

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Инженерная школа ядерных технологий
Направление подготовки 01.04.02 «Прикладная математика и информатика»
Отделение экспериментальной физики

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Моделирование распространения эпидемий с использованием сетей

УДК 519.17:573:616-036.22

Студент

Группа	ФИО	Подпись	Дата
0ВМ02	Дышеков Астемир Анзорович		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семёнов М.Е.	к.ф.-м.н., доцент		

Со-руководитель (по разделу «Концепция стартап-проекта»)

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Рыжкова М.В.	Д.Э.Н., доцент		

КОНСУЛЬТАНТЫ:

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Федорчук Ю.М.	Д.Т.Н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семёнов М.Е.	к.ф.-м.н., доцент		

Томск – 2022 г.

Планируемые результаты обучения по ООП

Код компетенции	Наименование компетенции
Профессиональные компетенции	
ПК(У)-1	Способен проводить научные исследования и получать новые научные и прикладные результаты самостоятельно и в составе научного коллектива
ПК(У)-2	Способен проводить поиск и анализ научной и научно-технической литературы по тематике проводимых исследований
ПК(У)-3	Способен разрабатывать и анализировать показатели качества информационных систем, используемых в производственной деятельности
ПК(У)-4	Способен планировать научно-исследовательскую деятельность, анализировать риски, управлять проектами, управлять командой проекта
ПК(У)-5	Способен преподавать математических дисциплин и информатики в образовательных организациях высшего образования
ПК(У)-6	Способен проектировать и организовывать учебный процесс по образовательным программам с использованием современных образовательных технологий
Общепрофессиональные компетенции	
ОПК(У)-1	Способен решать актуальные задачи фундаментальной и прикладной математики
ОПК(У)-2	Способен совершенствовать и реализовывать новые математические методы решения прикладных задач
ОПК(У)-3	Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности
ОПК(У)-4	Способен комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности
Универсальные компетенции	
УК(У)-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий
УК(У)-2	Способен управлять проектом на всех этапах его жизненного цикла
УК(У)-3	Способен организовывать и руководить работой команды, выработывая командную стратегию для достижения поставленной цели
УК(У)-4	Способен применять современные коммуникативные технологии, в том числе на иностранном языке, для академического и профессионального взаимодействия
УК(У)-5	Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия
УК(У)-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Инженерная школа ядерных технологий
Направление подготовки 01.04.02 «Прикладная математика и информатика»
Отделение экспериментальной физики

УТВЕРЖДАЮ:

Руководитель ООП

_____ Семенов М.Е.

(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
ОВМ02	Дышекову Астемиру Анзоровичу

Тема работы:

Моделирование распространения эпидемий с использованием сетей	
Утверждена приказом директора (дата, номер)	

Срок сдачи студентом выполненной работы:	31.05.2022
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	<i>Статистические данные распространения коронавирусной инфекции Covid-19.</i>
---------------------------------	--

Перечень подлежащих исследованию, проектированию и разработке вопросов	<ol style="list-style-type: none"> 1. Провести сравнительный анализ литературы по теме диссертации. 2. Разработка базовой версии математической модели на основе концепции SIR. 3. Разработка модели случайного графа с кликовой структурой. 4. Реализация модели в виде программного кода. 5. Верификация модели на статистических данных.
---	--

Перечень графического материала	<ol style="list-style-type: none"> 1. Схема математической модели. 2. Графики по анализу свойств случайного графа с кликовой структурой. 3. Пояснительные иллюстрации графа или его частей. 4. Графики с сопоставлением результатов моделирования и данных статистики.
--	--

Консультанты по разделам выпускной квалификационной работы

(если необходимо, с указанием разделов)

Раздел	Консультант
Концепция стартап-проекта	Рыжкова М.В.
Социальная ответственность	Федорчук Ю.М.
Иностранный язык	Панамарёва А.Н.

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	15.03.2022 г.
---	---------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семенов Михаил Евгеньевич	к. ф.-м. н., доцент		15.03.2022 г.

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0ВМ02	Дышеков Астемир Анзорович		15.03.2022 г.

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«КОНЦЕПЦИЯ СТАРТАП-ПРОЕКТА»**

Студенту:

Группа	ФИО
ОВМ02	Дышеков Астемир Анзорович

Школа	ИЯТШ	Направление	01.04.02 «Прикладная математика и информатика»
Уровень образования	Магистратура		

Перечень вопросов, подлежащих разработке:

<i>Проблема конечного потребителя, которую решает продукт, который создается в результате выполнения НИОКР (функциональное назначение, основные потребительские качества)</i>	
<i>Способы защиты интеллектуальной собственности</i>	
<i>Объем и емкость рынка</i>	
<i>Современное состояние и перспективы отрасли, к которой принадлежит представленный в ВКР продукт</i>	
<i>Себестоимость продукта</i>	
<i>Конкурентные преимущества создаваемого продукта</i>	
<i>Сравнение технико-экономических характеристик продукта с отечественными и мировыми аналогами</i>	
<i>Целевые сегменты потребителей создаваемого продукта</i>	
<i>Бизнес-модель проекта</i>	
<i>Производственный план</i>	
<i>План продаж</i>	

Перечень графического материала:

<i>При необходимости представить эскизные графические материалы (например, бизнес-модель)</i>	
---	--

Дата выдачи задания для раздела по линейному графику	01.03.2022 г.
---	----------------------

Задание выдал консультант по разделу «Концепция стартап-проекта» (со-руководитель ВКР):

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Рыжкова М.В.	Д.э.н., доцент		01.03.2022 г.

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
ОВМ02	Дышеков А.А.		01.03.2022 г.

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа	ФИО
0ВМ02	Дышеков Астемир Анзорович

ШКОЛА	иятш	Отделение школы (НОЦ)	
Уровень образования	Магистр	Направление/специальность	01.04.02 Прикладная математика и информатика

Тема дипломной работы: «Моделирование распространения эпидемий с использованием сетей»

Исходные данные к разделу «Социальная ответственность»:	
<p>1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения</p>	<p>Объектом исследования являются SIR-подобная модель на основе графов. Алгоритм модели реализован на языке программирования python.</p> <p>Разработанная модель позволяет прогнозировать эпидемиологическую обстановку по распространению новой коронавирусной инфекции.</p>
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<p>1. Производственная безопасность</p> <p>1.1. Анализ выявленных вредных факторов</p> <ul style="list-style-type: none"> • Природа воздействия • Действие на организм человека • Нормы воздействия и нормативные документы (для вредных факторов) • СИЗ коллективные и индивидуальные <p>1.2. Анализ выявленных опасных факторов :</p> <ul style="list-style-type: none"> • Термические источники опасности • Электробезопасность • Пожаробезопасности 	<p>1. Вредные факторы:</p> <p>1.1 Недостаточная освещенность;</p> <p>1.2 Нарушения микроклимата, оптимальные и допустимые параметры;</p> <p>1.3 Шум, ПДУ, СКЗ, СИЗ;</p> <p>1.4 Повышенный уровень электромагнитного излучения, ПДУ, СКЗ, СИЗ;</p> <p>2. Опасные факторы:</p> <p>2.1 Электроопасность; класс электроопасности помещения, безопасные номиналы I, U, R_{заземления}, СКЗ, СИЗ; Проведен расчет освещения рабочего места; представлен рисунок размещения светильников на потолке с размерами в системе СИ;</p>
<p>2. Экологическая безопасность:</p> <ul style="list-style-type: none"> • Выбросы в окружающую среду • Решения по обеспечению экологической безопасности 	<p>Воздействие объекта на атмосферу и гидросферу отсутствует. Рассмотрено влияние на литосферу при утилизации компьютера, используемого в разработке, и люминесцентных ламп освещения.</p>
<p>3. Безопасность в чрезвычайных ситуациях:</p> <ul style="list-style-type: none"> • перечень возможных ЧС при разработке и эксплуатации проектируемого решения; • разработка превентивных мер по предупреждению ЧС; 	<p>Возможная ЧС: сильные морозы зимой и штормовая погода летом (сильный порывистый ветер),</p>

<ul style="list-style-type: none"> • разработка действий в результате возникшей ЧС и мер по ликвидации её последствий. 	<p>приводящие к аварии на электро-, тепло-коммуникациях, водоканале, транспорт.</p>
<p>4. Перечень нормативно-технической документации.</p>	<ul style="list-style-type: none"> • СанПин 2.2.2/2.4.1340-03 • СанПин 2.2.4.548-96 • СанПин 2.2.4.3359-16 • СНиП 2.01.02-85 • СП 52.13330.2016 • СН 2.2.4/2.1.8.562-96 • ГН 2.2.5.1313-03 • ГОСТ 54 30013-83 • ГОСТ 12.1.004-91 • ГОСТ Р 55710-2013 • ГОСТ Р 55102-2012

Дата выдачи задания для раздела по линейному графику	26.03.2022 г..
---	----------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ТПУ	Федорчук Ю.М.	д.т.н.		26.03.2022 г.

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
ОВМ02	Дышеков А.А.		26.03.2022 г.

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Инженерная школа ядерных технологий
Направление подготовки 01.04.02 «Прикладная математика и информатика»
Отделение экспериментальной физики
Период выполнения весенний семестр 2021/2022 учебного года

Форма представления работы:

Магистерская диссертация

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
01.03.2022	Выдача задания	5
14.03.2022	Обсуждение структуры работы	5
01.04.2022	Обзор литературы	10
15.04.2022	Создания модели случайного графа	15
01.05.2022	Создание базовой версии модели	30
10.05.2022	Модификация модели	15
25.05.2022	Верификация модели	10
01.06.2022	Написание пояснительной записки ВКР	10

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семёнов М.Е.	к.ф.-м.н., доцент		

СОГЛАСОВАНО:

Руководитель ООП

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семёнов М.Е.	к.ф.-м.н., доцент		

Реферат

Выпускная квалификационная работа содержит 98 страницы, 18 рисунков, 12 таблиц, 44 источников, 3 формулы.

Ключевые слова: компартментальная модель, случайные графы, эпидемиологическая модель, коронавирус, клики.

Объектом исследования является проблема прогнозирования эпидемиологической обстановки по распространению коронавирусом инфекции Covid-19.

Целью работы является разработка и реализация SIR-подобной модели с применением сетей в виде программного продукта для прогнозирования распространения новой коронавирусной инфекции.

Область применения: эпидемиология.

В представленной работе разработана и реализована новая модель, позволяющая прогнозировать динамику распространения новой коронавирусной инфекции, с применением случайного графа с кликовой структурой. Модель реализована в виде программного продукта, который способен спрогнозировать эпидемиологическую обстановку в отдельном городе, используя статические данные по данному населенному пункту.

Оглавление

Введение.....	12
1. Обзор литературы.....	15
2. Применение графов.....	17
2.1 Основные понятия теории графов.....	17
2.2 Случайные графа Эрдеша-Реньи.....	18
2.3 Случайный граф с кликовой структурой.....	19
2.4 Анализ свойств случайного графа с кликовой структурой	20
3. Описание модели.....	26
3.1 Комpartmentальные модели.....	26
3.2 Базовая модель SEIRS	28
3.3 Модель SEIRS с иммунным статусом.....	30
3.4 Учёт ограничительных мер	31
3.5 Реализации модели в python	36
3.6 Верификация модели	37
Заключение	42
4.1 Концепция стартап-проекта	43
4.1 Проблема конечного потребителя.....	43
4.2 Способы защиты интеллектуальной собственности	44
4.3 Объем и емкость рынка	46
4.4 Современное состояние и перспективы отрасли	47
4.5 Конкурентные преимущества создаваемого продукта	48
4.6 Целевые сегменты потребителей создаваемого продукта.....	50
4.6.1 Налаживание контактов с заказчиками из госсектора.....	51
4.6.2 Требования для участия в государственных закупках.....	51
4.7 Цена программного продукта	52
4.8 Бизнес-модель проекта	55
4.8.1 Первоначальные инвестиции.....	56
4.8.2 Стоимость подписке к базе данных	58
4.9 Система налогообложения и оценка эффективности.....	59

4.10 План продаж	60
4.11 Заключение разделу	62
5.1 Социальная ответственность	63
5.1 Производственная безопасность	64
5.1.1 Отклонение показателей микроклимата в помещении	64
5.1.2 Превышение уровней шума	65
5.1.3 Повышенный уровень электромагнитных излучений	66
5.1.4 Поражение электрическим током	68
5.1.5 Освещённость.....	69
5.1.5 Пожарная опасность	73
5.2. Экологическая безопасность.....	75
5.3. Безопасность в чрезвычайных ситуациях	76
5.3. Заключение по разделу «Социальная ответственность»	77
Список публикаций студента.....	78
Список литературы	79
Приложение А	84

Введение

Пандемия новой коронавирусной инфекции Covid-19, начавшаяся в 2019 года, всё ещё продолжается. Волны пандемии следуют одной за другой. Появление вакцин не привело к окончанию пандемии, которая наносит значительный экономический ущерб. Особенно сильно уязвим человеческий капитал, воспроизводство которого требует больших вложений ресурсов и времени. Для эффективной борьбы с распространением коронавируса необходимы относительно точные прогнозы по эпидемиологической обстановке. Службы здравоохранения должны быть своевременно готовы к новой волне коронавируса. Запасы медикаментов, баллонов с кислородом, аппаратов искусственной вентиляции лёгких должно быть достаточно для преодоления пика заболеваний. Число сотрудников младшего, среднего и высшего медицинского персонала мобилизованных для борьбы с распространением заболевания также должно быть достаточным. Запасы коечного фонда должны слегка превышать потенциально-максимальное число одновременно больных, находящийся в тяжелом состоянии требующие стационарного лечения в стенах медицинских учреждений. Региональные органы власти и службы контроля должны своевременно вводить ограничительные меры для уменьшения скорости роста числа инфицированных. Все данные мероприятия являются очень дорогостоящими или наносят косвенный экономический ущерб (ограничительные меры). Поэтому объём данных мер должен быть обоснован для минимизации затрат. Например, квалифицированные медицинские сотрудники являются ограниченным ресурсом, который можно лишь перераспределить для решения вызовов перед здравоохранением. Невозможно за короткие сроки увеличить число врачей в стране. Поэтому очень важно на борьбу с пандемией выделять данный ресурс в минимально необходимых объёмах. Чем больше врачей направляются для борьбы с пандемией, тем менее доступной становится здравоохранение для многих членов общества, что может провоцировать излишнюю смертность по другим заболеваниям. Поэтому региональным органам власти, региональным министерствам здравоохранения и региональным

управлениям Роспотребнадзора необходимы точные прогнозы по распространению новой коронавирусной инфекции для рационального применения имеющихся ресурсов.

Компартментальные модели или SIR-подобные модели являются наиболее популярными моделями для прогнозирования распространения инфекций. Концепция данных моделей состоит в том, что всё население делится на группы, исходя из их инфекционного статуса. Например, в базовой модели SIR это восприимчивые (*Susceptible*), инфицированные (*Infected*) и переболевшие (*Recovered*). Название данных моделей формируются из первых букв возможных статусов. В данной работе предлагается объединить в единую математическую модель концепцию компартментальных моделей и случайных графов, имитирующие общественные связи.

Целью настоящей магистерской диссертации является разработка и реализация SIR-подобной модели с применением сетей в виде программного продукта для прогнозирования распространения новой коронавирусной инфекции. Для достижения данной цели были поставлены следующие задачи:

- Изучение существующих SIR-подобных моделей;
- Изучение существующих моделей случайных графов;
- Создание модели случайного графа с кликовой структурой;
- Создание математической модели SEIRS на основе случайного графа;
- Реализация модели SEIRS в Wolfram Mathematica и Python;
- Оптимизация программного кода для расчётов графа с большим числом вершин (более 1 млн.);
- Добавление механизмов ограничительных мер в базовую версию модели;
- Моделирование распространения коронавируса в различных городах и сопоставление полученных результатов с данными статистики (верификация);
- Анализ результатов и дальнейшая модернизация кода и модели.

Научная и практическая новизна и значимость работы. Применение в SIR-модели сетевой структуры на основе графа. Предложена модель случайного графа с кликовой структурой.

Результаты работы. В представленной работе разработана и реализована модель, позволяющая прогнозировать динамику распространения новой коронавирусной инфекции с применением сетевых данных. Модель реализована в виде программного продукта, который способен спрогнозировать эпидемиологическую обстановку в отдельном городе, используя статические данные по данному населенному пункту.

1. Обзор литературы

Существуют довольно много разных математических моделей [1-11], позволяющих прогнозировать и изучать процесс распространения эпидемий в человеческом социуме на конкретном этапе эпидемии или в срезе их эволюции. Разумеется, с появлением новой коронавирусной инфекции COVID-19 математики начали предлагать разные модели для описания процесса распространения данной инфекции. Первые модели начали появляться с момента первой вспышки в Китае. Пандемия ещё не закончилась, поэтому разработка и усовершенствование моделей также интенсивно продолжается.

В работах [12-15] использовалась модель логистического роста Ферхюльста или её модификации. Данная модель предполагает, что на начальном этапе распространения эпидемии рост числа заболевших происходит экспоненциально. Но из-за ограничения численности популяции скорость заражения падает, и число инфицированных выходит на плато.

Компартментальные модели (SIR-подобные) являются наиболее популярными для прогнозирования эпидемиологической обстановки: SIR [16-18], SIRS, SEIR, SAIR [19-22] и их модификации [23-30]. Практически во всех данных публикациях модели базируются на системе дифференциальных уравнениях. В основном данные модели различаются поправочными параметрами (для учёта специфики распространения коронавируса и фактор среды) и способами расчёта коэффициентов в уравнениях. Во всех работах прослеживается недостаток чистых компартментальных моделей, которые делают допущение, что среда распространения инфекции является однородной и каждый элемент связан с друг другом. В работе [21] авторы делают выводы, что компартментальные модели с постоянными коэффициентами плохо применимы для прогнозирования распространения COVID-19.

Российские учёные в своей работе [22] предлагают моделировать распространение коронавируса как суперпозицию множества волн на основе дискретного логистического уравнения Фейгенбаума. Авторы считают невозможным использовать для прогнозирования коронавируса модели, в

концепции которых существует лишь один пик заболеваемости. В работе предлагается методика определения локальных волн и расчёта их параметров. Добавление новых волн происходит, как только расчёты начинают значительно отклоняться от действительных значений. Авторы также подвергают критике модели, базирующиеся на статистических данных, так как в них изначально существует погрешность в 20-80%, связанных со сбором данных.

Бразильский математик в своей работе [23] предложил свою сетевую модель для прогнозирования коронавируса, которая похожа на нашу модель. В модели также предлагается считать вершины графа за отдельных индивидов, а ребра графа связями между ними. Модель также использует принципы SIR модели. Автор модели из Бразилии предлагает рассматривать не один граф, а несколько, каждый из которых описывает связи разного уровня: семья, работа, транспорт и др. Это отличается от концепции нашей модели, которая базируется на одном едином графе. Автор работы также отмечает недостаток сетевого подхода модели – необходимость больших вычислительных мощностей. В работе он производил расчёты с применением графов, число вершин которых не превышает 100000. Для масштабирования прогнозов на большие популяции в работе предлагается умножать полученные результаты на коэффициенты, которые равны отношению результатов моделирования для малой популяции к данным статистики для большой популяции.

2. Применение графов

2.1 Основные понятия теории графов

Граф – математический объект, состоящий из двух множеств. Одно из них – любое конечное множество, его элементы называются вершинами графа. Другое множество состоит из пар вершин, эти пары называются ребрами графа. Если множество вершин графа обозначено буквой V , множество ребер – буквой E , а сам граф – буквой G , то пишут $G = (V, E)$. Для обозначения числа вершин и числа ребер графа обычно используют буквы n и m [24].

Говорят, что ребро (a, b) соединяет вершины a и b . В графе может быть не более одного ребра, соединяющего две данные вершины. Ребро типа (a, a) , т.е. соединяющее вершину с ней же самой, называют петлей. Иногда петли разрешаются, иногда запрещаются. В последнем случае говорят, что рассматриваются графы без петель. Если ребра – упорядоченные пары $(a, b) \neq (b, a)$, то такой граф называется ориентированным (сокращенно орграф), если же неупорядоченные $(a, b) \equiv (b, a)$, то неориентированным. Если в графе есть ребро (a, b) , то говорят, что вершины a и b в нем смежны. Ребро $e = (a, b)$ инцидентно каждой из вершин a и b , а каждая из этих вершин инцидентна ребру e [24].

Существует много способов представить граф, назовем только самые распространенные. Все они иллюстрируются на примере одного графа:

1. Перечисление элементов. Исходя из определения, для того, чтобы задать граф, достаточно перечислить его вершины и ребра (т.е. пары вершин) $V = \{a, b, c, d\}$; $E = \{(a, b), (a, c), (b, d)\}$;
2. Изображение. Если граф не слишком большой, его можно нарисовать. В неориентированном графе ребра изображают линиями, соединяющими смежные вершины, в ориентированном – стрелками.
3. Матрица смежности. Пусть G – граф с n вершинами, пронумерованными числами от 1 до n . Матрица смежности – это таблица с n строками и n столбцами, в которой элемент, стоящий на пересечении строки с номером i и столбца с номером j , равен 1, если вершины с номерами i и j смежны, и 0, если они не смежны;

4. Матрица инцидентности. Пусть G – граф, вершины которого пронумерованы числами от 1 до n , а ребра – числами от 1 до m . В матрице инцидентности строки соответствуют вершинам, а столбцы – ребрам. На пересечении строки с номером i и столбца с номером j стоит 1, если вершина с номером i инцидентна ребру с номером j , и 0 в противном случае;

5. Списки смежности. Этот способ часто используется для компьютерного представления графов. Состоит он в том, что для каждой вершины задается список всех смежных с ней вершин. В структурах данных, применяемых в программировании, списки смежности могут быть реализованы как массив линейных списков или как словари.

Полный граф – граф, в котором каждые две вершины смежны. Подграф – это граф состоящий из множества вершин и ребер, являющихся подмножествами вершин и ребер основного графа. Из любого графа не пустого графа можно выделить подграф путём удаления вершин и ребер. Кликкой графа называется множество вершин, порождающее полный подграф, т.е. множество вершин, каждые две из которых смежны. Число вершин в клике наибольшего размера называется кликовым числом графа [24].

2.2 Случайные графа Эрдеша-Реньи

Под случайными графами подразумевают модели реальных сетей, которые позволяют описывать свойства и воспроизводить структура этих сетей. Исходя из свойств реальной структуры подбирается случайный граф наиболее подходящий. В данной работе, необходимо было подобрать модель графа, которая наиболее хорошо воспроизводит общественные связи с точки зрения распространения инфекций.

Случайные графы Эрдеша-Реньи [26] образуются следующим образом:

- фиксируют N вершин графа;
- любую пару вершин случайно, с равной вероятностью p , связывают ребром.

2.3 Случайный граф с кликовой структурой.

