодных условиях позволяет улучшить прогноз как минимум на 14 %.

Другой подход к увеличению качества прогноза — фильтрация аномальных выбросов. Наблюдается наложение на исходный процесс распространения заболевания сильного шума, бороться с которым практически невозможно, так как он вызван особенностями сбора статистики. Одним из проявлений шума является наличие в статистических данных резких изменений — аномальных выбросов. Для снижения их влияния на прогноз использовалось интервальное оценивание. По результатам решения задачи регрессии рассчитывались доверительные интервалы и все наблюдения, не попавшие в эти интервалы, отбрасывались и не учитывались в ходе повторного решения задачи регрессии для построения прогноза. Показано, что в некоторых случаях применение фильтрации аномальных выбросов может улучшить качество прогноза на 10–15 % по сравнению с однофакторной регрессией.

Для оценки допустимого срока прогнозирования в ходе решения задачи регрессии были рассчитаны априорные оценки (дисперсия прогноза) и апостериорные оценки (ошибка прогнозирования) качества прогноза, выполненного для различных промежутков времени. Прогноз приемлемого качества может быть получен на срок до месяца вперед. Ошибка прогнозирования на 4 недели вперед уже в 3 раза выше ошибки при прогнозе на неделю вперед.

В **третьем разделе** работы представлена предложенная стохастическая многоагентная дискретно-событийная модель распространения гриппа на территории одного города (или района). Агентом в модели является человек, принадлежащий к одной из возрастных групп: 0–2 года, 3–6 лет, 7–14 лет, 15–24 года, 25–39 лет, 40–64 года, 65 лет и старше. Число агентов каждой возрастной группы в модели определяется демографическими данными по моделируемой

территории. Возрастная группа агента характеризует такие важные факторы, определяющие вероятность заболевания, как количество контактов с другими агентами в день и возможные места пребывания человека (школа, работа и т. д.).

Основные два типа активных объектов в модели — агенты и локации. Локация это объект, моделирующий возможное местонахождение агента. Каждый агент ежедневно согласно своему внутреннему расписанию событий проводит определенное время в различных локациях. Например, агент от 7 до 14 лет может посещать такие локации как «дом», «школа» и «транспорт». Выделим 8 типов локаций: дом, детские ясли, детский сад, школа, внешкольные занятия, вуз (а также средние профессиональные учебные заведения и проч.), работа, общественный транспорт. При необходимости, этот список может быть пополнен. Количество различных мест пребывания каждого типа в модели определяется на основе общего количества агентов в модели и средней численности группы агентов, обычно посещающей данную локацию. Распределение агентов по всем локациям (кроме общественного транспорта) фиксировано (то есть школьник каждый день ходит в один и тот же класс, студент — в один и тот же вуз и т. д.). Все множество локаций образует окружающую среду для агентов.

Этапы протекания заболевания у агента изображены на рис. 2 с помощью стейтчарта (карты состояний). Смысл основных состояний агентов аналогичен принятому в классических моделях типа SEIR. Состояния «stateX» и «stateY» введены, чтобы промоделировать стадии инкубационного периода, пока больной не заразен и то время, когда он уже может заражать других агентов. Состояния «stateA», «stateB» и «stateC» моделируют постепенное снижение способности к распространению заболевания по мере выздоровления агента.

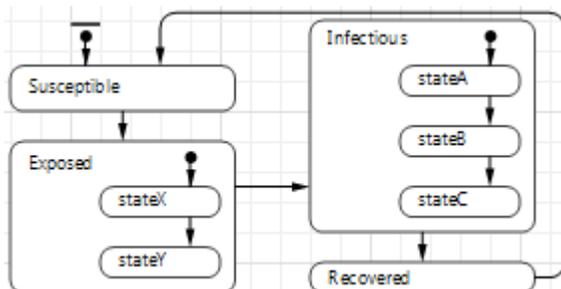


Рисунок 2 — Протекание заболевания у агента

Переходы между всеми состояниями, кроме перехода из «Susceptible» в «Exposed», осуществляются по таймауту. Все таймауты параметрически настраиваемы под моделируемый вид заболевания. Так

как рассматриваемая модель ориентирована на грипп А, агент проводит в состоянии «Exposed» в среднем от 1 до 3 дней, в состоянии «Infectious» от 6 до 8 дней, в состоянии «Recovered» 4 недели.