Модель Эрдеша-Реньи плохо отражает структура общества, а именно неоднородность связей. В графе Эрдеша-Реньи связи распределены равномерно. В реальной структуре общества имеются социальные ячейки (семья или домохозяйство). Внутри семей имеем более высокую плотность связей. Каждый индивид связан между собой. Если представить домохозяйство как подграф, то это будет клика по определению. Таким образом для воссоздания структуры общества необходимо генерировать граф, каждая вершина которой является частью клики. В данной работе кликой будем считать полный подграф с 2 и более вершин. Средняя статистическая семья, состоящая, из 2 родителей и 2 детей, может быть представлена кликой размерностью 4. Каждый индивид кроме внутренних связей в семье, разумеется, имеет постоянные внешние. Чаще всего это связи с друзьями, другими родственниками и коллегами. Для описания данных связей хорошо подходит модель Эрдеша-Реньи. Поэтому было предложено взять случайный граф Эрдеша-Реньи за основу и достроить в нём клики.

За основу берется случайный граф Эрдеша-Реньи G , который генерируется случайно с N вершинами и $d(m) \cdot N/2$ рёбрами, где $d(m)$ – средняя степень вершины m . Далее данный граф изменяется через цикл, который достраивает клики в нём. В ходе цикла происходит генерация списка ребер для формирования клик. На каждой итерации генерируется случайное число $k \sim U(2, max)$, вершины с номерами $i, i+1, \dots, i+k$ замыкаются в клику на k вершинах. Итерации происходят до тех пор пока из оставшихся вершин можно сформировать клику указанного размера, т.е. пока $N-i > max$, где max – кликовое число, которое задается перед генерацией графа. После выхода из цикла оставшиеся вершины также достраиваются до клики. На выходе из цикла получаем случайный граф с кликами, которые располагаются по главной диагонали матрицы смежности вершин. Случайное число k может генерироваться по любому закону распределения. В простом случае имеет равномерное распределение числа вершин клик в графе от 2 до max .

2.4 Анализ свойств случайного графа с кликовой структурой

Как уже отмечалось предложенная нами модель случайного графа с кликовой структурой базируется на графе Эрдеша-Реньи, однако ключевые параметры графов отличны. Поэтому проведём сравнение изначального графа Эрдеша-Реньи и графа с достроенными кликами по следующим характеристикам:

- количество ребер;
- средняя степень вершин;
- диаметр графа;
- коэффициент кластеризации.

Чтобы каждая вершина случайного графа являлась элементом клики, происходит достраивание рёбер для формирования полных подграфов. Поэтому конечный граф должен иметь больше рёбер, чем изначальный граф. Зададим случайный граф Эрдеша-Реньи с $N=2000$ вершин и $N_e=10000$ рёбер и сопоставим с графом с кликовой структурой, которые базируется на данном графе. На рис. 1 представлено зависимость числа рёбер от максимально значения кликового числа случайного графа (параметр max). Наблюдается линейная зависимость числа рёбер от параметра max .

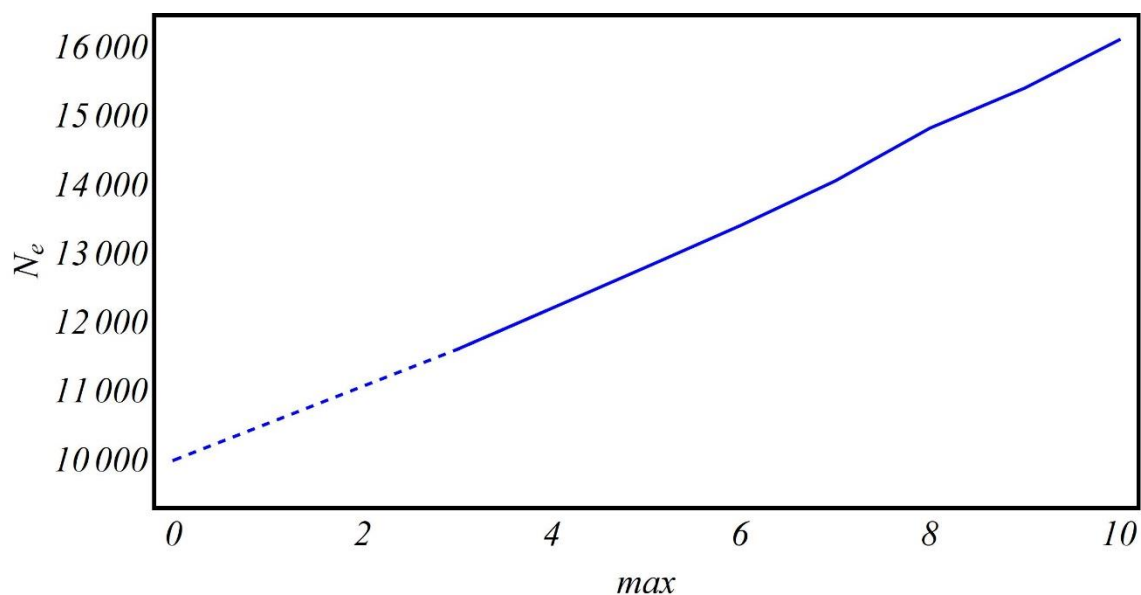


Рисунок. 1. Зависимость числа рёбер от значения максимального кликового числа

Скорость роста числа новых рёбер при формировании случайного графа зависит непосредственно от средней степени вершин (при $n = const$). Чем меньше изначально рёбер в графе, тем больше рёбер необходимо достроить для формирования клик. На рис. 2 подтверждается данная гипотеза. По оси ординат располагается отношение числа рёбер в графе с кликами к числу рёбер изначального графа Эрдеша-Реньи. Как видим, в относительных значениях также сохранится линейный рост с увеличением клейкого числа. Рисунок 2 прекрасно демонстрирует, что угол наклона линейной функции (скорость роста числа рёбер) непосредственно зависит от средней степени верши $d(m)$ в изначальном графе.

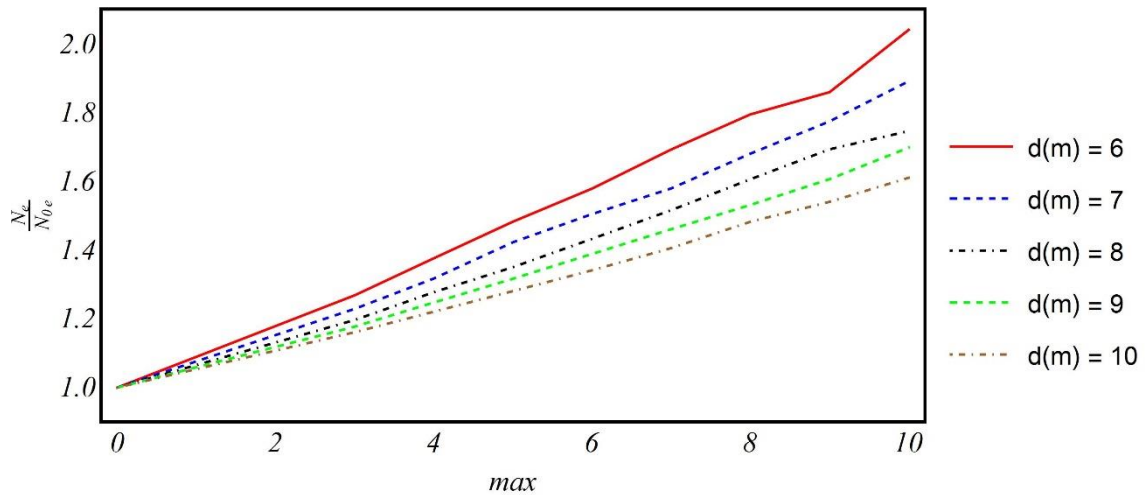


Рисунок. 2. Динамика изменения числа рёбер при увеличении максимального кликового числа

Среднюю степень вершин графа можно определить как среднее арифметическое степени каждой вершины графа:

$$d(m) = \frac{1}{n} \sum_{i=1}^N d(v_i) \quad (1)$$

Однако легче воспользоваться теоремой о рукопожатиях [24], согласно которому сумма степеней вершин равняется удвоенному числу рёбер. Тогда формула 1 принимает вид:

$$d(m) = \frac{1}{n} \sum_{i=1}^n d(v_i) = \frac{2m}{n} \quad (2)$$

При достраивании клик в случайном графе Эрдеша-Реньи количество вершин n остается неизменным, увеличивается лишь число рёбер m . Согласно формулы 2 $d(m)$ пропорционально m , поэтому характер изменения средней степени вершин графа будет таким же, как для числа рёбер, и выше сделанные выводы можно распространить и на данный параметр. График на рис. 3 подтверждает данный вывод.

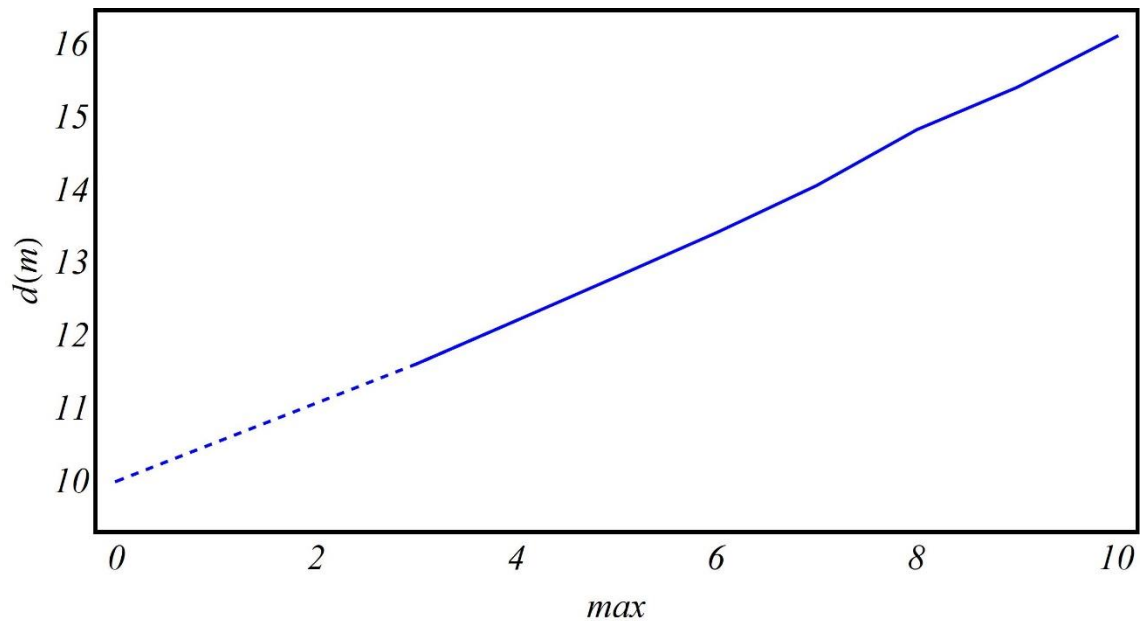


Рисунок. 3. Зависимость средней степени вершин графа от значения максимального кликового числа

В теории графов коэффициент кластеризации характеризует степень сгруппированность узлов графа. Тем выше данный параметр, тем больше число подграфов близких к кликам. Факты свидетельствуют о том, что в большинстве реальных сетей, и в частности в социальных сетях, узлы, как правило, создают тесно сплоченные группы, характеризующиеся относительно высокой плотностью связей. Эта вероятность как правило превышает среднюю вероятность случайного установления связи между двумя узлами. Как раз по

данному параметру случайный граф Эрдеша-Реньи не подходит для описания реальных сетей.

Коэффициент кластеризации можно рассчитать на двух уровнях: на локальном и глобальном. Локальный коэффициент кластеризации оценивает плотность связей в окрестности конкретной вершины. Напомним, что окрестностью вершины называет множество вершин смежных ей. Локальный коэффициент кластеризации рассчитывается, как отношение числа рёбер в окрестности вершины к максимально возможному числу рёбер в данной же окрестности. Максимальное число рёбер между n вершинами равняется $n(n-1)/2$ (число рёбер полного графа). На рис. 4 представлен пример расчёта локального коэффициента кластеризации для красной вершины при разном числе рёбер в окрестности. Если все смежные вершины соединены между собой, данный параметр равен 1, а если вовсе отсутствуют связи, то равен 0.

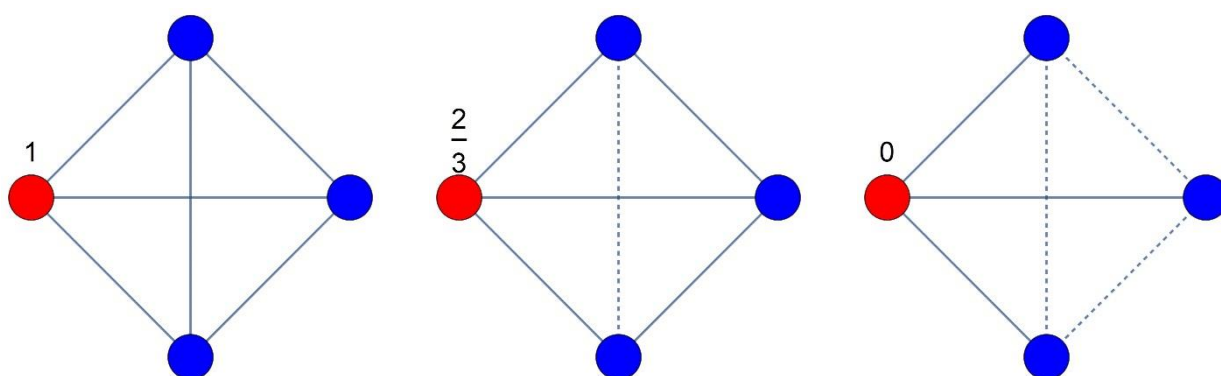


Рисунок. 4. Значения локального коэффициента кластеризации для красной вершины

Глобальный коэффициент кластеризации рассчитывается через триплеты узлов (просто триплеты или тройки). Триплет – это соединённые 3 вершины. Тройка может быть замкнутой (есть 3 ребра) и открытой (2 ребра). Глобальный коэффициент кластеризации равен отношению числа замкнутых триплетов в графе к общему числу триплетов (открытых и замкнутых). Нахождение триплетов – это довольно сложная математическая задача. Поэтому часто глобальный коэффициент кластеризации заменяют на усреднённое значение

локальных коэффициентов. Эти параметры не совсем равноценны, так как в усреднённом коэффициенте кластеризации больший вес имеют вершины с низкой степенью, в отличие от глобального коэффициента, где больший вклад вносят вершины с высокой степенью. В предложенной нами модели случайного графа с кликовой структурой разброс степеней вершин не такой значителен (не равен нескольким порядкам), поэтому для сравнительного анализа можем воспользоваться усреднённым значением локальных коэффициентов кластеризации.

Были рассчитаны коэффициенты кластеризации случайного графа с кликовой структурой при различных значениях кликового числа. Полученные данные были соотнесены с коэффициентом кластеризации изначального графа Эрдеша-Реньи. На рис. 5 по оси ординат представлено отношение коэффициентов кластеризации графа с кликами и Эрдеша-Реньи. Данное отношение показывает насколько выше кластеризация предложенного нами случайного графа по отношению к графу Эрдеша-Реньи. По графику видно, что при максимальном кликовом числе равном 5 коэффициент кластеризации возрастает на порядок. Коэффициент кластеризации прямо пропорционален кликовому числу.

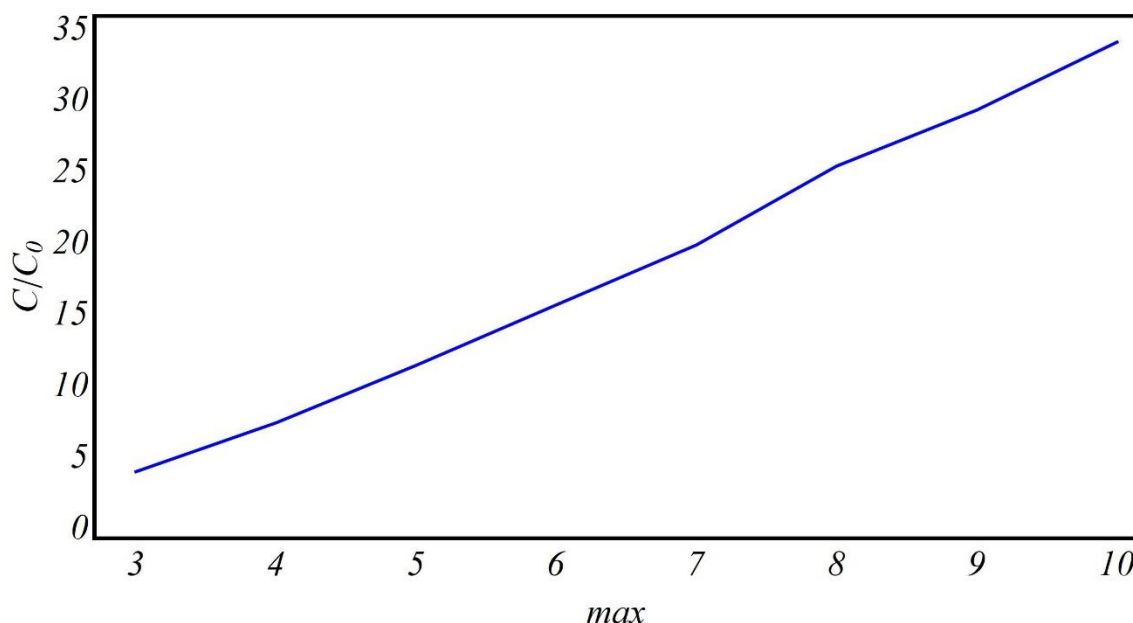


Рисунок. 5. Зависимость кратности увеличения степени кластеризации от значения максимального кликового числа

При расчётах диаметров случайных графов с кликовой структурой с разным числом вершин было обнаружено, что данное значение не зависит от числа вершин и в большинстве соответствует 6. Что на удивление, хорошо согласуется с «Теорией о шести рукопожатиях», согласно которой любой индивид на Земле связан с другим индивидом через число людей, которое не превышает 5.

На основе провиденного анализа можно заключить, что модель случайного графа с кликовой структурой имеет свойства близкие к реальной структуре общественных связей, поэтому является более оптимально для нашей модели, чем граф Эрдеша-Реньи.

3. Описание модели

3.1 Комpartmentальные модели

Комpartmentальные модели — это методика, используемая для упрощения математического моделирования инфекционных заболеваний. Принцип данных моделей заключается в том, что вся популяция делается на группы исходя из своего эпидемиологического статуса. Для удобства каждая группа обозначается метками – например, S, I или R.

SIR–подобные модели получили заслуженную популярность в силу простоты построения и использования. Их применение позволяет точно моделировать эпидемии гриппа и других заболеваний в больших городах, вводить новые параметры и анализировать разные сценарии. В основе данных лежит разделение всей популяции на группы, исходя из инфекционного статуса. В модели SEIRS всё население разделено на 4 группы: восприимчивые, контактные, инфицированные и выздоровевшие [27]. Поэтому в модели 4 ключевых параметра:

- $S(t)$ — численность восприимчивых индивидов в момент времени t ;
- $E(t)$ — численность контактных (находящихся в инкубационном периоде) индивидов в момент времени t ;
- $I(t)$ — численность инфицированных (с подтверждённым заболеванием) индивидов в момент времени t ;
- $R(t)$ — численность переболевших (имеющие иммунитет после болезни) индивидов в момент времени t .

На элементарном уровне данную модель можно реализовать классическим способом как система дифференциальных уравнений:

$$\begin{aligned}\frac{dS}{dt} &= -\beta \frac{S}{N} E + \sigma R \\ \frac{dE}{dt} &= -\beta \frac{I}{N} S - \alpha E \\ \frac{dI}{dt} &= \alpha E - \gamma I \\ \frac{dR}{dt} &= \gamma I - \sigma R\end{aligned}\tag{3}$$

где β — коэффициент интенсивности контактов индивидов с последующим инфицированием; γ — коэффициент интенсивности выздоровления инфицированных индивидов, величина, α — величина обратная среднему инкубационному периоду заболевания, σ — величина обратная средней длительности действия иммунитета.

Первое уравнение системы означает, что изменение числа здоровых (и при этом восприимчивых к заболеванию) индивидуумов уменьшается со временем пропорционально числу контактов с инфицированными и увеличивается за счёт тех, у кого истёк действие иммунитета. После контакта происходит заражение, восприимчивый переходит в состояние контактных. Контактные являются главными переносчиками заболевания. Второе уравнение показывает, что скорость увеличения числа заразившихся растёт пропорционально числу контактов здоровых и контактных и уменьшается по мере перехода последних в статус открыто болеющих. Третье уравнение демонстрирует, что число подтверждённых больных пропорционально числу контактах и уменьшается по мере выздоровления. Четвёртое уравнение демонстрирует, что число выздоровевших в единицу времени пропорционально числу инфицированных и уменьшается с течением времени из-за прекращения действия иммунитета. Иначе говоря, каждый заболевший через некоторое время должен поправиться. Таким образом, заболевание в модели SEIRS развивается по схеме: восприимчивые – контактные – инфицированные – выздоровевшие – восприимчивые.

Главным преимуществом реализации SIR-подобных моделей через систему дифференциальных уравнений является простота реализации расчётов. Расчёты легко масштабируются на популяцию любой размерности: тысячи, миллионы или миллиарды. Это не влияет на производительность расчётов. Однако за простоту и универсальность приходится платить. Крайне сложно учесть специфику социума конкретной области. Фактически для модели нет различия между город в Европе и городом в Индии или США. Сложно через систему дифференциальных уравнений учесть половозрастной состав населения.

Учесть мобильность населения. Сложно такие различные процессы заложить в коэффициентах. Поэтому было и предложено использовать функциональные возможности графов с концепцией SIR-моделей.

3.2 Базовая модель SEIRS

Предлагаемая SEIRS модель базируется на концепциях SIR-подобных моделей. В основе модели лежит граф G , который генерируется выше описанным алгоритмом с n вершинами и $d(m) \cdot n/2$ рёбрами, где $d(m)$ – средняя степень вершины m . Вершина графа сопоставлена представителю популяции, которая в целом имитируется графом. Ребра графа характеризуют устойчивые (частые) связи между людьми в обществе. Каждая вершина имеет категориальный статус: *восприимчивый*, *инфицированный* или *переболевший*. Для визуализации статуса вершины мы использовали цвета: синий, красный и зелёный соответственно (рис. 5).

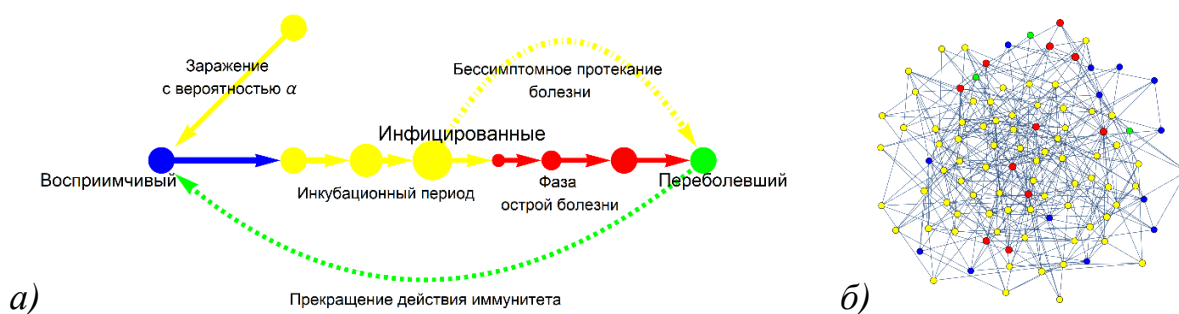


Рисунок. 5. (а) Общая схема модели SEIRS, (б) пример графа G при $n=100, d(m)=5$

В классической модели SIR люди со статусом *восприимчивые* при контакте с *инфицированными* также могут заболеть, и наоборот – инфицированные могут выздороветь. В предложенной нами модели, если синяя вершина является смежной красной, то она может быть инфицированной с вероятностью $0 \leq P(\alpha) \leq 1$ за одни сутки. Второй важной характеристикой вершины является её вес. Вес синих вершин постоянен и равен нулю. У других вершин вес динамически изменяется и равен количеству суток, которое они прибывают в определенном статусе. Например, если вес красной вершины равен 7, то данный индивид инфицирован уже в течении 7 суток. Подсчёт количества суток

необходим для учёта выздоровления. Вес красных вершин сопоставлен со средней продолжительностью болезни - 21 суток. Когда продолжительность болезни достигает значения средней продолжительности болезни, индивид выздоравливает и цвет вершины изменяется на зелёный.

Модель SIR довольно ограниченная, существует мало инфекций, которые можно моделировать с помощью этой модели. Однако, SIR имеет множество производных моделей, которые дополняют её функциональные возможности. Одной из таких моделей является SEIR, в которой используется ещё один статус людей – *контактный* (exposed). В этой модели водится важное понятие как инкубационный период - промежуток времени от момента инфицирования до появления открытых, явных симптомов. Часто в инкубационный период человек может уже заражать (как в случае Covid-19), что сильно осложняет своевременное отчуждение больного и контроль распространения инфекции становится затруднительным. В нашей модели мы предлагаем добавить новый статус, характеризующих больных в инкубационном периоде (жёлтые вершины, рис. 5б). В этом случае характеристики жёлтых вершин очень схожа с красными. Ключевым отличием является, то что после достижения жёлтой вершиной веса равного инкубационному периоду - 5 суток - они изменяют статус на *инфицированный* (красный), а не на *переболевший* (зелёный). Также в предлагаемой модели было введено следующее условие: заражать могут только *контактные* (жёлтые), тем самым имитируются процесс изоляции больного с открытыми симптомами. Далее в предлагаемой модели мы ввели параметр, характеризующий долю популяции, которые переболеют полностью в скрытой форме (20% популяции). Модель SIRS является развитием SIR модели, в которой введено временное ограничение на действия иммунитета – средняя длительность действия иммунитета принята равной 700 суток, т.е. в модели SIRS выздоровевшие через некоторое время могут снова стать восприимчивыми к болезни. Это очень характерно для гриппа, а возможно и для Covid-19. В предлагаемой модели мы использовали указанную особенность SIRS модели и ввели пересчет веса зелёных вершин: по достижению веса средней длительности

действия иммунитета цвет вершины меняется с зелёного на синий, вес вершины принимается равным нулю. Таким образом, мы получаем цикл из возможных статусов индивида, что позволяет учитывать возможную повторность заболевания при моделировании.

3.3 Модель SEIRS с иммунным статусом

В базовой модели SEIRS переболевшие не могут заразиться, пока не пройдет время равное длительности иммунитета. В реальности всегда есть вероятность повторного заражения коронавирусом спустя небольшой интервал времени после выздоровления. Поэтому было предложено вести функцию иммунитета, которая зависит от времени и принимает значения от 0 до 1. Функция характеризует вероятность заражения, и её можно задавать через линейную функцию:

$$f_{im}(t) = \begin{cases} 1 - \frac{t - t_0}{T}, & \text{при } t - t_0 \leq T \\ 0, & \text{при } t - t_0 > T \end{cases} \quad (3)$$

Где t_0 – день выздоровления, T – длительность действия иммунитета. Время выздоровления вносится как один из атрибутов вершин графа. С учётом иммунитета вероятность заражения вершины теперь равняется $P(\alpha) (1 - f_{im}(t))$. С добавлением функции иммунитета восприимчивые и переболевшие стали отличаться лишь значением атрибута t_0 , поэтому из названия модели можно убрать R – SEIS.

С данной версией модели появился смысл пропускать статические данные через граф, чтобы характеристики сети стали наиболее близки к реальным данным. Также такое «обучение» графа позволяет рассчитать значения параметра α . Фактически так решаем обратную задачу. Параметр α можно посчитать как отношение числа инфицированных за сутки к числу возможных инфицированных. На рис. 6 представлены рассчитанные значения данного параметра во времени, согласно статистики по городу Томск с момента первых заражений. Как видим в графике имеется разброс значений, который в первую

очередь обусловлен погрешностью данных статистики. Особенная большая погрешность на начальном этапе пандемии, из-за отсутствия методики и тестов для быстрой диагностики коронавируса. На графике также виден последний всплеск заразности коронавируса, обусловленный новым штаммом омикрон.

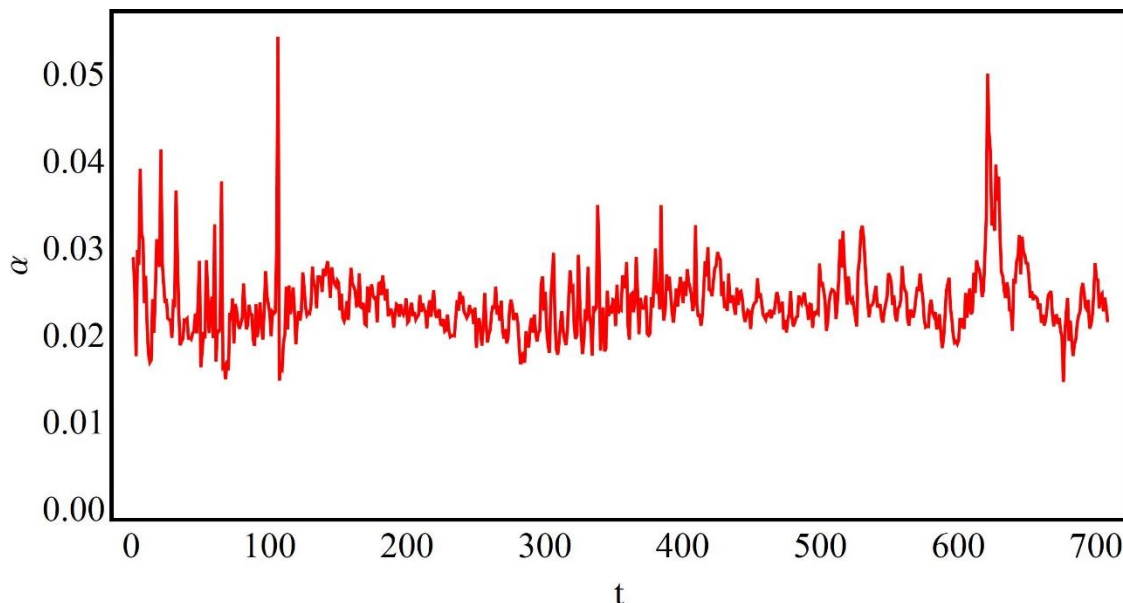


Рисунок. 6. Рассчитанные значения параметра α , исходя из статических данных по городу Томск

3.4 Учёт ограничительных мер

В классическом варианте SIR-моделей вирус распространяется в относительно идеальных условиях без ограничителей, что конечно далеко от реальности. В нашей модели было предложено использовать несколько механизмов имитирующие те или иные инструменты ограничения распространения инфекции.

Для учёта мероприятий, влияющих на вероятность заражения, на начальном этапе развития модели был введён коэффициент ограничительных мер, равный отношению вероятности заражения с учётом мероприятия к вероятности заражения без учёта мероприятия. Данный коэффициент меньше 1, если мероприятие ограничивает распространение, и больше 1 в противном случае. Эксперт исходя из эмпирических наблюдений и имеющихся данных статистики даёт оценку степени влияния тех или иных ограничительных мер и событий на

вероятность заражения индивида. На рис. 7а представлено сопоставление динамики зараженных $I(t)$ в городе Томск согласно официальной статистики и модельных результатов с применением данного коэффициента ограничительных мер, значение которых представлены на графике. В целом наблюдается хорошая сходимость результатов моделирования динамики инфицирования со статистикой. Расхождение в начальный период (месяц - май) обусловлено отсутствием методов массового тестирования на коронавирус в начале пандемии. В летний период (июнь-август) данные официальной статистики превышают результаты. На рисунке представлены данные по Томску на начальном этапе пандемии, когда начались применять жёсткие ограничительные меры. Значения коэффициента подбирались, исходя из эмпирической оценки наблюдателя происходящих событий.

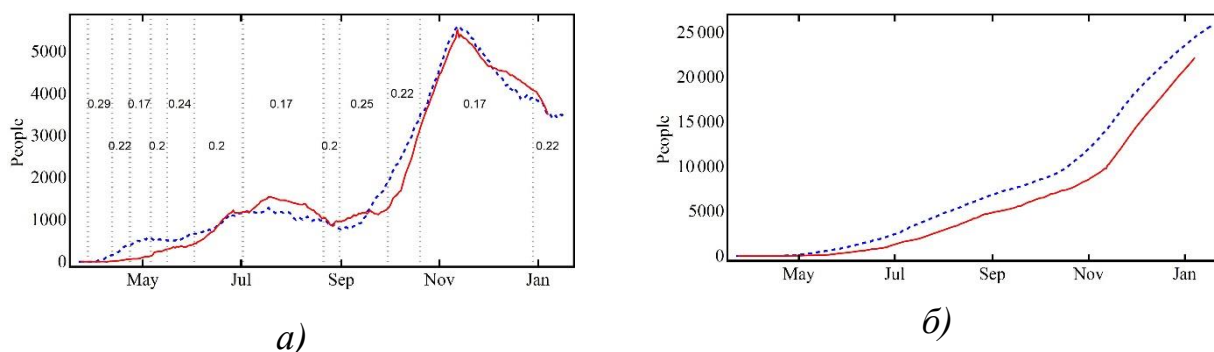
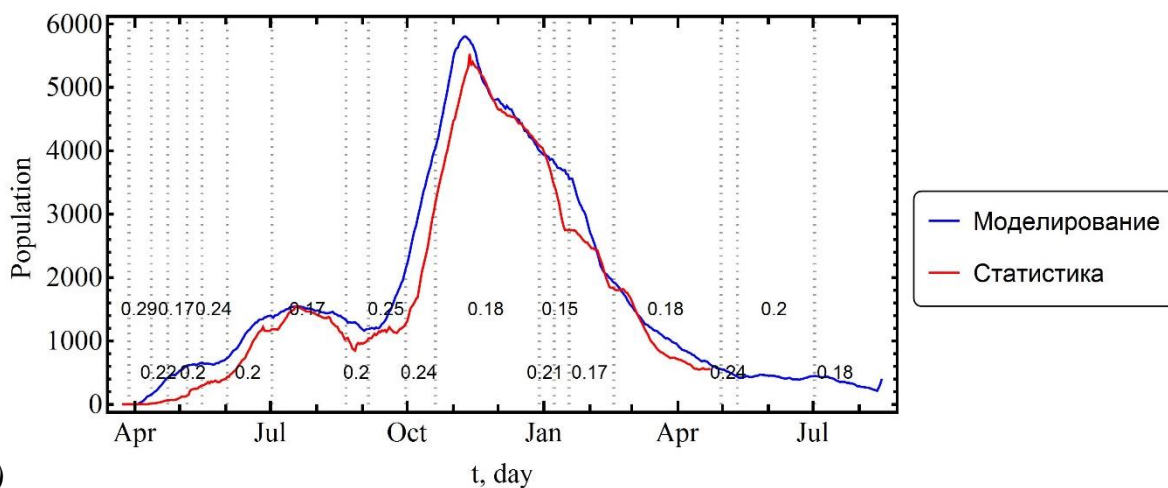
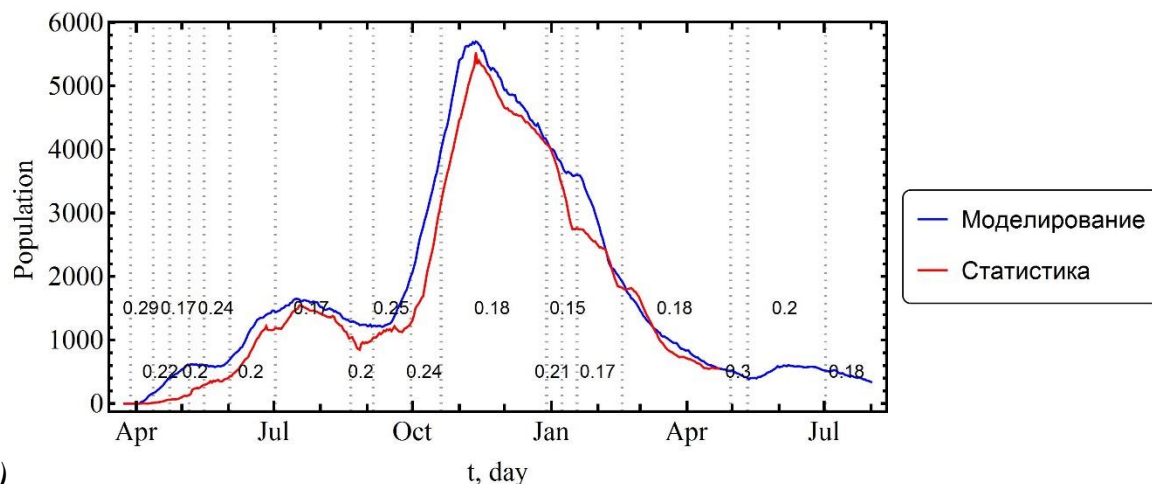


Рисунок. 7. Динамика изменения (а) числа инфицированных и (б) выздоровевших в Томске с марта 2020 г. по январь 2021 г.: официальные данные (красная линия) и результатам моделирования (пунктирная линия)

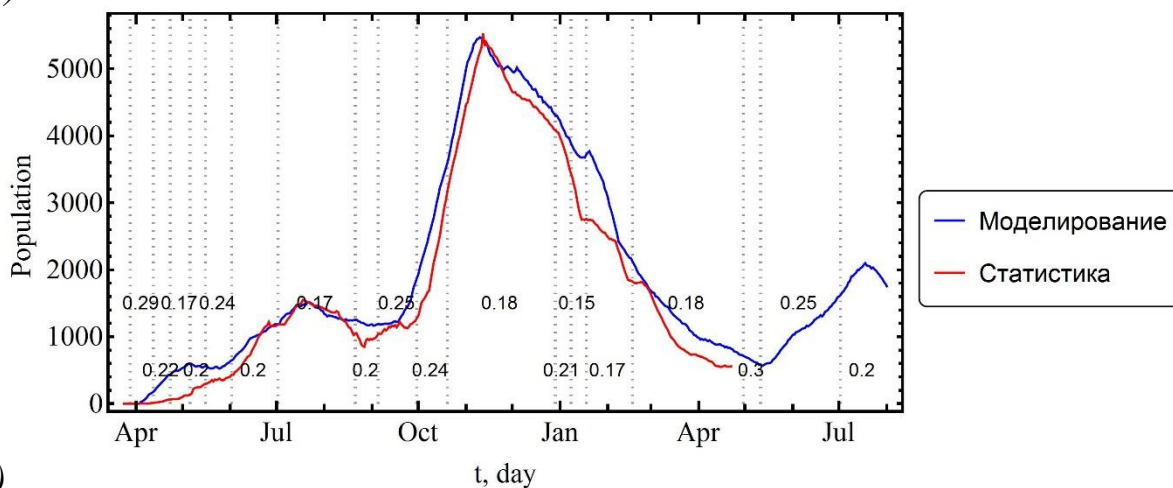
В апреле 2021 года были сделаны прогнозы по динамике инфицированных с использованием коэффициента ограничительных мер (см. рис.8). Были сделаны 3 предположения: ограничительные меры ужесточат, оставят без изменений или ослабят. «Пессимистический» вариант в конечном итоге оказался наиболее близок к действительности, так как ограничения были смягчены, а тёплая погода способствовала нарушению оставшихся.



a)



б)



в)

Рисунок 8. Прогнозы по заболеваемости на май 2021 года: оптимистичный (а), умеренный (б) и пессимистичный (в)

Продемонстрированные результаты моделирования показывают, что коэффициент ограничительных мер можно применять для учёта мероприятий и событий, влияющих на распространение коронавируса. Однако данный метод учёта ограничений имеет большой субъективный характер и сильно зависит от

предположений человека, использующего модель. Поэтому было принято решение отойти от данного способа и внести в модель другие механизмы учёта ограничений распространения коронавируса.

Первым барьером, ограничивающий распространение инфекции является бдительность и осторожность индивида. Для учёта бдительности индивидов был добавлен фактор окружения – μ . Данный фактор равен доли смежных вершин, которые официально не инфицированы (имеют статус 0 или 1). Вероятность заражения в этом случае равняется: $P(\alpha\mu) = P(\alpha)P(\mu)$. На рис. 9 представлены примеры расчёта значения фактора окружения в зависимости от инфекционного статуса смежных вершин.

Таким образом мы имитируем бдительность здоровой вершины.

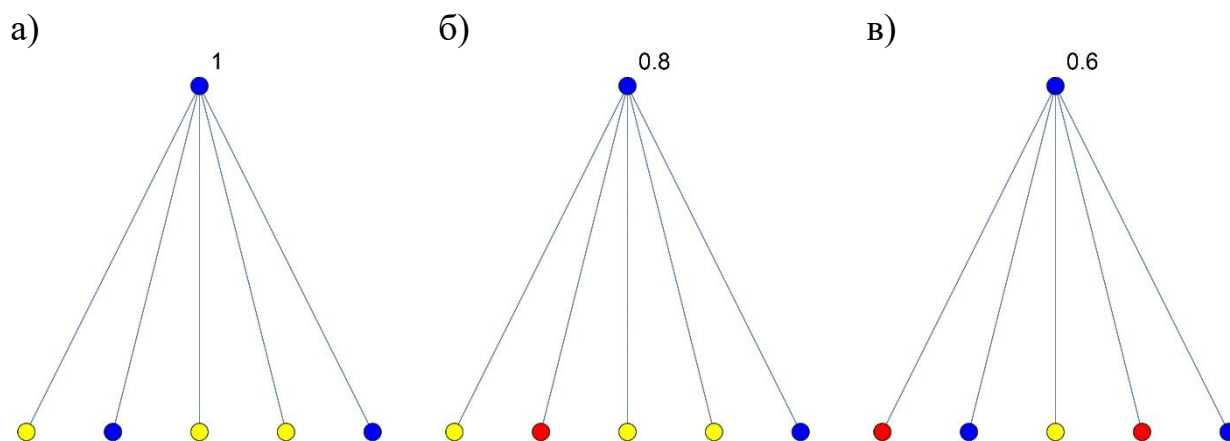


Рисунок 9. – Примеры значений фактора окружения вершины в зависимости от окружения (красный – инфицированный, синий – здоровый, жёлтый – инфицированный в инкубационном периоде)

На рис. 10 представлено сравнение результатов расчётов динамики распространения коронавируса при учёте и не учёте фактора окружения. При этом другие отрицательные факторы не учитывались. Моделирование производилось на идентичных графах. Как видно фактор окружения значительно снижает скорость распространения на начальных этапах волны пандемии. Однако при больших значениях заболеваемости прекращает оказывать значительное влияние.

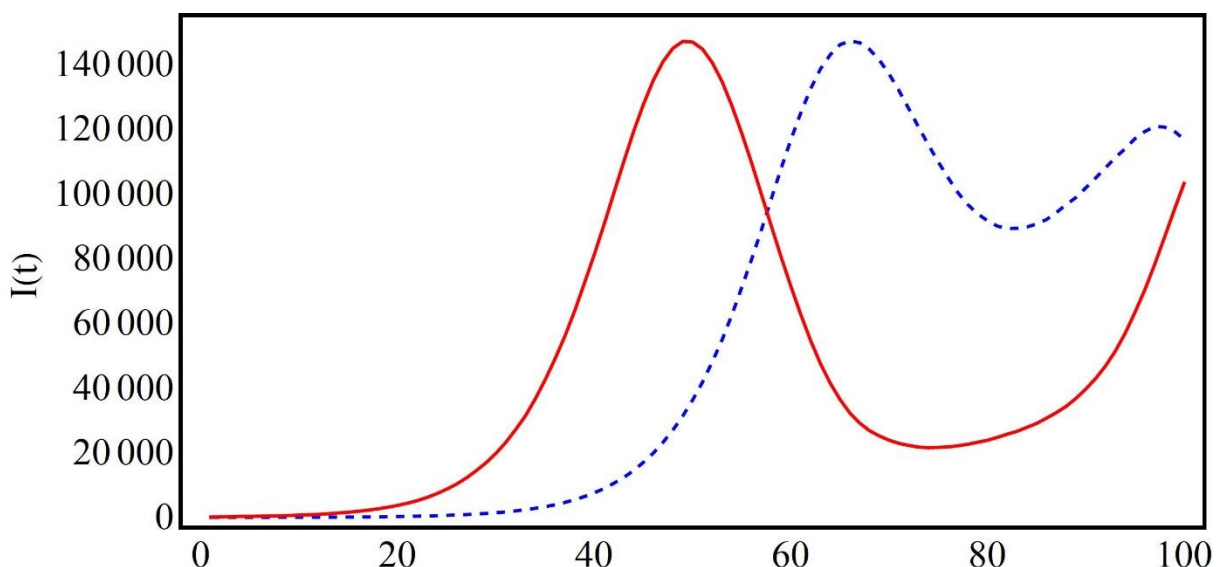


Рисунок 10. Динамика числа инфицированных без учёта (красная линия) и с учётом (пунктирная линия) фактора окружения в модели

На распространение инфекции значительно влияет эффективность выявления больных. В нашей модели, когда вершина переходит в стадию острой болезни, то происходит проверка всех контактировавших с инфицированным. И с вероятностью $P(\beta)$ могут быть определены те, кто инфицирован, но ещё в инкубационном периоде (вершины со статусом 1). Данный параметр, как и фактор окружения прекрасно демонстрируют, какие инструментальные возможности имеет модель с применением сетевых структур данных.

На рис. 11 приведено сравнение результатов расчётов с при различных значениях эффективности выявления контактных (вершины со статусом 1) β . Как видно, эффективное выявление контактах значительно затрудняет распространение инфекции. Это один способов борьбы с пандемией. Такой путь продемонстрировал Китай, которые имеет одни из самых низких показатель заражения.