Переход из состояния «Susceptible» в «Exposed» происходит в случае передачи заболевания от одного агента к другому, которая может произойти в результате события «контакт». Это событие вызывается с определенной интенсивностью больными агентами (находящимися в состояниях «stateY» или «Infectious»). При «контакте» заболевание может быть передано с определенной вероятностью только здоровому агенту (в состоянии «Susceptible»), находящемуся в той же самой локации, что и больной агент. Вероятность передачи заболевания зависит от стадии заболевания больного агента, восприимчивости здорового агента и текущего времени года.

Время года (погодные условия) в модели учитывается с помощью сезонных коэффициентов. Каждый месяц модельного времени разбит на две части, каждой соответствует свой коэффициент, регулирующий вероятность достаточного контакта. Значения сезонных коэффициентов вычисляются в процессе калибровки модели на основе текущих данных заболеваемости гриппом.

Общая структура модели изображена на рис. 3. Множество агентов и локаций взаимодействуют друг с другом с помощью корневого объекта, несущего преимущественно вспомогательные функции. Перечислим основные виды событий (динамики) в модели, формирующие эпидемическую обстановку:

- событие «контакт», порождаемое больным агентом;
- события, связанные с протеканием заболевания у агента;
- события, связанные с процессами вакцинации агентов (действие вакцины поэтапно и описано с помощью стейтчарта);
- события перемещения агентов между локациями (порождаются локациями, в которых находятся агенты);
- события, связанные с сезонными изменениями параметров модели, регулирующими зависимость заболеваемости от погоды и т. д.;
- события введения и отмены карантина в определенных локациях, обычно регулируемые пользователем модели.

В качестве исходных данных для построенной МРЗ выступает уровень заболеваемости гриппом. Основными выходными показателями является число больных горожан в каждый момент модельного времени, а также экономический ущерб от заболеваемости гриппом, рассчитанный на основе средней взвешенной стоимости одного случая заболевания.

В **четвертом** разделе работы описаны детали реализации компьютерной МРЗ и ее основные режимы работы. Построение МРЗ на языке программирования высокого уровня не представляется

целесообразным ввиду излишней сложности такой задачи. Рассмотрены среды агентного моделирования, используемые для моделирования распространения заболеваний: SOARS, NetLogo, Swarm. Единственной на сегодняшний день системой моделирования, удовлетворяющей всем требованиям построенной МРЗ, является отечественная среда AnyLogic, в которой и была реализована модель.

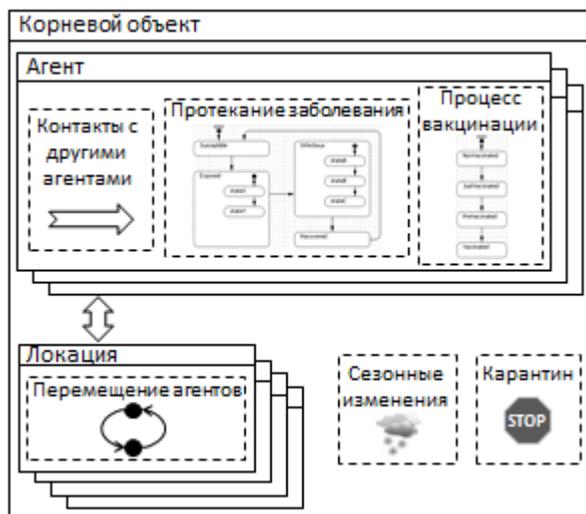


Рисунок 3 — Структура агентной модели распространения гриппа

Созданный программный комплекс, помимо классов, обеспечивающих работу непосредственно модели, включает классы вычислительных экспериментов, поддерживающих различные режимы и сценарии работы модели, а также способы анализа результатов моделирования. Все эксперименты объединены в единое приложение — систему прогнозного моделирования развития эпидемий. Базовые эксперименты с моделью — динамический и стохастический.