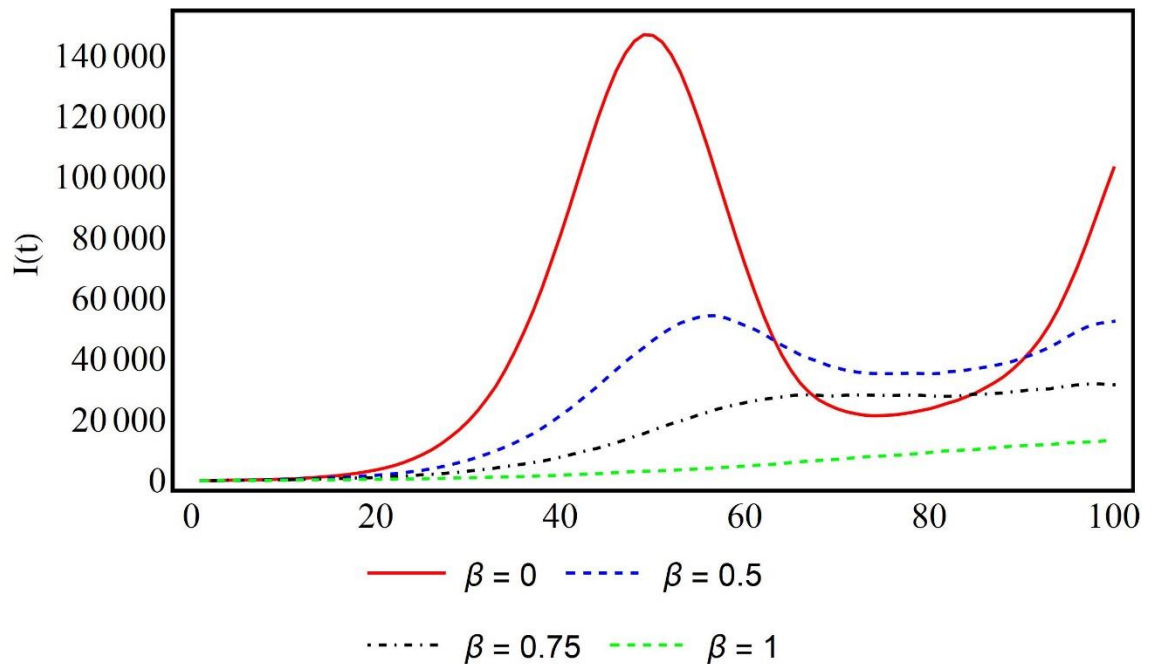


Рисунок 11. Динамика числа инфицированных при различных значениях эффективности выявления заразившихся

С появлением вакцин, они стали новым барьером в распространении коронавируса. С ростом числа вакцинированных и переболевших возникает эффект коллективного иммунитета, при котором вирусу сложно передаться восприимчивым. Учесть долю вакцинированных можно 2 способами. Можно непосредственно задать процент вакцинированных при генерации случайного графа. Или пропустить через граф статистику вакцинации, что дает сеть, иммунный статус которого наиболее близок к действительности.

3.5 Реализации модели в python

Для реализации модели SEIRS в Python было предложено использовать библиотеку NetworkX [28]. Библиотека NetworkX предназначена для создания, манипуляции и изучения структуры, динамики и функционирования сложных сетевых структур. библиотека свободно может оперировать весьма большими сетевыми структурами, уровня графа с 10 миллионами узлов и 100 миллионами дуг между ними. В виду того, что он базируется на низкоуровневой структуре данных языка Python под названием «словарь-словарей», память расходуется эффективно, графы хорошо масштабируются, мало зависят от особенностей операционной системы в которой выполняется скрипт и отлично

подходят для популярного на данный момент направления по анализу данных из социальных сетей и графов.

В NetworkX граф задаётся из себя словарём вершин и словарём рёбра. Сам граф, каждая вершина и каждое ребро в словаре может иметь сколь угодно количество атрибутов. Например, вес и цвет вершина. Зная имя вершины (чаще номер) можно узнать значения нужного атрибута и его заменить. Реализация модели базируются на данной функциональной возможности. Каждая вершина графа в модели имеет атрибут, значения которого отвечает за инфекционный статус. Для удобства значения данного атрибута взяты не в виде букв аббревиатуры названия модели, а использовались числа: 0 (S), 1 (E), 2 (I), 3 (R). Вторым важным атрибутом вершин графа является время нахождения в данном состоянии. При переходе из одного состояния в другое данный атрибут обнуляется. После одной итерации (равный одному дню) увеличивается на единицу.

3.6 Верификация модели

Для проверки работоспособности и точности модели было произведено моделирование распространения коронавируса для Томской и Новосибирской агломерации. Для моделирования Томской агломерации использовался граф с параметрами: $n = 800000$, $d(m) = 10$ и $max = 6$. Для Новосибирской агломерации: $n = 2200000$, $d(m) = 10$ и $max = 6$. Каждый граф был обучен на статистических данных по заболеваемости до точки начала моделирования. Это позволило оценить значение параметра $P(\alpha) = 0.028$. Статистические данные, на которых обучаются сети, вносят различие между графами разных городов. На рис. 12 представлены сравнение результатов моделирования и статистики. Данные статистики были взяты с [29]. Как видно результаты моделирования качественно совпадают на начальном этапе волны. Расположение максимумов статистики и расчётов примерно одинаково. Максимальные количественные значения инфицированных в статистике и в моделировании не сильно отличаются и лежат в пределах погрешности статистики. На спаде волны наблюдаем расхождение

статистики и моделирование. Статистическая кривая более быстро спадает. Предположительно это обусловлено переводам ряда категорий населения на дистанционный формат работы. В моделировании это не было учтено. Сходства и отличия моделирования и статистики по Новосибирску такие же как у Томска.

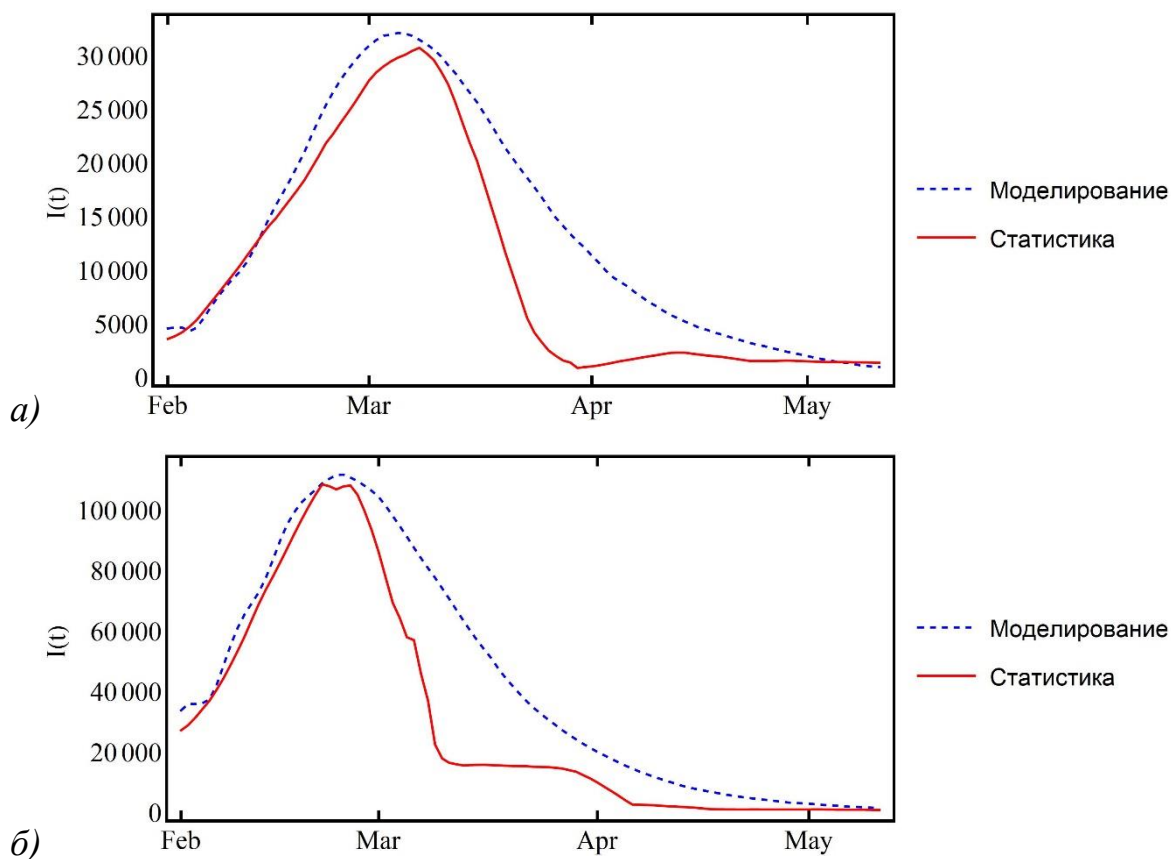


Рисунок 12. Сравнение результатов моделирования и статистики по числу инфицированных по Томску (а) и Новосибирску (б)

Для подтверждения гипотезы причины расхождения данных статистики и результатов моделирование было произведено повторные расчеты, но с учётом введения ограничительных мер. Результаты представлены на рис. 13. Результаты моделирования в этом случае хорошо сопоставимы с данными статистики. Это демонстрирует то, что механизмы учёта ограничительных мер в модели работают. А предложенная гипотеза о влиянии перевода части граждан на удалённую работу является истинной.

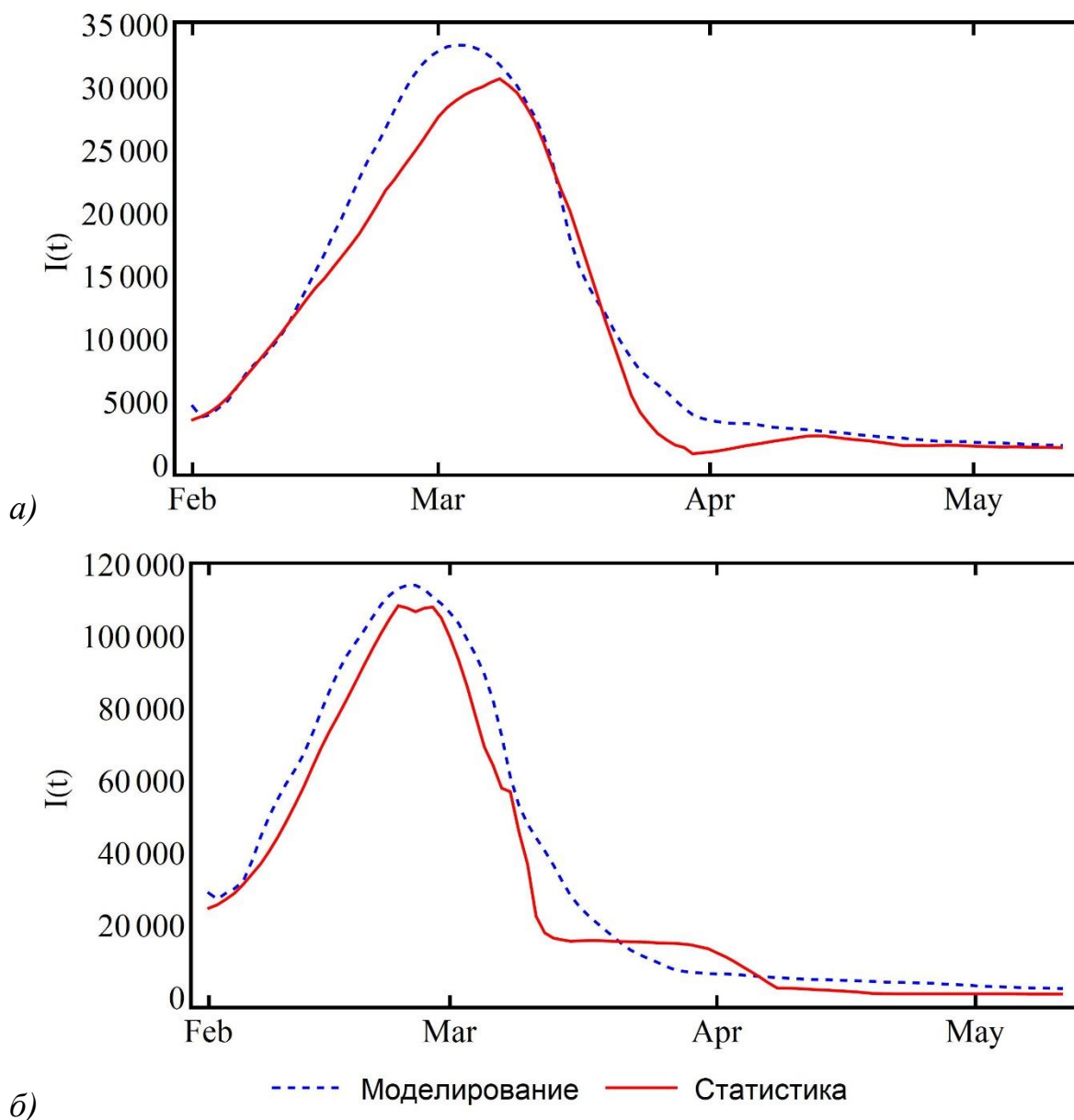


Рисунок 13. Сравнение результатов моделирования (с учётом ограничительных мер) и статистики по числу инфицированных по Томску (а) и Новосибирску (б)

Так как наши вычислительные мощности были ограничены 12 ГБ оперативной памяти, не удалось произвести расчёты для 2 двух крупнейших городских агломераций России: Москва и Санкт-Петербург. Поэтому были взяты 2 других города-миллионика – Красноярск и Нижний Новгород. Для моделирования Красноярской агломерации использовался граф с параметрами: $n = 1900000$, $d(m) = 10$ и $max = 6$. В ходе прогонки статистических данных через

граф было получено значение параметра $P(\alpha) = 0.029$. Результаты моделирования представлены на рис. 14. По графику видим, что статистический пик волны заражения коронавирусом по Красноярску совпадёт с результатами моделирования. Есть небольшие расхождения, которые можно объяснить запозданием и неточностью статистических данных. А в целом численно и количественно графики совпадают в пределах погрешности.

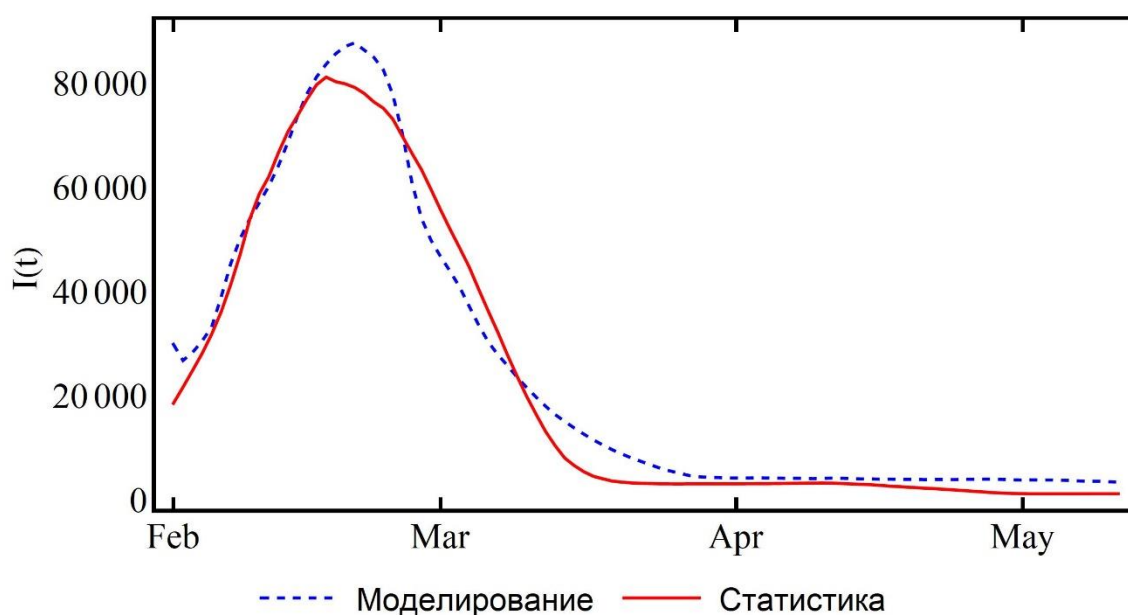


Рисунок 14. Сравнение результатов моделирования (с учётом ограничительных мер) и статистики по числу инфицированных по Красноярску

Для моделирования Нижегородской агломерации использовался граф с параметрами: $n = 2300000$, $d(m) = 10$ и $max = 6$. статистических данных через граф было получено значение параметра $P(\alpha) = 0.0285$. Результаты моделирования представлены на рис. 15. На графики статистические данные меньше результатов моделирования примерно на 15% в пике. Однако отметим, что результаты моделирования хорошо качественно описывают характер роста заболеваемости и его дальнейший спад. Мы наблюдаем быстро рост, и относительно долгий спад. Это имеется и на статической кривой, и на кривой результатов моделирования. Расхождения в количественных значениях можно объяснить тем, что мы не учли ограничительные меры, которые нам не известны

в момент проведения прогноза. Также стоит отметить, что нельзя брать статистические данные как истину, относительно которой можно рассчитать точность нашей модели. Однако данные статистики позволяет подтвердить, что модель способна описывать закономерности при распространении коронавируса.

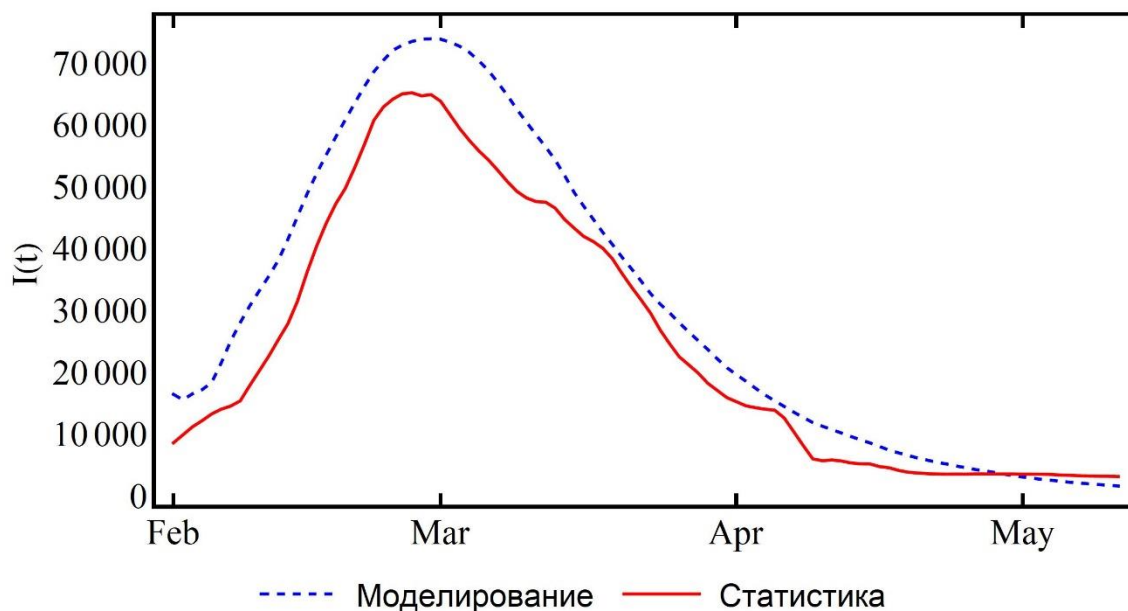


Рисунок 15. Сравнение результатов моделирования и статистики по числу инфицированных по Нижнему Новгороду

По сопоставленным результатам моделирования и данных статистики можно заключить, что предложенная модель способна описывать закономерности распространения коронавируса в определенном населенном пункте. Модель позволяет делать прогнозы на 100 дней вперед. Однако надо учитывать, что следует ежедневно обновлять расчеты, исходя из-за свежих данных статистики, так как могут произойти случайные события, которые сложно предугадать (появление нового штамма), а тем более учесть в модели на 100 дней вперед.

Заключение

Разработана модель на основе сетей позволяющая описывать распространение коронавирусной инфекции. Разработанная модель случайного графа с кликовой структурой позволяет имитировать реальные связи в обществе, по которым распространяется инфекция. Применение графа позволяет использовать концепцию SIR моделей, не вводя большого количества поправочных параметров. При этом параметры разработанной модели имеют постоянное значение в процессе моделирования распространения коронавируса в отдельном населённом пункте. Использование статистических данных для обучения графа позволило решить обратную задачу по нахождению колючего параметра нашей модели, а также позволило максимально приблизить свойства графа к свойствам реальной сети.

Модель была реализована в виде программного кода на языке Python. В ходе верификации модели было получено качественная и количественная сходимость результатов моделирования и данных статистики, что подтверждает способность модели описывать реальные процессы распространения инфекций.

Реализованная модель кроме предсказательной способности обладает высоким исследовательским потенциалом. Применение графа предоставляет множество инструментов для исследования процессом распространения инфекции. Что позволяет изучить влияние различных факторов на сценарий развития пандемии.

4.1 Концепция стартап-проекта

4.1 Проблема конечного потребителя

Пандемия новой коронавирусной инфекции Covid-19, начавшаяся в 2019 года, всё ещё продолжается. Волны пандемии следуют одной за другой. Появление вакцин не привело к окончанию пандемии, которая наносит значительный экономический ущерб. Особенно сильно уязвим человеческий капитал, воспроизводство которого требует больших вложений ресурсов и времени. Для эффективной борьбы с распространением коронавируса необходим относительно точный прогноз по эпидемиологической обстановке. Службы здравоохранения должны быть своевременно готовы к новой волне коронавируса. Запасы медикаментов, баллонов с кислородом, аппаратов искусственной вентиляции лёгких должно быть достаточно для преодоления пика заболеваний. Число сотрудников младшего, среднего и высшего медицинского персонала мобилизованных для борьбы с распространением заболевания также должно быть достаточным. Запасы коечного фонда должны слегка превышать потенциально-максимальное число одновременно больных, находящихся в тяжелом состоянии требующие стационарного лечения в стенах медицинских учреждений. Региональные органы власти и службы контроля должны своевременно вводить ограничительные меры для уменьшения скорости роста числа инфицированных. Все данные мероприятия являются очень дорогостоящими или наносят косвенный экономический ущерб (ограничительные меры). Поэтому объём данных мер должен быть обоснован для минимизации затрат. Например, квалифицированные медицинские сотрудники являются ограниченным ресурсом, который можно лишь перераспределить для решения вызовов перед здравоохранением. Невозможно за короткие сроки увеличить число врачей в стране. Поэтому очень важно на борьбу с пандемией выделять данный ресурс в минимально необходимых объёмах. Чем больше врачей направляются для борьбы с пандемией, тем менее доступной становится здравоохранение для большего членов общества, что может провоцировать излишнюю смертность по другим заболеваниям. Поэтому региональным

органам власти, региональным министерствам здравоохранения и региональным управлениям Роспотребнадзора необходимы точные прогнозы по распространению новой коронавирусной инфекции для рациональное применяя имеющихся ресурсов.

Счетная палата представила свою оценку федеральных расходов на борьбу с пандемией: в 2020 году они составили 2,86 трлн руб. — 12,5% общих расходов бюджета [30]. В 2021 году федеральный бюджет потратил 850 млрд руб. на лишь на антиковидные здравоохранные меры, согласно данным от министра финансов РФ Антона Силуанова. Также стоит отметить, что в среднем на борьбу с ковидом тратилось по 2 млрд. рублей в день, а в вовремя 4-й волной пандемии сумма достигала 4 млрд. в день. Если даже точные прогнозы способны уменьшить данные расходы лишь пару процентов, то это суммы на сотни млн. рублей. Таким образом созданный продукт способен выйти на рынок объёмом десятков млн. рублей. Отметим, что это грубые оценки. Реальные значения объёма рынка скорее всего имеют большие значения.

В представленном проекте разработана и реализована модель, позволяющая прогнозировать динамику распространения новой коронавирусной инфекции с применение сетевых данных. Модель реализована в виде программного продукта, который способен спрогнозировать эпидемиологическую обстановку в отдельном городе, используя статические данные по данному населенному пункту.

В ходе верификации модели было подтверждена её работоспособной и относительно высокая точность в прогнозах.

4.2 Способы защиты интеллектуальной собственности

Модель, реализованная в виде программного продукта, является комплексным техническим решением, включающим в себя программный код, уникальный дизайн, определенные функциональные возможности. Все это, согласно части 1 статьи 1259 Гражданского кодекса Российской Федерации,

является объектом интеллектуальной собственности. Защищенный объект интеллектуальной собственности необходим в следующих ситуациях:

- при передаче(продаже) объекта интеллектуальной собственности другим юридическим или физическим лицам;
- передача права использования программного кода для проектирования других сервисов;
- в спорных ситуациях, где авторство оспаривается.

Согласно части 1 статьи 1259 ГК РФ, любая программа для ЭВМ, не зависимо на какой платформе оно реализовано можно приравнять к литературному произведению, а значит на него действуют положения об авторском праве. Важно понимать, что авторское право защищает не саму идею, а форму, в которой эта идея воплощена, в данном случае защита применяется к программному коду, к элементам интерфейса и к результатам работы модели. Период защиты: в течение всей жизни автора, переживших других соавторов, и семидесяти лет, считая с 1 января года, следующего за годом смерти.

Для того, чтобы зафиксировать свое авторство, необязательно регистрировать мобильное приложение в ответственном за это органе (Роспатенте). Авторское право возникает сразу после официальной публикации разработки. Однако, юристы советуют проводить процедуру регистрации, т.к. она позволяет правообладателю более эффективно защищать свои интересы при возникновении споров. Для получения свидетельства о регистрации необходимо подать соответствующее заявление в Роспатент, вместе с ним предоставить короткое описание ПО (реферат для программы ЭВМ) и часть кода регистрируемого программного обеспечения.

Заявление о регистрации программы для ЭВМ или базы данных можно подать через официальный портал Госуслуг [31]. Физлицам необходима заплатить госпошлину в размере 3000 рублей. Для юридического лица это сумма составляет 4500 рублей. Внеся необходимые данные о программе, вопрос пойдет

на рассмотрение. После положительного ответа будет получена Свидетельство о государственной регистрации программы для ЭВМ или базы данных.

Зарегистрированное свидетельство будет являться основным доказательством исключительных прав на объект интеллектуальной собственности. Свидетельство будет давать возможность официально передавать/продавать авторские права на ПО. В случае нарушения, лица, незаконно использовавшие объект интеллектуальной собственности будут привлечены к гражданско-правовой, административной и уголовной ответственности.

Основным ноу-хау модели заключается применение сетевых технологий. Всё это может быть «спрятано» в коде программного обеспечения. Поэтому реверс-инжиниринговый анализ для раскрытия ноу-хау с целью повторения – малоэффективен.

4.3 Объем и емкость рынка

Рынок продажи прогноз эпидемиологической обстановки можно новым. Поэтому сложно проблематично численно оценить объём и емкость рынка. Особенностью рынка является её сильная зависимость от характера дальнейшего протекания пандемии. Если распространение коронавируса будет сохранить такой же периодический характер, а её летальность будет также высокой, то будет спрос на данные касательно прогноза инфицирования. Основным крупным потребителем можно отметить государство. Конечно возможны продажа базы данных с прогнозами частным компаниям, чья деятельность так или иначе подвергается тем или иным рискам, связанных с распространением новой коронавирусной инфекции. Однако рынок государственных закупок является более привлекательным для нашего проекта.

Рынка услуг по предоставлению прогнозов и исследований по эпидемиологической обстановке как уже отмечали фактически не существует на данный момент. Однако постоянные угроза пандемии создают условия её

формирования. Главным заинтересованным потребителем данных услуг на таком рынке будет является государство. Фактически рынок будет ориентирован на одного потребителя, что определяет её характер. Ёмкость рынка будет определяться в этом случае суммой, которую государство готово потратить на получение данных услуг. Российская федерация состоит из 85 субъектов, каждый из которых является потенциальным потребителем наших. Если провести грубый анализ трат государства [32] на исследования связанных с влиянием новой коронавирусной инфекции на социальные и экономические аспекты жизнедеятельности отдельных субъектов, то видим суммы от 300000 до 2000000 рублей за одно исследование [32]. Если пересчитать на все субъекты, то мы имеем потенциальный объём рынка от 25.5 млн. руб. до 170 млн. руб. Наиболее реалистичным объёмом рынка можно считать 40 млн. рублей. Провести более точные и объективные оценки объёма рынка представляется невозможным, из-за её несформированности, а также из-за её специфики в виде государства как фактически единственного потребителя предлагаемых нами услуг.

4.4 Современное состояние и перспективы отрасли

По большинству прогнозов ВОЗ и министерства здравоохранения РФ пандемия коронавируса прекратится в 2023-2024 годах. Однако все эксперты единогласны, что распространение новой коронавирусной инфекции не прекратится. Инфекция примет сезонный характер. Также невозможно исключить появление новых опасных штаммов.

Прекращение пандемии конечно уменьшит спрос на прогнозы. Однако вирус всё ещё долго будут представлять опасность для человечества и постоянный мониторинг является необходимой мерой для предотвращения новых «взрывных» волн ковида. Также стоит учитывать фактор памяти людей о тяжелых этапах пандемии, что влияет на одобрение трат на мониторинг и прогнозирования коронавирусной инфекции.

Отдельно стоит отметить, что созданная и реализованная модель может быть адаптирована к другим инфекционным заболеваниям, такие как вирус гриппа или оспы обезьян. Поэтому спрос на данные по прогнозированию инфекций всегда будет.

4.5 Конкурентные преимущества создаваемого продукта

Конкурентным преимуществом реализованного продукта является гибкость созданной модели, которая способна имитировать реальные связи в обществе. Это позволяет производить манипуляции над моделью, отражающие реальные мероприятия по борьбе с распространением коронавируса. Например, в нашей модели мы легко можем выделять отдельные социальные и профессиональные группы и проводить манипуляции над ними. Можно выделить группу людей, работающих в IT-сфере и перевести их на удаленную работу, прервав прямые внешние связи в модели при этом сохранив связи внутри семьи (домохозяйства). Практически нет других моделей позволяющие производить такие эксперименты. Для большинства существующих моделей не доступны такие инструменты. Созданная модель применяется на уровне населённых пунктов, что позволяет получать более точные данные. Конкурентные модели чаще всего берут популяции крупных регионов или стран. И зачастую предполагает однородность распространение болезни, что в реальности не так. Разумеется, с помощью таких моделей можно даже кое-что объяснить в распространении пандемии. Но предсказать что-то в отдаленном будущем они могут крайне приблизительно.

Существуют конкурентные модели, базирующиеся на методах машинного обучения. Для примера можно взять модель американца Youyang Gu, которая была отмечена во множества журналах самого высокого уровня [33]. Даже американское государство использует его модель. Его модель как и множество подробных показывают высокую предсказательную способность на небольших промежутках времени. Однако они сильно зависят от чистоты изначальных статистических данных. Разумеется, существуют различные

методы очистки данных, но стоит отметить, что статистические данные могут иметь ошибку более 100% и часто некорректно могут отражать действительность. Также данные модели зависят не только от чистоты статистических данных, но и от их количества. Многие статистические данные сложно или невозможно получить в России. Созданная модель также использует статистические данные, но это не является обязательным. Количество параметров необходимо из статистики намного ниже, чем в моделях на методах машинного обучения. Модель легко позволяет уменьшать и увеличивать доверие к данным статистики при расчетах.

Российские ученые в своей работе [22] предлагают моделировать распространение коронавируса как суперпозицию множества волн на основе дискретного логистического уравнения Фейгенбаума. Данная модель также показала высокую прогностическую способность и зависит от минимального количества параметров. Это является одновременно и преимуществом и недостатком модели. Использование небольшого числа параметров делает модель простой в использовании. Однако модель становится очень ограниченной. Она теряет прямую связь с действительностью и не применима для исследования распространения инфекции. Поэтому она почти не применима в плане создания рекомендаций для противодействия распространению коронавируса.

На основе уже озвученного выше, можно выделить несколько конкретных преимуществ реализованного продукта:

- высокая предсказательная способность модели;
- разнообразные инструменты манипуляций с моделью, позволяющие отражать реальные мероприятия;
- возможность модели перенастроить на другие инфекционные заболевания;
- программный код отечественного производства и не зависит от импортного ПО.

4.6 Целевые сегменты потребителей создаваемого продукта

Основным потребителем, на которой нацелен данный продукт, является государство. Поэтому развитие бизнеса на основе предложенного продукта будет происходить по модели B2G. Государственный сектор России имеет своеобразную национальную специфику. С одной стороны, порядок проведения тендеров (госзакупок) чётко регламентирован и позиционируется открытым, конкурентам и гласным (кроме закупок спецхарактера). Однако зачастую за инициацией этих тендеров стоят множество формальных и не формальных контактов бизнеса и государства. Потому фактически невозможно продать новый продукт государству, не наладив контакты с представителями государства.

Сегмент B2G имеет характерные особенности:

- Решения принимаются по сложной многоуровневой системе.
- Закупки совершаются .
- Сложные финансовые механизмы осуществления закупок.
- Длительные отношения правительства и частных компаний.
- Большие объемы.
- Административный ресурс вовлекается в сотрудничество с частными фирмами.

Получают государственные заказы в основном крупные компании и концерны, которые проверены временем и располагают мощными производственными ресурсами. Малым компания часто выполняют небольшие IT-проекты, где масштабность не играет большего значения.

4.6.1 Налаживание контактов с заказчиками из госсектора

Из изученных общедоступных материалов можно выделить ряд рекомендаций по налаживанию взаимоотношений с представителями государства:

- Заводить знакомства и общаться в сети. Можно писать письма министру, депутату, действительному советнику, работнику администрации и предложите свои варианты решения его задач;
- Записываться на приёмы к представителям власти, для заявления о желании работать с госсектором;
- Стать участником молодежной школы политики, бизнеса или науки;
- Позиционировать свой проект в рамках проводимой государством политики (импортозамещение, развитие информационных технологий, поддержка инновационных технологий, поддержка науки малого бизнеса и тд.)

Рекомендуется начинать налаживать отношения с регионалами представителем государства, а лишь потом выходить на федеральных.

Для продвижения товара планируем предлагать субъектам России попользоваться нашей пробной подписки в течении 1-2 месяцев. Это позволит оценить насколько необходимы наши услуги государству. В случаи положительного отклика на продукт это позволит запустить процесс организации тендера.

4.6.2 Требования для участия в государственных закупках

Тендеры производиться на сайте государственных закупок zakupki.gov.ru.

Для участия необходимо зарегистрироваться на данном портале.

Главное требование – юридическое соответствие, как и при сотрудничестве с крупными организациями. Подрядчик должен быть действующей организацией или ИП (ИП крайне редко допускай к участию в

тендерах). Идеальный вариант – ООО. Минимальный набор документов для проверки: устав, свидетельство о регистрации, приказ о назначении гендиректора. Самозянятым сотрудничать с государством невозможно.

У организации не должно быть многочисленных судебных процессов и, разумеется, фирма не сможет взаимодействовать с государством, если хотя бы раз была признана недобросовестным поставщиком услуг. Очень важен предыдущий опыт агентства.

Очень ценят госструктуры и рекомендации. Часто государственные учреждения обращаются на прямую по рекомендациям своих представителей.

4.7 Цена программного продукта

Себестоимость программного продукта в первую очередь определяется оплатой труда программиста-разработчика и амортизационными расходами. По данным рекрутингового сайта HeadHunter (hh.ru) заработная плата разработчика python уровня middle начинается от 65 000 рублей.

Для разработки программного продукта требуется выполнить следующие работы:

- сбор и анализ данных;
- проектирование, разработка и верификация модели;
- реализация и тестирование программного продукта;
- интеграционное тестирование программного продукта;
- создание документации и инструкции.

Для сбора, документирования и анализа требований модели выделяется обычно до 3 недель. Дальнейшие сроки зависят от объема работ, определенного на первом этапе. По опыту, в подобных проектах этап проектирования и разработки модели и этап реализации и тестирования могут частично перекрываться по времени и производиться по гибкой методологии итерациями. Приблизительный срок выполнения этих работ – 20 недель. На интеграционное

тестирование программного продукта, устранение недочетов, написание документации и инструкции потребуется до 4 недель.

Итого, на разработку программного продукта и доведения его до версии, готовой к выходу на рынок потребуется до 27 недель или 6,3 месяца.

Таблица 1. Расчет затрат на заработную плату разработчика

Наименование	Количество (чел)	Сумма за месяц (руб.)	Сумма с выплатами (руб.)	Сумма за полгода (руб.)
Менеджер проекта + программист python уровня middle	1	65000	84630	506280

Для работы разработчика требуется приобретение рабочего места. Расчет стоимости одного рабочего места представлен в Таблице 2 (цены взяты с сайта www.komus.ru и www.dns-shop.ru).

Таблица 2. Расчет рабочего места программиста

Наименование	Количество	Цена (руб.)
Письменный стол «Арго» эргономичный	1 шт.	7 665
Кресло офисное EasyChair Comfort GTP	1 шт.	5790
Компьютер для работы с сетевыми данными (бюджетная сборка)	1 шт.	146700
Лампа настольная	1 шт.	1700
Итого		161855

Приобретенное оборудование для рабочих мест относится ко второй амортизационной группе со сроком полезного использования до 3 лет или 36 месяцев.

Норма амортизации оборудования:

$$\text{Норма амортизации} = \frac{100\%}{\text{Срок полезного использования}} = \frac{100\%}{36} = 2.8\%$$

Ежемесячная сумма амортизации оборудования составит:

$$161855 \cdot 0.028 = 4532 \text{ руб.}$$

Амортизационные расходы на время реализации продукта:

$$4532 \cdot 6 = 27192 \text{ руб}$$

Разработка программного продукта будет производиться на дому программиста. Поэтому в себестоимости не учитывается траты на аренду помещения и услуги связи. Итого, себестоимость программного продукта составит:

$$27192 + 506280 = 533472 \text{ руб.}$$

С учётом закладываемой прибыли в 25% стоимость программного продукта составит примерно 670000 рублей.

4.8 Бизнес-модель проекта

Бизнес-модель проекта может быть показана с помощью модели Остервальдера. Созданный программный продукт будет использован для создания обширной базы данных по прогнозам распространения инфекционных заболеваний по городам и регионам.

Таблица 3. Бизнес модель по Остервальдеру

Ключевые партнеры Отсутствуют	Ключевые виды деятельности Предоставление базы данных по прогнозам по распространению инфекционных заболеваний	Ценностные предложения Единая база данных по прогнозам по различным городам и регионам с высокой точностью за счёт еденной обученной сети.	Взаимоотношения с клиентами (государством) Наложение прямых контактов с региональными и федеральными представителями министерства здравоохранения. Участие в конкурсах поддерживаемых государством.	Потребительские сегменты Главным потребителем продукта является государственные медицинские и здравоохранные учреждения.
	Ключевые ресурсы Программный продукт на основе модели с использованием сетевых технологий		Каналы сбыта Через тендеры на сайте государственных закупок zakupki.gov.ru	
Структура издержек Инвестиции в проект – 560000 руб. Операционные издержки – 206460 руб.			Потоки поступления доходов стоимость одной годовой подписки – 408000 руб.	

Для развития программного продукта, его дальнейшей поддержки и создания базы данных потребуется команда специалистов в следующем минимальном ролевом составе: менеджер проекта и разработчик python уровня middle и один аналитик данных. По данным рекрутингового сайта HeadHunter

(hh.ru) заработная плата разработчика python уровня middle и аналитика данных начинается от 65 000 рублей. Так как стартап проект на начальном этапе, как правило, ограничен в финансовых ресурсах, заработная плата сотрудникам возможна с опционом. В этом случае зарплата на начальном этапе – минимальна. В то же время у сотрудников появляется заинтересованность в конечном результате. Для ведения бухгалтерии на начальном этапе можно воспользоваться услугами бухгалтерских бюро. Итого, команда на начальном этапе составляет 2 человека. При этом менеджер проекта является непосредственно разработчиком программного продукта и модели. Дальнейшие оценки инвестиций рассчитывались исходя из этого количества сотрудников.

4.8.1 Первоначальные инвестиции

Для работы сотрудников требуется приобретение рабочих мест. Расчет стоимости одного рабочего места представлен в Таблице 2 (цены взяты с сайта www.komus.ru и www.dns-shop.ru).

Производим расчет одноразовых затрат для старта проекта (Таблица 4) расчет затрат на содержание предприятия (Таблица 5). Затраты на заработную плату производим исходя из минимальных уровней зарплат специалистов (Таблица 6).

Таблица 4. Расчет одноразовых затрат для старта проекта

Наименование	Количество	Цена (руб.)	Сумма (руб.)
Рабочее место программиста	2 шт.	161855	323710
Регистрация ООО	1 шт.	4000	4000
Создание интернет-сайта	1 шт.	25000	25000
Итого			352710

Таблица 5. Расчет затрат для содержания предприятия

Наименование	Сумма за месяц (руб.)	Сумма за год (руб.)
Аренда офиса 30 кв. м	22000	264000
Затраты на коммуникацию (Интернет в офис 100 Мбит/с для организаций)	5200	62400
Накладные расходы на содержание офиса, включая обслуживание расчетного счета	10000	120000
Итого	37200	446400

Таблица 6. Расчет затрат на заработную плату

Наименование	Количество (чел)	Сумма за месяц (руб.)	Сумма с выплатами (руб.)	Сумма за год (руб.)
Менеджер проекта + разработчик python уровня middle	1	65000	84630	1012560
Аналитик данных	1	65000	84630	1012560
Итого			169260	2031120

Размер первоначальных инвестиций будем считать равным сумме одноразовых затрат фирмы, затрат на содержание предприятия и выплат заработных плат в течении 1 месяца:

$$352710+37200+169260=559170.$$

Стоимость создания программного продукта не учитывается в первоначальных инвестициях, а будет учитываться при расчёте стоимости доступе базы данных. Итого, для открытия фирмы потребуется 560000 изначальных инвестиций.

4.8.2 Стоимость подписке к базе данных

Реализоваться база данных по прогнозам эпидемиологической обстановки будет производится в форме годовой подписки. Субъект РФ напрямую или через федеральный центр объявляет о тендере на приобретении продукта (базы данных). Имея единственный вариант, субъект РФ приобретает нашу подписку на базу данных. РФ состоит из 85 субъектов с разным числом населения. Для удобства расчётов стоимость подписки для каждого субъекта будем считать равным. Себестоимость ежемесячной подписки будем рассчитывать, как сумму ежемесячных затрат и амортизационных расходов делённое на число подписок.

Ежемесячные затраты состоят из расходов на аренду и содержание офиса (таблица 5) и выплату зарплат (таблица 6):

$$37200 + 169260 = 206460 \text{ руб.}$$

Приобретенное оборудование для рабочих мест относится ко второй амортизационной группе со сроком полезного использования до 3 лет или 36 месяцев. Стоимость оборудования:

Таблица 7. Стоимость оборудования

Наименование	Количество	Цена (руб.)	Сумма (руб.)
Рабочее место программиста	2 шт.	161855	323710

Норма амортизации оборудования:

$$\text{Норма амортизации} = \frac{100\%}{\text{Срок полезного использования}} = \frac{100\%}{36} = 2.8\%$$

Ежемесячная сумма амортизации оборудования составит:

$$323710 \cdot 0.028 = 9064 \text{ руб.}$$

Будем считать, что стоимость программного продукта будет распределена на первые 12 месяцев функционирования предприятия:

$$\frac{670000}{12} = 55834 \text{ руб.}$$

Итого ежемесячные амортизационные расходы:

$$55834 + 9064 = 64898 \text{ руб.}$$

Месячная себестоимость создания и содержания базы данных:

$$64898 + 206460 = 271358 \text{ руб.}$$

В Таблице 8 представлена месячная себестоимость подписки в зависимости от числа проданных подписок, а также годовая стоимость подписки с учётом 25% наценки.

Таблица 8. Зависимость стоимости подписки от их числа

Количество активных подписок	Себестоимость подписки на 1 мес. (руб.)	Стоимость подписки на 1 мес. (руб.)	Стоимость годовой подписки (руб.)
5	54272	68000	816000
10	27135	34000	408000
15	18091	17000	204000
20	13568	8500	102000

4.9 Система налогообложения и оценка эффективности

На первоначальном этапе выбираем упрощенную систему налогообложения (УСН). В УСН доступны два вида налогов: 6% от выручки и 15% от прибыли (то есть доходы минус расходы). Для определения оптимального вида налога производится расчет ежемесячной прибыли в зависимости от количества годовых подписок. В таблице 9 представлены результаты расчётов. Как видно, при цене годовой подписки равной 408000 желательно иметь 10 активных подписок. Точка безубыточности для предприятия равна 8 годовым подпискам. При 10 годовых подписок выгоднее

платить налог от прибыли, чем от выручки. Также по Таблице 9 видно, что инвестиции окупятся в течении года, если за первый год реализовать 10 подписок.

Таблица 9. Чистая прибыль предприятия при разных видах налогов

Количество подписок	6% от выручки, руб.	15% от прибыли, руб.
5	-1338696	-1216296 руб.
10	578904	700148
15	2496504	2434148
20	4414104	4168148
25	6331704	5902148

При 10 активных годовых подписок рентабельность инвестиций составит:

$$ROI = \text{Чистая прибыль} / \text{Сумма инвестиций} * 100\% = 700148 / 560000 * 100\% = 125 \%$$

4.10 План продаж

Первые 2 месяца реализации проекта будут потрачены на работу с 1 субъектом РФ, который готов будет приобрести нашу годовую подписку через систему государственных закупок. Далее планируется наращивать количество активных подписок за счёт продвижения в новых субъектах. По таблице 10 видно, что планируем за первые полгода реализовать лишь 2 подписки. Такое малое количество подписок обоснованно тем, что на начальном этапе необходимо будет продвигать совсем неестественных продукт, а также необходимо будет изучить механизмы запуска тендера. За второе полугодие планируется реализовать ещё 6 подписок. Достижимость данного числа продаж обусловлено тем, что после 2 подписок мы будем иметь рекомендации, а также уже приобретённый опыт участия в государственных закупках. Также необходимо достигнуть 8 активных подписок, чтобы выйти на окупаемость.

На рис.16а представлен план по количеству активных подписок. При соблюдении данного плана мы получим динамику прибыли представленный на рис. 16б. На первый взгляд мы имеем значительную отрицательную динамику, однако надо учесть, что в расчётах учитываются расходы на создания модели, а также зарплата менеджера проекта. Поэтому реальные расходы в первые месяцы не потребуют дополнительных привлечений средств. Также потенциальные расходы можно снизить за счёт того, что не привлекать сотрудника (аналитика данных) в первые 6 месяцев реализации проекта. Это позволит сэкономить 508000 рублей на зарплате. На этом основании можно заключить, что динамика прибыли, представленная на графике, является допустимой при реализации проекта.