Динамический эксперимент — простейший из доступных пользователю системы. В его процессе модель выполняется один раз, интерактивно для пользователя, что позволяет наглядно продемонстрировать работу агентной МРЗ. Существует возможность проанализировать реакцию модели на выбранную комбинацию административных мер снижения инфекционной заболеваемости — в текущей версии модели доступно проведение вакцинации выбранной доли населения, а также введение карантина.

Стохастический эксперимент объединяет несколько режимов работы модели, требующих ее многократного исполнения. Так как агентная МРЗ — стохастическая, каждое ее исполнение представляет

с собой одну из реализаций случайного процесса, определяемого моделью. Для получения достоверных результатов моделирования в большинстве случаев требуется усреднение нескольких реализаций данного процесса.

Первый режим стохастического эксперимента предназначен для многократного исполнения модели с целью получения усредненного прогноза развития эпидемической обстановки. Второй режим позволяет проанализировать распределение частот определенного события в модели с помощью построения гистограммы. Таким образом можно оценить риски превышения заданных эпидемических порогов при определенных мерах противодействия заболеваемости. Третий режим основан на поисковых процедурах оптимизатора OptQuest, встроенного в среду AnyLogic, и позволяет выбрать наилучшую стратегию вакцинации населения в установленных условиях. Четвертый режим поддерживает работу других экспериментов и нужен для поиска и сохранения заданного состояния модели.

Пятый раздел работы посвящен анализу построенной имитационной МРЗ и содержит описание соответствующих экспериментов с моделью. Значительного объема работ потребовала калибровка модели — тонкая настройка ее параметров, обеспечивающая максимальное приближение результатов расчетов к данным натуральных измерений. Калибровка модели проводилась поэтапно.

На начальном этапе был проведен предварительный анализ чувствительности модели к изменениям значений ее отдельных параметров, по результатам которого были выделены две группы калибровочных параметров. Первую группу образуют коэффициенты восприимчивости агентов к инфекции. Эти параметры влияют на вероятность достаточного контакта агентов каждой возрастной группы в локации каждого типа. В основу калибровки коэффициентов восприимчивости положен тот факт, что в течение года в Санкт-Петербурге сохраняется приблизительно одинаковое соотношение долей больных индивидов различных возрастных групп от общего числа больных. Для поиска коэффициентов, обеспечивающих такое соотношение, был создан отдельный эксперимент, работающий с использованием процедур оптимизатора OptQuest.

Вторая группа калибровочных параметров — это сезонные коэффициенты, регулирующие время года и погоду в модели. Разработан вычислительный эксперимент, позволяющий настроить все сезонные коэффициенты в году так, чтобы поведение модели было максимально близко к наблюдаемой статистике заболеваемости. По результатам данного эксперимента была проверена корректность модели. Среднеквадратическая оценка разницы между общим числом больных агентов в каждый момент модельного времени и количеством реально

больных гриппом горожан в тот же момент времени составила 0,28 % от максимального числа одновременно больных горожан за рассмотренные 2006–2007 гг., что позволяет считать модель адекватной.

С помощью специального эксперимента был проведен анализ чувствительности настроенной модели, по результатам которого можно сделать вывод об устойчивости модели к отклонениям ее параметров.

Одним из важнейших является эксперимент по построению прогноза. Каждый год процесс распространения гриппа в Санкт-Петербурге несколько изменяется, и известные исторические данные могут быть использованы только в качестве начального приближения ожидаемого уровня заболеваемости. Эти особенности, также как и текущая погода, в модели регулируются с помощью сезонных коэффициентов. Необходимо поддерживать их значения в соответствии с наблюдаемой эпидемической обстановкой в городе.