Таблица 10. Расчёт выручки и издержек за первый год

Месяц	Количество проданных подписок за данный месяц	Количество активных подписок	Выручка	Издержки	Прибыль
1	0	0	0	271358	-271358
2	0	0	0	542716	-542716
3	1	1	408000	814074	-406074
4	0	1	408000	1085432	-677432
5	1	2	816000	1356790	-540790
6	0	2	816000	1628148	-812148
7	1	3	1224000	1899506	-675506
8	1	4	1632000	2170864	-538864
9	0	4	1632000	2442222	-810222
10	1	5	2040000	2713580	-673580
11	1	6	2448000	2984938	-536938
12	2	8	3264000	3256296	7704

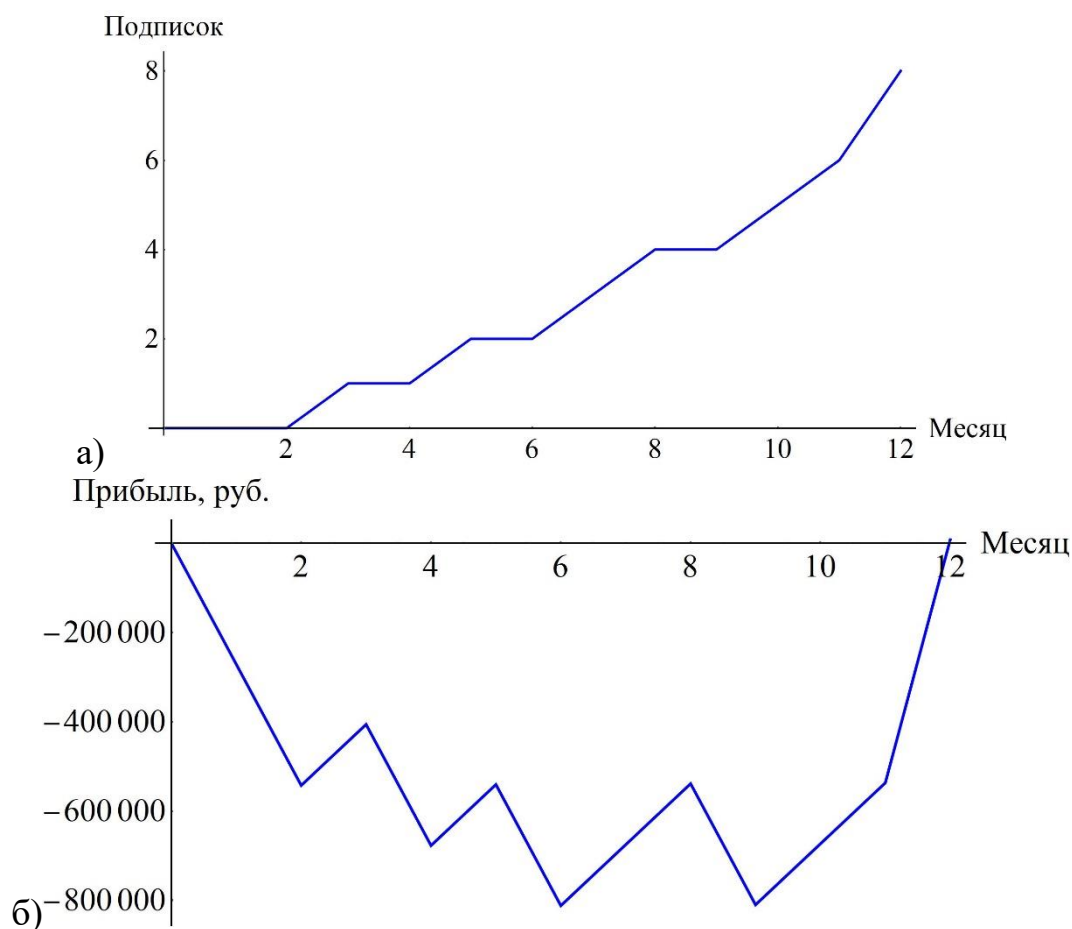


Рисунок. 16. План по количеству активных подписок (а) и динамика изменения прибыли (б) в первые 12 месяцев реализации проекта

4.11 Заключение разделу

Предложенный проект является инновационным, конкурентноспособным продуктом, способным решить реальные проблемы потребителей. Использование продукта приносит потребителю ощутимый экономический эффект. С точки зрения инвестиционной привлекательности проект не требует больших финансовых вложений. В то же время имеет высокую рентабельность инвестиций – 25 %.

5.1 Социальная ответственность

Социальная ответственность – ответственность отдельного ученого и научного сообщества перед обществом. Первостепенное значение при этом имеет безопасность применения технологий, которые создаются на основе достижений науки, предотвращение или минимизация возможных негативных последствий их применения, обеспечение безопасного как для испытуемых, как и для окружающей среды проведения исследований.

Целью данной работы является разработка SIR-подобной модели с применением сетей, имитирующие общественные связи индивидов, по которым распространяются инфекционные заболевания, в частности новая коронавирусная инфекция. Данная модель позволяет прогнозировать эпидемиологическую обстановку в городах с высокой точностью. Применение инструментария теории графов позволяет относительно легко учитывать различные факты, влияющие на распространение инфекции. Модель реализована в виде программного кода на языке python. Потенциальными потребителями разработанной модели являются региональные службы по контролю над эпидемиологической обстановкой (например, управление Роспотребнадзора по Томской области). В качестве основного места выполнения работы рассматривается рабочее место стажёра IT-компании, расположенное в одном из офисных зданий Томска и оснащенное техникой для разработки системы: дисплей, клавиатура, компьютерная мышь, системный блок.

Целью данного раздела является анализ вредных и опасных факторов возникающих при проведении исследования и обоснование мер защиты от них. В разделе также рассматриваются вопросы безопасности в чрезвычайных ситуациях, анализ факторов, влияющих на обеспечение экологической безопасности.

5.1 Производственная безопасность

5.1.1 Отклонение показателей микроклимата в помещении

Проанализируем микроклимат в помещении, где находится рабочее место. Микроклимат производственных помещений определяют следующие параметры: температура, относительная влажность, скорость движения воздуха. Эти факторы влияют на организм человека, определяя его самочувствие.

Оптимальные и допустимые значения параметров микроклимата приведены в таблице 11 и 12.

Таблица 11 – Оптимальные нормы микроклимата

Период года	Температура воздуха, С°	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	19-23	40-60	0.1
Теплый	23-25		0.2

Таблица 12 – Допустимые нормы микроклимата

Период года	Температура воздуха, С°		Относительная влажность воздуха, %	Скорость движения воздуха, м/с
	Нижняя допустимая граница	Верхняя допустимая граница		
Холодный	15	24	20-80	<0.5
Теплый	22	28	20-80	<0.5

Температура в теплый период года 23-25°С, в холодный период года 19-23°С, относительная влажность воздуха 40-60%, скорость движения воздуха 0,1 м/с.

Общая площадь рабочего помещения составляет 24 м², объем составляет 84 м³, количество рабочих мест 2. По СанПиН 2.2.2/2.4.1340-03 санитарные нормы составляют 6,5 м² и 20 м³ объема на одного человека. В данном случае на 1 человека приходится 12 м² и 42 м³ объёма. Исходя из приведенных выше

данных, можно сказать, что количество рабочих мест соответствует размерам помещения по санитарным нормам.

После анализа габаритных размеров рассмотрим микроклимат в этой комнате. В качестве параметров микроклимата рассмотрим температуру, влажность воздуха, скорость ветра.

В помещении осуществляется естественная вентиляция посредством наличия легко открываемого оконного проема (форточки), а также дверного проема. По зоне действия такая вентиляция является общеобменной. Основным недостатком - приточный воздух поступает в помещение без предварительной очистки и нагревания. Согласно нормам СанПиН 2.2.2/2.4.1340-03 объем воздуха необходимый на одного человека в помещении без дополнительной вентиляции должен быть более 40 м³ [34]. В нашем случае объем воздуха на одного человека составляет 42 м³, из этого следует, что дополнительная вентиляция не требуется. Параметры микроклимата поддерживаются в холодное время года за счет систем водяного отопления с нагревом воды до 100°С, а в теплое время года – за счет кондиционирования, с параметрами согласно [35]. Нормируемые параметры микроклимата, ионного состава воздуха, содержания вредных веществ должны соответствовать требованиям [36].

5.1.2 Превышение уровней шума

Одним из наиболее распространенных в производстве вредных факторов является шум. Он создается рабочим оборудованием, преобразователями напряжения, рабочими лампами дневного света, а также проникает снаружи. Шум вызывает головную боль, усталость, бессонницу или сонливость, ослабляет внимание, память ухудшается, реакция уменьшается.

Основным источником шума в комнате являются компьютерные охлаждающие вентиляторы и. Максимальный уровень шума при исправности устройств охлаждения достигает 48.4 дБА. Согласно СанПиН 2.2.2 / 2.4.1340-03, при выполнении основных работ на ПЭВМ уровень шума на рабочем месте не должен превышать 82 дБА [37].

При значениях выше допустимого уровня необходимо предусмотреть средства индивидуальной защиты(СИЗ) и средства коллективной защиты (СКЗ) от шума.

Средства коллективной защиты:

1. устранение причин шума или существенное его ослабление в источнике образования;
2. изоляция источников шума от окружающей среды (применение глушителей, экранов, звукопоглощающих строительных материалов);
3. применение средств, снижающих шум и вибрацию на пути их распространения;

Средства индивидуальной защиты;

1. применение спецодежды и защитных средств органов слуха: наушники, беруши, антифоны.

Для поддержания в норме уровня шума от систем охлаждения ПЭВМ необходимо своевременно проводить чистку и осмотр составных устройств компьютерных.

5.1.3 Повышенный уровень электромагнитных излучений

Источником электромагнитных излучений в нашем случае являются дисплеи ПЭВМ. Монитор компьютера включает в себя излучения рентгеновской, ультрафиолетовой и инфракрасной области, а также широкий диапазон электромагнитных волн других частот. Согласно СанПиН 2.2.2/2.4.1340-03 напряженность электромагнитного поля по электрической составляющей на расстоянии 50 см вокруг ВДТ не должна превышать 25В/м в диапазоне от 5Гц до 2кГц, 2,5В/м в диапазоне от 2 до 400кГц [38]. Плотность магнитного потока не должна превышать в диапазоне от 5 Гц до 2 кГц 250нТл, и 25нТл в диапазоне от 2 до 400кГц. Поверхностный электростатический потенциал не должен превышать 500В [38]. В ходе работы использовалась ПЭВМ типа Acer Nitro со следующими характеристиками: напряженность электромагнитного поля 2,2В/м; поверхностный потенциал составляет 450 В [6].

При длительном постоянном воздействии электромагнитного поля (ЭМП) радиочастотного диапазона при работе на ПЭВМ у человеческого организма сердечно-сосудистые, респираторные и нервные расстройства, головные боли, усталость, ухудшение состояния здоровья, гипотония, изменения сердечной мышцы проводимости. Тепловой эффект ЭМП характеризуется увеличением температуры тела, локальным селективным нагревом тканей, органов, клеток за счет перехода ЭМП на теплую энергию.

Предельно допустимые уровни облучения [38]:

- а) до 10 мкВт/см² , время работы (8 часов);
- б) от 10 до 100 мкВт/см² , время работы не более 2 часов;
- в) от 100 до 1000 мкВт/см² , время работы не более 20 мин. при условии пользования защитными очками;
- г) для населения в целом ППМ не должен превышать 1 мкВт/см².

Защита человека от опасного воздействия электромагнитного излучения осуществляется следующими способами:

СКЗ

1. защита временем;
2. защита расстоянием;
3. снижение интенсивности излучения непосредственно в самом источнике излучения;
4. экранирование источника;
5. защита рабочего места от излучения;

СИЗ

1. Очки и специальная одежда, выполненная из металлизированной ткани (кольчуга). При этом следует отметить, что использование СИЗ возможно при кратковременных работах и является мерой аварийного характера. Ежедневная защита обслуживающего персонала должна обеспечиваться другими средствами.

2. Вместо обычных стекол используют стекла, покрытые тонким слоем золота или олова.

5.1.4 Поражение электрическим током

К опасным факторам можно отнести наличие в помещении большого количества аппаратуры, использующей однофазный электрический ток напряжением 220 В и частотой 50Гц. По опасности электропоражения комната с ПЭВМ относится к помещениям без повышенной опасности, так как отсутствует повышенная влажность, высокая температура, токопроводящая пыль и возможность одновременного сприкосновения токоведущих элементов с заземленными металлическими корпусами оборудования [39].

Каждому необходимо знать меры медицинской помощи при поражении электрическим током. В любом рабочем помещении необходимо иметь медицинскую аптечку для оказания первой медицинской помощи.

Поражение электрическим током чаще всего наступает при небрежном обращении с приборами, при неисправности электроустановок или при их повреждении.

Для освобождения пострадавшего от токоведущих частей необходимо использовать непроводящие материалы. Если после освобождения пострадавшего из-под напряжения он не дышит, или дыхание слабое, необходимо вызвать бригаду скорой медицинской помощи и оказать пострадавшему доврачебную медицинскую помощь:

- обеспечить доступ свежего воздуха (снять с пострадавшего стесняющую одежду, расстегнуть ворот);
- очистить дыхательные пути;
- приступить к искусственной вентиляции легких (искусственное дыхание);
- в случае необходимости приступить к непрямому массажу сердца.
- любой электроприбор должен быть немедленно обесточен в случае:
- возникновения угрозы жизни или здоровью человека;
- появления запаха, характерного для горячей изоляции или пластмассы;
- появления дыма или огня;
- появления искрения;

- обнаружения видимого повреждения силовых кабелей или коммутационных устройств.

Для защиты от поражения электрическим током используют СИЗ и СКЗ.

Средства коллективной защиты:

1. Заземление источников электрического тока;
2. Использование щитов, барьеров, клеток, ширм, а также заземляющих и шунтирующих штанг, специальных знаков и плакатов.

Средства индивидуальной защиты:

Использование диэлектрических перчаток, изолирующих клещей и штанг, слесарных инструментов с изолированными рукоятками, указатели величины напряжения, калоши, боты, подставки и коврики.

5.1.5 Освещённость

Источником недостаточной освещенности рабочей зоны является низкий уровень естественного и искусственного освещения в помещении рабочей зоны. Разработка и эксплуатация программного кода по моделированию эпидемиологической обстановки предполагает длительный контакт с монитором ЭВМ. Недостаточное освещение увеличивает нагрузку на зрительные органы, что может привести к профессиональным заболеваниям зрительных органов, увеличивает утомляемость и нагрузку на нервную систему. Согласно СанПиН 2.2.4.3359-16 разработка модели и её эксплуатация относятся к разрядку III и подразряду Г зрительных работ (работы высокой точности) [40].

Рабочие столы следует размещать таким образом, чтобы видеодисплейные терминалы были ориентированы боковой стороной к световым проемам, чтобы естественный свет падал преимущественно слева. Искусственное освещение в помещениях для эксплуатации ПЭВМ должно осуществляться системой общего равномерного освещения. В производственных и административно-общественных помещениях, в случаях преимущественной работы с документами, следует применять системы комбинированного освещения.

Освещенность на поверхности стола в зоне размещения рабочего документа должна быть 300 - 500 лк. Освещение не должно создавать бликов на поверхности экрана. Освещенность поверхности экрана не должна быть более 300 лк. Следует ограничивать прямую блескость от источников освещения, при этом яркость светящихся поверхностей (окна, светильники и др.), находящихся в поле зрения, должна быть не более 200 кд/м².

Расчёт общего равномерного искусственного освещения горизонтальной рабочей поверхности выполняется методом коэффициента светового потока, учитывающим световой поток, отражённый от потолка и стен [41]. Длина помещения $A = 6$ м, ширина $B = 4$ м, высота $H = 3,5$ м. Высота рабочей поверхности над полом $h_p = 1,0$ м. Согласно СНиП 52.13330.2016 необходимо создать освещенность не ниже 300 лк, в соответствии с разрядом зрительной работы [42].

Площадь помещения: $S = A \times B$,

где A – длина, м; B – ширина, м.

$$S = 6 \times 4 = 24 \text{ м}^2$$

Коэффициент отражения свежепобеленных стен с окнами, без штор $\rho_c = 50\%$, свежепобеленного потолка $\rho_{п} 70\%$. Коэффициент запаса, учитывающий загрязнение светильника, для помещений с малым выделением пыли равен $K_z = 1,5$. Коэффициент неравномерности для люминесцентных ламп $Z = 1,1$.

Выбираем лампу дневного света ЛД-40, световой поток которой равен ФЛД = 2600 Лм. Выбираем светильники с люминесцентными лампами типа ОДОР-2-40. Этот светильник имеет две лампы мощностью 40 Вт каждая, длина светильника равна 1227 мм, ширина – 265 мм.

Интегральным критерием оптимальности расположения светильников является величина λ , которая для люминесцентных светильников с защитной решёткой лежит в диапазоне 1,1–1,3. Принимаем $\lambda = 1,1$, расстояние светильников от перекрытия (свес) $h_c = 0,3$ м.

Высота светильника над рабочей поверхностью определяется по формуле:

$$h = H - h_n - h_p = 3,5 - 1 - 0,3 = 2,2 \text{ м}$$

где h_n – высота светильника над полом, высота подвеса,

h_p – высота рабочей поверхности над полом.

Наименьшая допустимая высота подвеса над полом для двухламповых светильников ОДОР: $h_n = 3,5 \text{ м}$.

Из формулы $\Phi_{л} = (E \cdot S \cdot K_z \cdot Z) / (N \cdot \eta)$ находим число ламп N

$$N = (E \cdot S \cdot K_z \cdot Z) / (\Phi_{л} \cdot \eta)$$

η определяем через индекс помещения по формуле:

$$i = \frac{A \cdot B}{h(A + B)} = \frac{6 \cdot 4}{2,2 \cdot 10} = 1,1$$

Коэффициент использования светового потока, показывающий какая часть светового потока ламп попадает на рабочую поверхность, для светильников типа ОДОР с люминесцентными лампами при $\rho_{п} = 70 \%$, $\rho_{с} = 50\%$ и индексе помещения $i = 1,1$ равен $\eta = 0,43$.

Тогда $N = (E \cdot S \cdot K_z \cdot Z) / \Phi_{л} \cdot \eta = (300 \cdot 24 \cdot 1,5 \cdot 1,1) / (2600 \cdot 0,43) = 11$ ламп;
Принимаем 12 ламп, при этом получается 6 светильников, т.е. 2 ряда по 3 светильника.

Из условий равномерности освещения определяем расстояния L_1 и $L_1/3$ и L_2 и $L_2/3$ по следующим уравнениям:

$$6000 = 2 \cdot L_1 + 2/3 \cdot L_1 + 3 \cdot 1227; L_1 = 870 \text{ мм}; L_1/3 = 290 \text{ мм};$$

$$4000 = L_2 + 2/3 \cdot L_2 + 2 \cdot 265; L_2 = 2082 \text{ мм}; L_2/3 = 694 \text{ мм};$$

На рис. 17 изображен план помещения и размещения светильников с люминесцентными лампами.

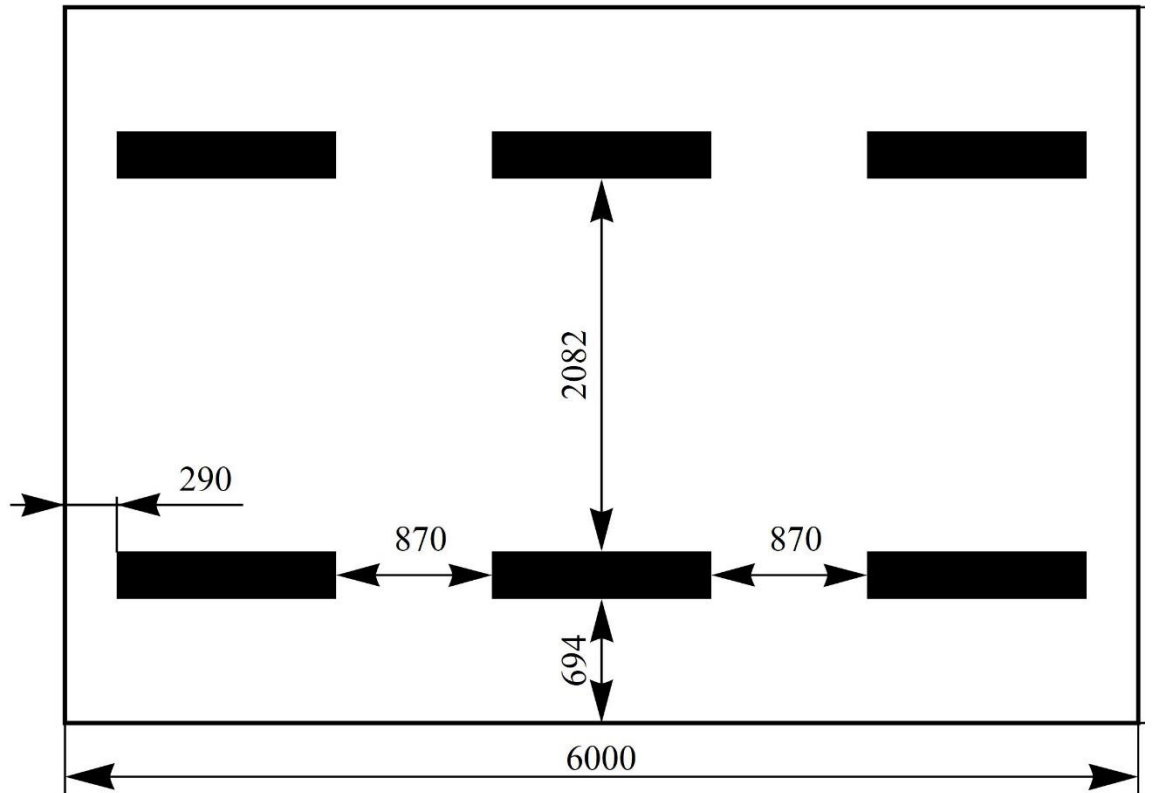


Рисунок 17 – План помещения и размещения светильников с люминесцентными лампами

Потребный световой поток люминесцентной лампы определяется по формуле: $\Phi_{л} = (E \cdot S \cdot K_3 \cdot Z) / N \cdot \eta = (300 \cdot 24 \cdot 1,5 \cdot 1,1) / (12 \cdot 0,43) = 2093,23$ лм

Делаем проверку выполнения условия:

$$-10\% \leq \frac{\Phi_{лд} - \Phi_{л}}{\Phi_{лд}} \cdot 100\% \leq 20\%$$

$$\frac{\Phi_{лд} - \Phi_{л}}{\Phi_{лд}} \cdot 100\% = \frac{2600 - 2093,23}{2600} \cdot 100\% = 19,5\% < 20\%$$

Таким образом, мы получили, что необходимый световой поток не выходит за пределы требуемого диапазона. Теперь рассчитаем мощность осветительной установки: $P = 12 \cdot 40 = 480$ Вт.

5.1.5 Пожарная опасность

По взрывопожарной и пожарной опасности помещения подразделяются на категории А, Б, В1-В4, Г и Д.

Согласно НПБ 105-03 помещение с ПВЭМ относится к категории В–горючие и трудно горючие жидкости, твердые горючие и трудно горючие вещества и материалы, вещества и материалы, способные при взаимодействии с водой, кислородом воздуха или друг с другом только гореть, при условии, что помещения, в которых находится, не относятся к категории наиболее опасных А или Б.