Для этого было предложено два схожих алгоритма. В качестве начальных приближений они используют прошлогодние значения сезонных коэффициентов. В обоих алгоритмах по мере поступления новой статистики заболеваемости выполняются следующие действия:

- проводится проверка качества и, при необходимости, калибровка сезонного коэффициента за промежуток времени, для которого собрана полная статистика заболеваемости (этот коэффициент понадобится только через год);
- оценивается соответствие действующего сезонного коэффициента свежим данным об уровне заболеваемости и, при необходимости, производится подстройка значения коэффициента;
- с помощью настроенной модели рассчитывается прогноз развития эпидемической обстановки.

Алгоритмы отличаются условиями, при которых принимается решение о корректировке значения коэффициента, и допустимыми диапазонами его изменения. Один алгоритм ориентирован на применение в условиях ординарного уровня заболеваемости в городе, другой — на быстро меняющуюся аномальную эпидемическую обстановку. На рис. 4 представлен внешний вид получаемого прогноза (широкая серая линия – результат численного эксперимента, тонкая черная линия – реально полученная статистика заболеваемости за соответствующий период времени).

Проведено сравнение качества прогноза, полученного с помощью агентной МРЗ и регрессионного анализа. В условиях осенней эпидемии гриппа, точность работы агентной модели в полтора раза выше, чем в случае всех предложенных методов, основанных на регрессионном анализе. Однако при усреднении всех результатов, полученных в течение одного года, качество прогнозирования на одну неделю вперед с

помощью решения задачи многофакторной регрессии на 25 % лучше результатов имитационного моделирования. При прогнозировании на две и более недели, точность агентной МРЗ выше как минимум на 15 %.

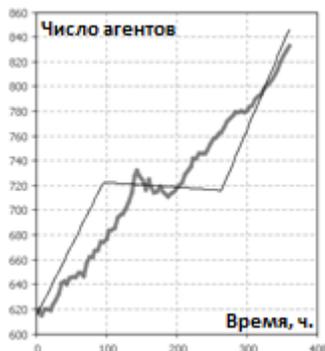


Рисунок 4 — Пример прогноза, полученного с помощью агентной модели распространения гриппа

Для учета при построении прогноза положительных качеств регрессионного анализа, разработанная система прогнозного моделирования развития эпидемий была дополнена модулем решения задачи регрессии, реализованном на языке программирования Java. Система работает на основе исходных данных, хранящихся в электронных таблицах Excel, в которые также записываются результаты экспериментов. Таким образом, был создан прототип системы поддержки принятия административных и санитарно-эпидемиологических решений для борьбы с инфекционными заболеваниями, основанный на имитационной модели.

В **заключении** приведены основные результаты, полученные в диссертационной работе, а также возможные направления развития исследования.

В **приложении А** представлена модель протекания локальных процессов заболевания в организме человека Г. И. Марчука, которая может быть использована в агентной модели территориального распространения заболеваний.

В **приложениях Б–И** подробно рассмотрены методы регрессионного анализа, применявшиеся для построения прогноза заболеваемости гриппом в Санкт-Петербурге, а также указаны оценки качества прогнозирования с их помощью.

В **приложениях К–М и Р** перечислены параметры агентной МРЗ, их значения по умолчанию, а также способы задания начального состояния агентов в модели.

В **приложении Н** содержится способ расчета экономического ущерба от заболеваемости гриппом на основе собранной статистики заболеваемости.

В **приложениях П и У** приведены результаты анализа чувствительности модели (предварительного и для настроенной модели).

В **приложениях С и Т** представлены результаты экспериментов по калибровке коэффициентов восприимчивости и сезонных коэффициентов модели.