По степени огнестойкости данное помещение относится к 1-й степени огнестойкости по СНиП 2.01.02-85 (выполнено из кирпича, которое относится к трудносгораемым материалам) [43].

Возникновение пожара при работе с электронной аппаратурой может быть по причинам как электрического, так и неэлектрического характера.

Причины возникновения пожара неэлектрического характера:

а) халатное неосторожное обращение с огнем (курение, оставленные без присмотра нагревательные приборы, использование открытого огня);

Причины возникновения пожара электрического характера: короткое замыкание, перегрузки по току, искрение и электрические дуги, статическое электричество и т. п.

Для локализации или ликвидации загорания на начальной стадии используются первичные средства пожаротушения. Первичные средства пожаротушения обычно применяют до прибытия пожарной команды.

Огнетушители водо-пенные (ОХВП-10) используют для тушения очагов пожара без наличия электроэнергии. Углекислотные (ОУ-2) и порошковые огнетушители предназначены для тушения электроустановок, находящихся под напряжением до 1000В. Для тушения токоведущих частей и электроустановок применяется переносной порошковый огнетушитель, например ОП-5.

В общественных зданиях и сооружениях на каждом этаже должно размещаться не менее двух переносных огнетушителей. Огнетушители следует

располагать на видных местах вблизи от выходов из помещений на высоте не более 1,35 м. Размещение первичных средств пожаротушения в коридорах, переходах не должно препятствовать безопасной эвакуации людей.

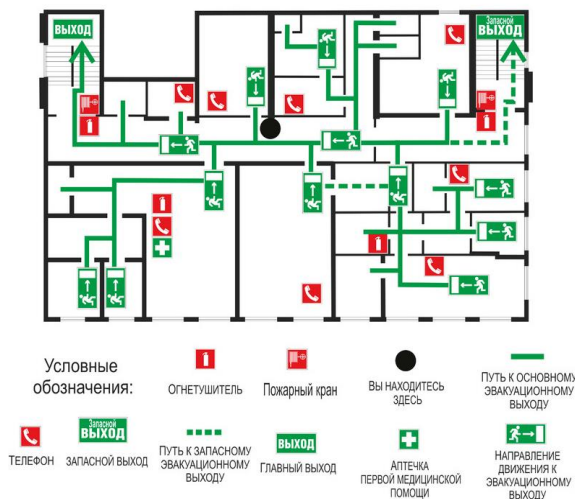
В качестве превентивных мер необходимо:

- Проводить систематические проверки пожарной безопасности в здании;
- Проведение пожарно-технического обследования здания с представителями пожарного надзора;
- Проведение инструктажей с работниками здания по вопросам пожарной безопасности;
- Проведение систематической проверки оборудования и средств пожаротушения, системы оповещения о пожаре.

В случае возникновения пожара необходимо:

- Сообщить о пожаре ответственному работнику или позвонить в службу пожаротушения;
- Отключить питание ПВЭМ;
- При возникновении угрозы жизни необходимо покинуть опасную зону согласно плану эвакуации (см. рис. 18);
- При возникновении небольшого пожара необходимо воспользоваться ручным огнетушителем и потушить возгорание.

Действия при пожаре Сохранять спокойствие!	
1	Сообщить по телефону: 01, 112 (сеп); - адрес объекта - место возникновения пожара - свое фамилию
2	Эвакуировать людей - ориентироваться по знакам направления движения - взять с собой пострадавших
3	По возможности принять меры по тушению пожара - использовать средства противопожарной защиты при необходимости - обесточить помещение
Действия при аварии Сохранять спокойствие!	
1	Сообщить по телефону: 01, 112(сеп); - адрес объекта - что случилось - имеются ли пострадавшие - свое фамилию
2	Локализовать аварию - прекратить развитие аварии - использовать средства защиты - обозначить место аварии
3	Эвакуировать людей - оказать помощь пострадавшим - ориентироваться по знакам - взять с собой пострадавших.



ГОСТ Р 12.2.143-2009 200-25-1440-ЖЗ

Рисунок 18 – План эвакуации

5.2. Экологическая безопасность

В компьютерах огромное количество компонентов, которые содержат токсичные вещества и представляют угрозу, как для человека, так и для окружающей среды. К таким веществам относятся:

- свинец (накапливается в организме, поражая почки, нервную систему);
- ртуть (поражает мозг и нервную систему);
- никель и цинк (могут вызывать дерматит);
- щелочи (прожигают слизистые оболочки и кожу);

Поэтому компьютер требует специальных комплексных методов утилизации. В этот комплекс мероприятий входят:

- отделение металлических частей от неметаллических;
- металлические части переплавляются для последующего производства;
- неметаллические части компьютера подвергаются специально переработке;

Исходя из сказанного выше перед планированием покупки компьютера необходимо [44]:

- Побеспокойтесь заранее о том, каким образом будет утилизирована имеющаяся техника, перед покупкой новой.
- Узнать насколько новая техника соответствует современным эко-стандартам и примут ее на утилизацию после окончания срока службы.

Утилизировать оргтехнику, а не просто выбрасывать на «свалку» необходимо по следующим причинам:

Во-первых, в любой компьютерной и организационной технике содержится некоторое количество драгоценных металлов. Российским законодательством предусмотрен пункт, согласно которому все организации обязаны вести учет и движение драгоценных металлов, в том числе тех, которые входят в состав основных средств. За несоблюдение правил учета, организация может быть оштрафована на сумму от 20000 до 30000 руб. (согласно ст. 19.14. КоАП РФ);

Во-вторых, предприятие также может быть оштрафовано за несанкционированный вывоз техники или оборудования на «свалку»;

Стадия утилизации, утилизируя технику мы заботимся об экологии: количество не перерабатываемых отходов минимизируется, а такие отходы, как пластик, пластмассы, лом черных и цветных металлов, используются во вторичном производстве. Электронные платы, в которых содержатся драгметаллы, после переработки отправляются на аффинажный завод, после чего чистые металлы сдаются в Госфонд, а не оседают на свалках. Таким образом утилизацию компьютера можно провести следующим образом:

- отделить металлические детали от неметаллов;
- разделить углеродистые металлы от цветмета;
- пластмассовые изделия (крупногабаритные) измельчить для уменьшения объема;

- кофир-порошок упаковать в отдельную упаковку, точно также, как и все проклассифицированные и измельченные компоненты оргтехники, и после накопления на складе транспортных количеств отправить предприятиям и фирмам, специализирующимся по переработке отдельных видов материалов.

5.3. Безопасность в чрезвычайных ситуациях

Природная чрезвычайная ситуация – обстановка на определенной территории или акватории, сложившейся в результате возникновения источника природной чрезвычайной ситуации, который может повлечь или повлек за собой человеческие жертвы, ущерб здоровью людей и (или) окружающей природной среде, значительные материальные потери и нарушение условий жизнедеятельности людей.

Производство находится в городе Томске с континентально-циклоническим климатом. Природные явления (землетрясения, наводнения, засухи, ураганы и т. д.), в данном городе отсутствуют.

Возможными ЧС на объекте в данном случае, могут быть сильные морозы и диверсия. Для Сибири в зимнее время года характерны морозы. Достижение

критически низких температур приводит к авариям систем тепло- и водоснабжения, сантехнических коммуникаций и электроснабжения, приостановке работы. В этом случае при подготовке к зиме следует предусмотреть:

1. газобаллонные калориферы (запасные обогреватели);
2. дизель или бензоэлектрогенераторы;
3. запасы питьевой и технической воды на складе (не менее 30 л на 1 человека);
4. теплый транспорт для доставки работников на работу и с работы домой в случае отказа муниципального транспорта. Их количества и мощности должно хватать для того, чтобы работа на производстве не прекратилась.

В летний теплый период в Томске причиной прекращения электроснабжения может стать штормовая погода с сильным порывистым ветром, который может повредить линии электропередач.

5.3. Заключение по разделу «Социальная ответственность»

По разделу «Социальная ответственность» были рассмотрены правовые и организационные вопросы обеспечения безопасности работников. Далее были проанализированы опасные и вредные факторы, которые могут оказывать воздействие на исполнителя исследования ВКР, а также средства защиты от этих факторов. Все это требуется для обеспечения комфортных и безопасных условий труда работников проекта.

Также в данном разделе рассмотрена экологическая безопасность (анализ воздействия объекта на литосферу). Следующим пунктом в разделе были рассмотрены наиболее возможные чрезвычайные ситуации. Наиболее вероятно чрезвычайная ситуация будет иметь природный характер, чаще всего это сильные морозы, способные повредить коммуникации и инфраструктуру города. Соответственно, в данном пункте были разработаны превентивные меры предосторожности по предупреждению ЧС и необходимые действия при возникновении ЧС.

Список публикаций студента

1. Дышеков, А. А. Моделирование распространения эпидемий с использованием сетей / А. А. Дышеков ; науч. рук. М. Е. Семенов // Перспективы развития фундаментальных наук : сборник научных трудов XVIII Международной конференции студентов, аспирантов и молодых ученых, г. Томск, 27-30 апреля 2021 г. : в 7 т. — Томск : Изд-во ТПУ, 2021. — Т. 3 : Математика. — [С. 19-21].
2. Дышеков, А. А. Верификация численного кода для расчёта характеристик дифракционного и переходного излучения / А. А. Дышеков ; науч. рук. Д. А. Шкитов // Перспективы развития фундаментальных наук : сборник научных трудов XVIII Международной конференции студентов, аспирантов и молодых ученых, г. Томск, 27-30 апреля 2021 г. : в 7 т. — Томск : Изд-во ТПУ, 2021. — Т. 1 : Физика. — [С. 109-111].
3. Дышеков А. А. Свойства терагерцового излучения от решётки, генерируемого релятивистским электронным пучком / А. А. Дышеков, Д. А. Шкитов // Физико-технические проблемы в науке, промышленности и медицине. Российский и международный опыт подготовки кадров : сборник тезисов докладов X Международной научно-практической конференции, г. Томск, 09 – 11 сентября 2020 г. — Томск : Ветер, 2020. — [С. 73-74].

Список литературы

1. M.J. Keeling, P. Rohani Modeling infectious diseases in humans and animals. / NJ: Princeton University Press, 2008. 408 pp
2. J.D. Murray Mathematical Biology: An introduction Springer. 2002.
3. W. Liu, H.W. Hethcote, S.A. Levin Dynamical behavior of epidemiological models with nonlinear incidence rates //Journal of mathematical biology. – 1987. – Т. 25. – №. 4. – С. 359-380.
4. Infectious diseases of humans: Dynamics and control. By R.M. Anderson and R.M. May, 757 pp. Oxford: Oxford University Press, 1991.
5. W.H. Hamer Epidemic disease in England. Lancet I, 733- 739. (1906)
6. W. Liu, S. Tang, Y. Xiao Model selection and evaluation based on emerging infectious disease data sets including A/H1N1 and Ebola //Computational and mathematical methods in medicine. – 2015. – Т. 2015.
7. H. Hsieh, J.Y. Lee, H.L. Chang SARS epidemiology modeling //Emerging infectious diseases. – 2004. – Т. 10. – №. 6. – С. 1165.
8. S. Kingsland The refractory model: The logistic curve and the history of population ecology //The Quarterly Review of Biology. – 1982. – Т. 57. – №. 1. – С. 29-52.
9. H.W. Hethcote The mathematics of infectious diseases //SIAM review. – 2000. – Т. 42. – №. 4. – С. 599-653
10. W.O. Kermack, A.G. McKendrick A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character 1927; 115(772):700–21.
11. M. Y. Li, J.S. Muldowney Global stability for the SEIR model in epidemiology //Mathematical biosciences. – 1995. – Т. 125. – №. 2. – С. 155-164.
12. Gabriele Martelloni, Gianluca Martelloni Analysis of the evolution of the Sars-Cov-2 in Italy, the role of the asymptomatics and the success of Logistic model// Chaos, Solitons and Fractals 140 (2020) 110150

- 13.P. Wang, X. Zheng, J. Li, B. Zhu Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics // *Chaos, Solitons and Fractals* 139 (2020) 110058 doi.org/10.1016/j.chaos.2020.110058.
- 14.J. Guo Theoretical Epidemic Laws Based on Data of COVID-19 Pandemic //medRxiv. – 2020.
- 15.Ke Wu, D. Darcet, Q. Wang, D. Sornette Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world// medRxiv preprint doi: <https://doi.org/10.1101/2020.03.11.20034363>
- 16.N.S. Barlow, S.J. Weinstein Accurate closed-form solution of the SIR epidemic model // *Physica D* 408 (2020) 132540
- 17.M. Cadoni, G. Gaeta Size and timescale of epidemics in the SIR framework// *Physica D* 411 (2020) 132626
- 18.T.T. Marinov, R.S. Marinova Dynamics of COVID-19 using inverse problem for coefficient identification in SIR epidemic models // *Chaos, Solitons & Fractals: X* 5 (2020) 100041
- 19.T. Odagaki Analysis of the outbreak of COVID-19 in Japan by SIQR model // *Infectious Disease Modelling* 5 (2020) 691e698
- 20.A.G.M. Neves, G. Guerrero Predicting the evolution of the COVID-19 epidemic with the A-SIR model: Lombardy, Italy and São Paulo state, Brazil// *Physica D* 413 (2020) 132693
- 21.Comunian, R. Gaburro, M. Giudici Inversion of a SIR-based model: A critical analysis about the application to COVID-19 epidemic // *Physica D* 413 (2020) 132674
- 22.Куркина Е.С., Кольцова Е.М. Математическое моделирование и прогнозирование распространения эпидемии коронавируса COVID-19 // Проектирование будущего. Проблемы цифровой реальности: труды 4-й Международной конференции (4-5 февраля 2021 г., Москва). — М.: ИПМ им. М.В.Келдыша, 2021 — С. 178-192.
- 23.M. Agrawal, M. Kanitkar, M. Vidyasagar SUTRA: An Approach to Modelling Pandemics with Asymptomatic Patients, and Applications to COVID-19. <https://arxiv.org/abs/2101.09158>

- 24.Алексеев В.Е, Захарова Д.В. Теория графов: Учебное пособие. – Нижний Новгород: Нижегородский госуниверситет, 2017. –119 с.
25. Erdős P., Rényi A. (1959) On random graphs I. Publ. Math. Debrecen. 6., 290–297.
26. Винницкий, Э., Уайт Р. Введение в моделирование инфекционных заболеваний: учебное пособие. – Оксфорд: Изд-во Оксфордского университета, 2010. – 369 с.
- 27.Edelstein-Keshet L. (2005) Mathematical Models in Biology. SIAM. Society for Ind. and App. Math. 184 p.
- 28.NetworkX [Электронный ресурс] – URL: <https://networkx.org/documentation/stable/index.html> (дата обращения 28.05.2022)
- 29.Яндекс Статистика. Коронавирус [Электронный ресурс] – URL: <https://yandex.ru/covid19/stat> (дата обращения 28.05.2022)
- 30.Оперативный доклад за 2020 год [Электронный ресурс] – URL: <https://ach.gov.ru/audit/oper-2020/> (дата обращения 19.05.2022).
- 31.Портал государственных услуг [Электронный ресурс] – URL: <https://www.gosuslugi.ru/16260/3/> (дата обращения 19.05.2022).
- 32.Официальный сайт Единой информационной системы в сфере закупок [Электронный ресурс] – URL: <https://zakupki.gov.ru/> (дата обращения 19.05.2022).
- 33.Lessons from the pandemic’s superstar data scientist, Youyang Gu [Электронный ресурс] – <https://www.technologyreview.com/2021/04/27/1023657/lessons-from-the-pandemics-superstar-data-scientist-youyang-gu/> (дата обращения 19.05.2022).
- 34.СанПиН 2.2.2/2.4.1340-03. «Гигиенические требования к персональнымэлектронно-вычислительным машинам и организации работы» // Электронный фонд правовой и нормативно-технической документации URL: <http://docs.cntd.ru/document/901865498/> (дата обращения: 03.05.2022).
- 35.СанПиН 2.2.4.548-96. «2.2.4. Физические факторы производственной среды. Гигиенические требования к микроклимату производственных помещений. Санитарные правила и нормы» // Электронный фонд правовой и нормативно-

технической документации. URL: <http://docs.cntd.ru/document/901704046/> (дата обращения: 03.05.2022).

36.ГН 2.2.5.1313-03. «Предельно допустимые концентрации (ПДК) вредных веществ в воздухе рабочей зоны»// Законодательство РФ. URL:https://legalacts.ru/doc/postanovlenie-glavnogo-gosudarstvennogo-sanitarnogo-vracha-rf-ot-30042003-n_8/

37.СН 2.2.4/2.1.8.562-96. «Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки. Санитарные нормы» // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/901703278> (дата обращения: 03.05.2022).

38.ГОСТ 54 30013-83. Электромагнитные излучения СВЧ. Предельно допустимые уровни облучения. Требования безопасности URL:

39.ГОСТ 12.1.004-91. Пожарная безопасность. Общие требования.

40.СанПиН 2.2.4.3359-16 «Санитарно-эпидемиологические требования к физическим факторам на рабочих местах» // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/420362948> (дата обращения: 03.05.2022).

41.ГОСТ Р 55710-2013 «Освещение рабочих мест внутри зданий. Нормы и методы измерений» // Электронный фонд правовой и нормативнотехнической документации. URL: <http://docs.cntd.ru/document/901704046/> (дата обращения: 03.05.2022).

42.СП 52.13330.2016 «Естественное и искусственное освещение» // Электронный фонд правовой и нормативно-технической документации. URL: <http://docs.cntd.ru/document/456054197> (дата обращения: 03.05.2022).

43.СНиП 2.01.02-85 (1991) «Противопожарные нормы»

44.ГОСТ Р 55102-2012 «Ресурсосбережение. Обращение с отходами. Руководство по безопасному сбору, хранению, транспортированию и разборке отработавшего электротехнического и электронного оборудования, за исключением ртутьсодержащих устройств и приборов» // Электронный фонд

правовой и нормативно-технической документации.
<http://docs.cntd.ru/document/1200104723> (дата обращения:03.05.2022).

URL:

Приложение А

Modeling the spread of epidemics using networks

Студент

Группа	ФИО	Подпись	Дата
0BM02	Дышеков Астемир Анзорович		

Руководитель ВКР:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семёнов М.Е.	к.ф.-м.н., доцент		

Консультант-лингвист отделения иностранных языков ШБИП:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИЯ	Панамарёва А.Н.	к.ф.н., доцент		

Introduction

The COVID-19 pandemic that began in 2019 is still ongoing. Pandemic waves follow one after another. The advent of a vaccine has not led to the end of a pandemic that is causing significant economic damage. Human capital is vulnerable to the virus. Its reproduction requires resources and time. To effectively combat the spread of coronavirus, an accurate forecast of the epidemiological situation is needed. Health services must be prepared for the new wave of health. Stocks of medicines, oxygen cylinders, ventilators should be enough to overcome the peak of diseases. The number of medical personnel mobilized to combat the spread of the disease should also be sufficient. The stock of beds should slightly exceed the potential maximum number of patients whose condition is serious and requires inpatient treatment. Regional authorities and control services should introduce restrictive measures in a timely manner to reduce the rate of growth in the number of infected people. All these measures are very expensive or cause indirect economic damage (restrictive measures). Therefore, the scope of these measures should be justified in order to minimize costs. For example, skilled health workers are limited resources that can only be reallocated to address specific health challenges. It is impossible to increase the number of doctors in a country in a short time. Therefore, it is very important to allocate this resource to fight against the pandemic in the minimum required number. The more doctors sent to fight the pandemic, the less accessible healthcare becomes for more members of society, which can provoke excess deaths from other diseases. Therefore, regional authorities, regional ministries of health and regional departments of Rospotrebnadzor need accurate forecasts on the spread of a new coronavirus infection for the rational use of available resources.

The purpose of this master's thesis is to develop and implement a SIR-like model using networks as a software product to predict the spread of a new coronavirus infection.

In the presented work, a model has been developed and implemented that allows predicting the dynamics of the spread of a new coronavirus infection using

network data. The model is implemented as a software product that is able to predict the epidemiological situation in a particular city using static data for a given locality.

1. Literature review

There are quite a few different mathematical models [1-11] that allow predicting and studying the process of the spread of epidemics in human society at a particular stage of them or in the context of their evolution. Of course, with the advent of the new coronavirus infection COVID-19, mathematicians began to offer different models for describing the process of the spread of this infection. The first models began to appear since the first outbreak in China. The pandemic is not over yet, so the development and improvement of models also continues intensively.

In [12-15], the Verhulst logistic growth model or its modifications were used. This model assumes that at the initial stage of the spread of the epidemic, the growth in the number of cases occurs exponentially. But due to population restrictions, the rate of infection falls and the number of infected people reaches a plateau.

Compartmental models (SIR-like) are the most popular for predicting the epidemiological situation: SIR [16-18], SIRS, SEIR, SAIR [19-22] and their modifications [23-30]. In almost all of these publications, the models are based on a system of differential equations. Basically, these models differ in correction parameters (to take into account the specifics of the spread of coronavirus and the environmental factor) and methods for calculating the coefficients in the equations. All papers lack pure compartmental models that assume that the infection environment is homogeneous and that each element is related to each other. In [21], the authors conclude that compartmental models with constant coefficients are poorly applicable to predicting the spread of COVID-19.

Russian scientists in their work [22] propose to model the spread of coronavirus as a superposition of many waves based on the Feigenbaum discrete logistic equation. The authors consider it is impossible to use models for predicting coronavirus, in the concept of which there is only one peak of incidence. The paper proposes a method for

determining local waves and calculating their parameters. The addition of new waves occurs as soon as the calculations begin to deviate significantly from the actual values. The authors also criticize models based on statistical data, as they initially have an error of 20-80% associated with data collection.

The Brazilian mathematician in his work [23] proposed his network model for predicting coronavirus, which is similar to our model. The model also proposes to consider the vertices of the graph as separate individuals, and the edges of the graph as links between them. The model also uses the principles of the SIR model. The author of the model from Brazil proposes to consider not one graph, but several, each of which describes relationships of different levels: family, work, transport, etc. This differs from the concept of our model, which is based on one single graph. The author of the work also notes a disadvantage of the network approach of the model - the need for large computing power. In his work, he made calculations using graphs whose number of vertices did not exceed 100,000. To scale the predictions to large populations, the paper proposes to multiply the results by coefficients. The coefficients are equal to the ratio of the statistical data for a large population to the simulation results for a small population.

2. Random graph with clique structure

Random graphs are the models of real networks that allow you to describe the properties and reproduce the structure of these networks. Based on the properties of the real structure, the most suitable random graph is selected. In this paper, it was necessary to choose a graph model that best reproduces public relations in terms of the spread of infections.

The classical random graph is the Erdős-Rényi graph [25]. Random Erdős-Rényi graphs are formed as follows:

- fix N graph vertices;
- any pair of vertices is randomly connected by an edge with equal probability p .