В **приложениях Ф–Ц** рассмотрены результаты прогнозирования с помощью агентной МРЗ и рассчитаны оценки качества полученного прогноза.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Предложены способы построения прогноза уровня заболеваемости с помощью решения задачи однофакторной и многофакторной регрессии. Сформулированы рекомендации по выбору регрессионных зависимостей и числа наблюдений. Модуль прогнозирования на основе регрессионного анализа был внедрен в существующую систему мониторинга уровня заболеваемости гриппом в Санкт-Петербурге.
2. Предложен способ фильтрации аномальных выбросов в статистике заболеваемости с помощью интервального оценивания при построении прогноза с помощью регрессионного анализа.
3. Агентный подход развит применительно к задаче моделирования распространения заболевания. Сформулированы особенности поведения агентов, свойства окружающего пространства.
4. Построена стохастическая многоагентная дискретно-событийная модель распространения гриппа А на территории одного города. Ее параметры могут быть настроены для моделирования не только гриппа, но и других видов инфекционных заболеваний.
5. Разработаны процедуры калибровки параметров агентной модели для достижения максимального приближения результатов работы модели к данным натуральных измерений заболеваемости на исследуемой территории.
6. Предложены алгоритмы прогнозирования уровня заболеваемости с помощью имитационной модели, основанные на динамической подстройке параметров модели для корректировки ее поведения в соответствии с текущей статистикой заболеваемости. Прогноз, получаемый с помощью агентной модели, до полутора раз точнее выполненного на основе регрессионного анализа.

7. Разработан стохастический эксперимент с имитационной моделью распространения гриппа, позволяющий оценить риск возникновения определенного события при заданных условиях. В частности, можно рассчитать вероятность получения определенного экономического ущерба при выбранных мерах борьбы с заболеваемостью.
8. Построен вычислительный эксперимент по выбору наилучшей стратегии вакцинации населения в сложившейся эпидемической обстановке. Возможен учет и других мер снижения заболеваемости, например карантина.
9. На основе разработанной модели и методов регрессионного анализа создан прототип программной системы поддержки принятия решений по управлению силами и средствами противодействия развитию эпидемий.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Основное содержание работы изложено в следующих публикациях (статьи 1–2 опубликованы в изданиях, включенных в перечень ВАК).

1. **Кондратьев М. А., Ивановский Р. И., Цыбалова Л. М. Применение агентного подхода к имитационному моделированию процесса распространения заболевания // Научно-технические ведомости СПбГПУ. Серия «Наука и образование». — 2010. — Т. 2, № 2. — С. 189–195.**
2. **Кондратьев М. А. Имитационное моделирование в медицине: многоагентная модель распространения гриппа // Компьютерные инструменты в образовании. — 2011. — № 4. — С. 32–36.**
3. Кондратьев М. А., Ивановский Р. И. Прогноз динамики распространения инфекционных заболеваний // XXXVIII Неделя науки СПбГПУ: Материалы международной научно-практической конференции. Ч. VIII. — СПб.: Изд-во Политехн. ун-та, 2009. — С. 161–163.
4. Kondratyev M. A. Decision support system for infectious disease spread containing // Preprints of 13th International Student Olympiad on Automatic Control. — Saint-Petersburg, 2010. — P. 87–90.
5. Кондратьев М. А., Изотов А. В. Оценка экономического ущерба заболеваемости гриппом на основе имитационного моделирования // Системный анализ в проектировании и управлении: сборник научных трудов XIV Междунар. науч.-практ. конф. Ч. 1. — СПб.: Изд-во Политехн. ун-та, 2010. — С. 180–182.

6. Кондратьев М. А. Имитационное моделирование как основа поддержки принятия решений в условиях неопределенности // Проблема комплексного обеспечения информационной безопасности и совершенствование образовательных технологий подготовки специалистов силовых структур. Межвузовский сборник научно-технических статей. — СПб.: СПбГУ ИТМО, 2010. — С. 61–68.
7. Кондратьев М. А. Совершенствование методологии моделирования распространения инфекционных заболеваний // Вычислительные, измерительные и управляющие системы: сборник научных трудов. — СПб.: Изд-во Политехн. ун-та, 2011. — С. 37–44.