The Erdős-Renyi model poorly reflects the structure of society, namely the heterogeneity of ties. In the Erdős-Rényi graph, the connections are evenly distributed. In a real society, the structure has social cells (family or household). Within families, we have a higher density of ties. Each individual is interconnected. If we present the household as a subgraph, then it will be a clique by definition. Thus, to recreate the structure of society, it is necessary to generate a graph, each vertex of which is part of a clique. Recall that a clique of a graph is a set of vertices that generates a complete subgraph. Thus, to recreate the structure of society, it is necessary to generate a graph, each vertex of which is part of a clique. In this paper, a clique is a complete subgraph with 2 or more vertices. An average statistical family consisting of 2 parents and 2 children can be represented by a clique of dimension 4. Each individual, in addition to internal ties in the family, of course, has constant external ones. Most often, these are connections with friends, other relatives and colleagues. To describe these relationships the Erdős-Rényi model is used. Therefore, it was proposed to take a random Erdős-Renyi graph as a basis and complete cliques in it.

The basis is a random Erdősche-Rényi graph G , which is randomly generated with N vertices and $d(m)N$ edges, where $d(m)$ is the average degree of the vertex m . Further, this graph is modified through a cycle that completes cliques in it. During the loop, a list of edges is generated to form cliques. At each iteration, a random number $k \sim U(2, max)$ is generated, vertices with numbers $i, i+1, \dots, i+k$ are closed in a clique on k vertices. Iterations continue until a clique of the specified size can be formed from the remaining vertices, i.e. while $N-i > max$, where max is the clique number that is set before graph generation. After the cycle ends, the remaining vertices also complete the clique. At the exit from the loop, we get a random graph with cliques that are located along the main diagonal of the vertex adjacency matrix. Random number k can be generated according to any distribution law. In a simple case, it has a uniform distribution of the number of clique vertices in the graph from 2 to max .

3. Base Model SEIRS

The model proposed by SEIRS (*Susceptible- Exposed - Infected -Recovered- Susceptible*) is based on the concepts of SIR-like models. The model is based on the graph G , which is generated by the algorithm described above with N vertices and $d(m)N$ edges, where $d(m)$ is the average degree of the vertex m . The vertex of the graph is associated with a representative of the population, which as a whole is presented by the graph. The edges of the graph characterize stable (frequent) connections between people in the society. Each vertex has a categorical status: *susceptible*, *infected*, or *recovered*. To visualize the vertex status, we used colors: blue, red and green, respectively (Fig. 1).

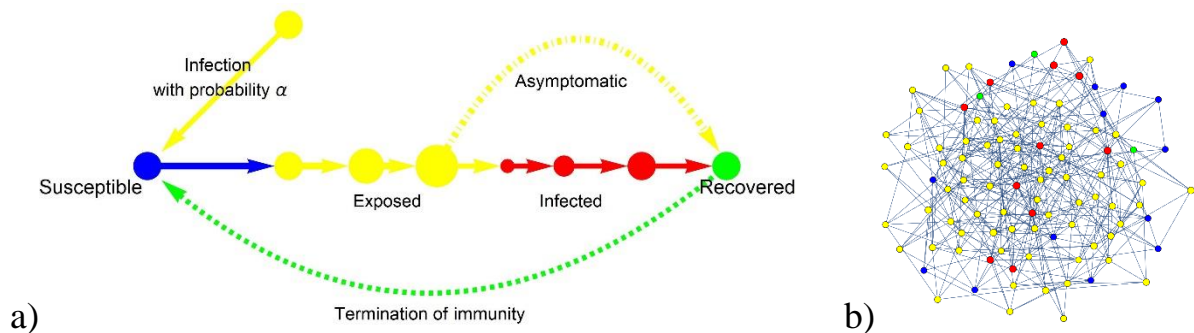


Figure 1. (a) General scheme of the SEIRS model, (b) example of a graph G with $n=100$, $d(m)=5$

In the classic SIR model, people with a susceptible status can also become ill when in contact with infected people, and vice versa, infected people can recover. In our model, if the blue vertex is adjacent to the red one, then it can be infected with a probability of $0 \leq P(\alpha) \leq 1$ in one day. The second important characteristic of a vertex is its weight. The weight of the blue vertices is constant and equal to zero. For other vertices, the weight changes dynamically and is equal to the number of days they arrive in a certain status. For example, if the weight of the red vertex is 7, this individual has already been infected for 7 days. Counting the number of days is necessary to account for recovery. The weight of the red peaks was compared with the average duration of

the disease - 21 days. When the disease duration reaches the value of the average disease duration, the individual recovers and the color of the vertex changes to green.

The SIR model is quite limited, it can describe only few infections. However, SIR has many derivative models that complement its functionality. One of these models is SEIR, which uses another status of people - exposed. In this model, an important concept is found as the incubation period - the period of time from the moment of infection to the appearance of overt symptoms. Often during the incubation period, a person can already infect (as in the case of Covid-19), which greatly complicates the timely isolation of the patient and control of the spread of infection becomes difficult. In our model, we propose to add a new status that characterizes patients in the incubation period (yellow vertices, Fig. 1b). In this case, the characteristics of the yellow vertices are very similar to the red ones. The key difference is that after the yellow vertex reaches a weight equal to the incubation period - 5 days – it changes the status to infected (red), and not to recovered (green). Also in the proposed model, the following condition was introduced: only contacts (yellow) can infect, thereby simulating the process of isolating a patient with open symptoms.

Further, in the proposed model, we introduced a parameter that characterizes the proportion of the population that will completely recover from the disease in a latent form (20% of the population). The SIRS model is a development of the SIR model, which introduces a temporary limitation on the effects of immunity - the average duration of immunity is taken to be 365 days, i.e. in the SIRS model, those who have recovered may become susceptible to the disease again after some time. This is very common with the flu, and possibly Covid-19 as well. In the proposed model, we used the indicated feature of the SIRS model and introduced a recalculation of the weight of green vertices: when the weight of the average duration of immunity is reached, the color of the vertex changes from green to blue, the weight of the vertex is taken equal to zero. Thus, we get a cycle of possible statuses of an individual, which allows us to take into account the possible recurrence of the disease in the simulation.

4. SEIRS model with immune status

In the basic SEIRS model, recovered patients cannot become infected until a time equal to the duration of immunity has passed. In reality, there is always a chance of re-infection with coronavirus after a short period of time after recovery. Therefore, it was proposed to conduct an immunity function, which depends on time and takes values from 0 to 1. The function characterizes the probability of infection, and it can be specified through a linear function:

$$f_{im}(t) = \begin{cases} 1 - \frac{t - t_0}{T}, & \text{when } t - t_0 \leq T \\ 0, & \text{when } t - t_0 > T \end{cases}$$

Where t_0 is the day of recovery, T is the duration of the immunity. The recovery time is entered as one of the attributes of the graph vertices. Taking into account immunity, the probability of vertex infection is now equal to $P(\alpha) (1 - f_{im}(t))$. With the addition of the immunity function, susceptible and recovered patients began to differ only in the value of the t_0 attribute, so R - SEIS can be removed from the model's name.

With this version of the model, it makes sense to pass static data through the graph so that the network characteristics become as close as possible to real data. Also, such "training" of the graph allows you to recalculate the value of the parameter α . The parameter α can be calculated as the ratio to the number of infected people a day before a possible infection. Figure 2 shows the calculated values of this parameter over time, according to the statistics for the city of Tomsk since the first infections. As you can see in the graph, it has scattered values, which is primarily due to the error of the statistics data. A particularly big mistake was at the initial stage of the pandemic, because there were no methods and tests for the rapid diagnosis of coronavirus. The graph also shows the latest surge in the infectivity of the coronavirus, caused by a new strain of omicron.

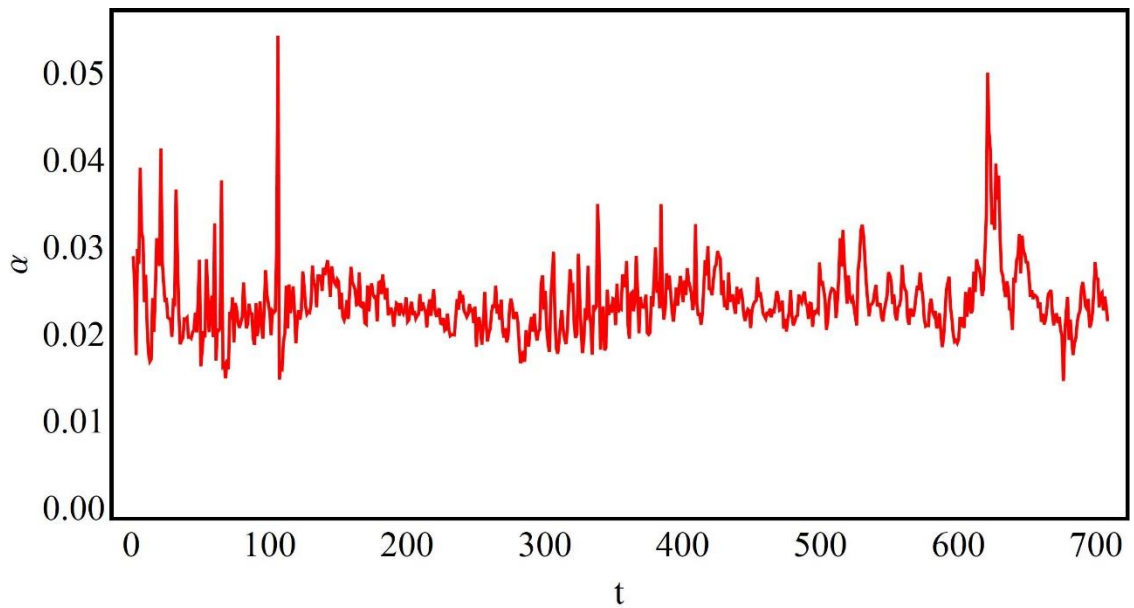


Figure 2. Calculated values of the parameter α based on statistics data for the city of Tomsk

5. Model implementations in Python

To implement the SEIRS model in Python, it was proposed to use the NetworkX library [28]. NetworkX library is designed to create, manipulate and study the structure, dynamics and functioning of complex network structures. The library can freely handle very large network structures, graph level with 10 million nodes and 100 million arcs between them. Since it is based on a low-level data structure of the Python language called "dictionary-dictionaries". This allows for efficient memory usage and good scaling of the graphs. This data structure is independent of the operating system and is well suited for graph analysis.

In NetworkX, a graph is defined by a vertex dictionary and an edge dictionary. The graph itself, each vertex and each edge in the dictionary can have any number of attributes. For example, the weight and color of the top. Knowing the name of the vertex (more often the number), you can find out the values of the desired attribute and replace it. The implementation of the model is based on this functionality. Each vertex of the graph in the model has an attribute, the values of which are responsible for the infectious status. For convenience, the values of this

attribute are not taken as letters of the abbreviation of the model's name, but numbers are used: 0 (S), 1 (E), 2 (I), 3 (R). The second important attribute of graph vertices is the time spent in a given state. When transitioning from one state to another, this attribute is reset to zero. After one iteration (equal to one day) it increases by one.

6. Accounting for restrictive measures

In the classical version of the SIR models, the virus spreads under relatively ideal conditions without restrictive ones, which, of course, is far from reality. In our model, we propose to use several mechanisms that imitate certain tools to limit the spread of infection.

To take into account the measures that affect the probability of infection, at the initial stage of the development of the model, a coefficient of restrictive measures was introduced, equal to the ratio of the probability of infection, taking into account the event, to the probability of infection without taking the event into account. This coefficient is less than 1 if the activity restricts distribution, and greater than 1 otherwise. Based on empirical observations and available statistical data, an expert assesses the degree of influence of certain restrictive measures and events on the likelihood of infection of an individual.

Figure 3a shows a comparison of the dynamics of the infected $I(t)$ in the city of Tomsk according to official statistics and model results using this coefficient of restrictive measures, the values of which are presented in the graph. In general, there is a good convergence of the results of modeling the dynamics of infection with statistics. The discrepancy in the initial period (month - May) is due to the lack of mass testing methods for coronavirus at the beginning of the pandemic. In the summer period (June-August), official statistics data exceed the results. The figure shows data for Tomsk at the initial stage of the pandemic, when severe restrictive measures began to be applied. The values of the coefficient were selected based on the observer's empirical assessment of the events taking place.

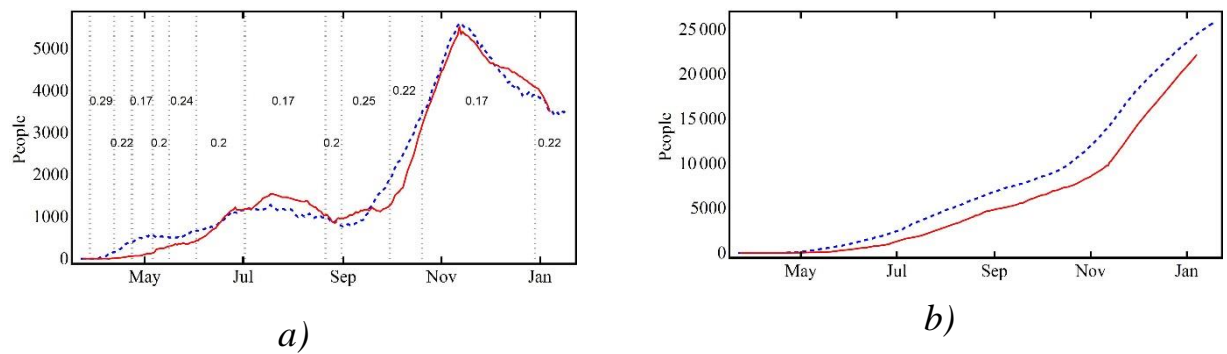


Figure 3. Dynamics of change in the number of (a) the infected and (b) the recovered in Tomsk from March 2020 to January 2021: official data (red line) and simulation results (dashed line)

In April 2021, forecasts were made on the dynamics of infected people using the coefficient of restrictive measures (see Fig. 3). Four assumptions were made:

- all lockdown measures are preserved;
- lockdown measures will soften;
- lockdown measures will be strengthened.

The "pessimistic" scenario ended up being the closest to reality, as restrictions were eased and warmer weather helped disrupt the remaining ones.

The demonstrated simulation results show that the coefficient of restrictive measures can be used to take into account activities and events affecting the spread of coronavirus. However, this method of accounting for constraints is highly subjective and highly dependent on the assumptions of a person using the model. Therefore, it was decided to move away from this method and introduce other mechanisms for taking into account the restrictions on the spread of coronavirus into the model.

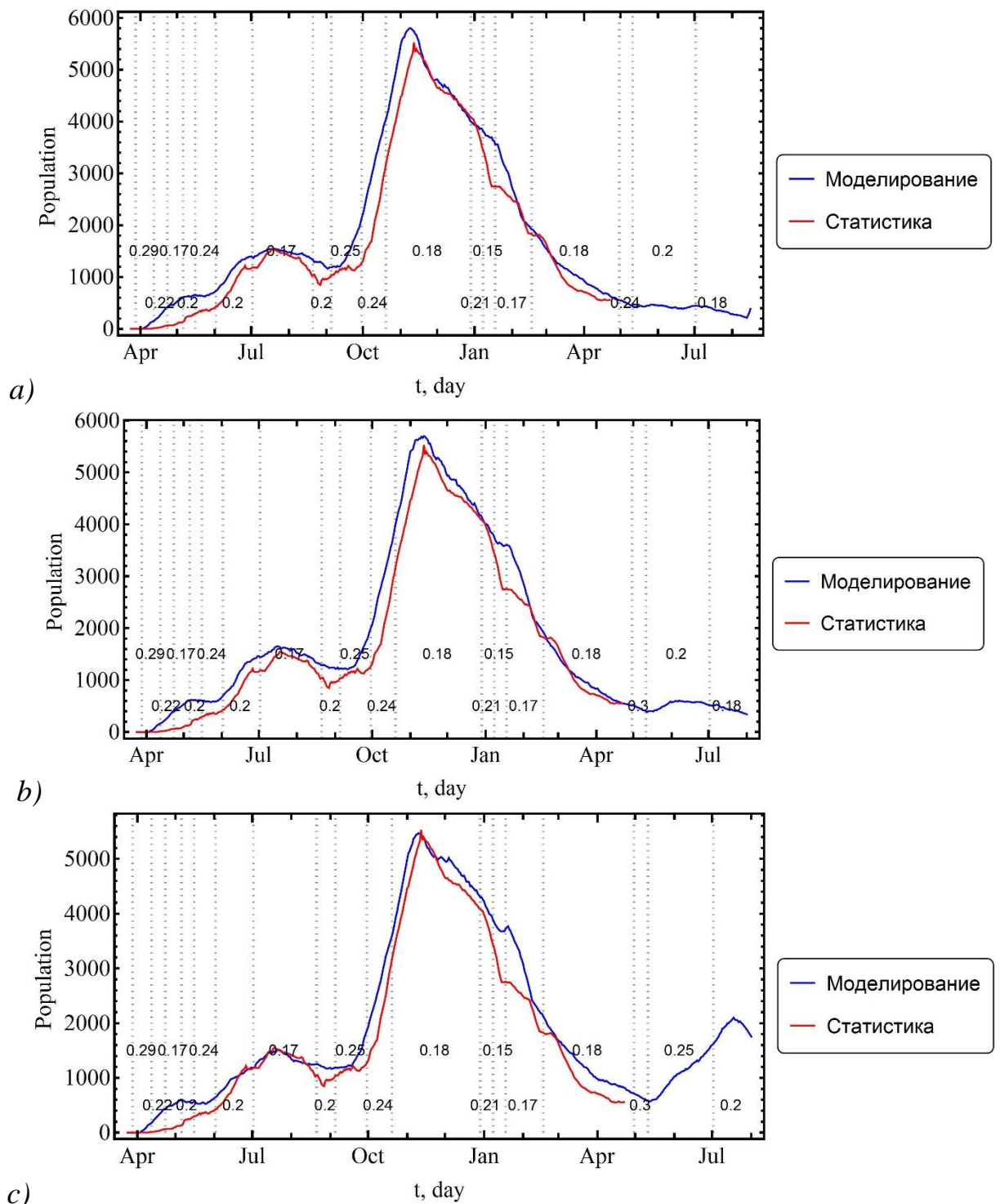


Figure 4. Morbidity forecasts for May 2021: optimistic (a), moderate (b) and pessimistic (c)

The first barrier that limits the spread of the infection is the vigilance and caution of the individual. To take into account the vigilance of individuals, the environmental factor μ was added. This factor is equal to the proportion of adjacent

vertices that are not officially infected (have status 0 or 1). The probability of the infection in this case is: $P(\alpha\mu)=P(\alpha)P(\mu)$.

The effectiveness of identifying patients significantly affects the spread of the infection. In our model, when the vertex goes into the stage of acute illness, all those who have been in contact with an infected person are checked. And with probability $P(\beta)$, those who are infected, but still in the incubation period (vertices with status 1), can be identified. This parameter, as well as the environment factor, perfectly demonstrates what instrumental capabilities the model has with the use of network data structures.

With the advent of vaccines, people have a new barrier to the spread of coronavirus. With an increase in the number of vaccinated and recovered patients, the effect of herd immunity occurs, in which it is difficult for the virus to be transmitted to susceptible ones. There are two ways to account for the proportion of those vaccinated. You can directly set the percentage of vaccinated when generating a random graph or run the vaccination statistics through the graph. This makes it possible to bring the characteristics of the graph closer to the real epidemiological situation.

7. Model Verification

To test the performance and accuracy of the model, we simulated the spread of coronavirus for the Tomsk and Novosibirsk agglomerations. Fig. 7 presents a comparison of simulation results and statistics. The statistics data were taken from [29]. As can be seen, the simulation results qualitatively coincide with the statistical data at the initial stage of the wave. The location of the maxima of statistics and calculations is approximately the same. The maximum quantitative values of infected people in statistics and in simulation do not differ much and lie within the statistical error. On the decline of the wave, we observe the divergence of the statics and simulation. The statistical curve falls off more quickly.

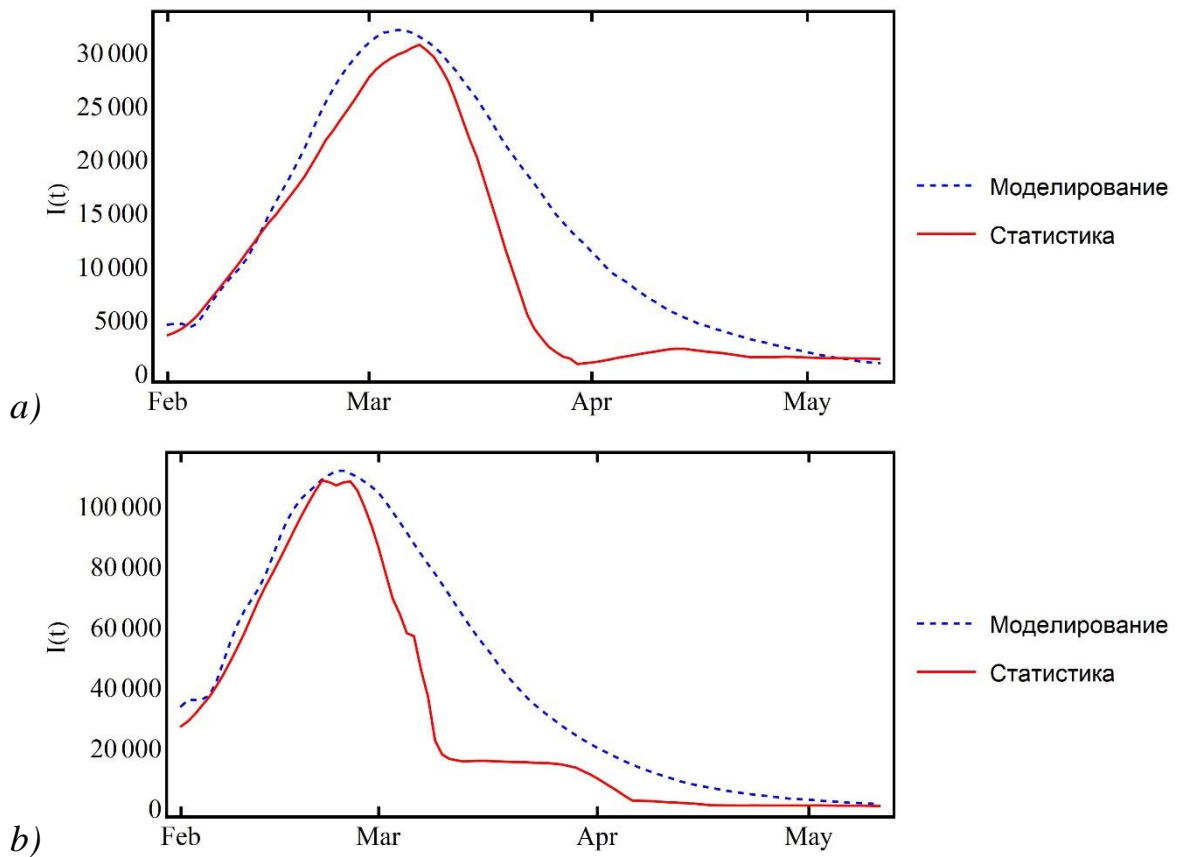


Figure 5. Comparison of the simulation results and statistics on the number of infected people in Tomsk (a) and Novosibirsk (b)

To confirm the hypothesis of the reason for the discrepancy between the statistical data and the simulation results, repeated calculations were made, but taking into account the introduction of restrictive measures. The results are presented in fig. 6. The simulation results in this case are well comparable with the statistical data. This demonstrates that the mechanisms for taking into account restrictive measures in the model work. The proposed hypothesis about the impact of the transfer of part of citizens to remote work is true.

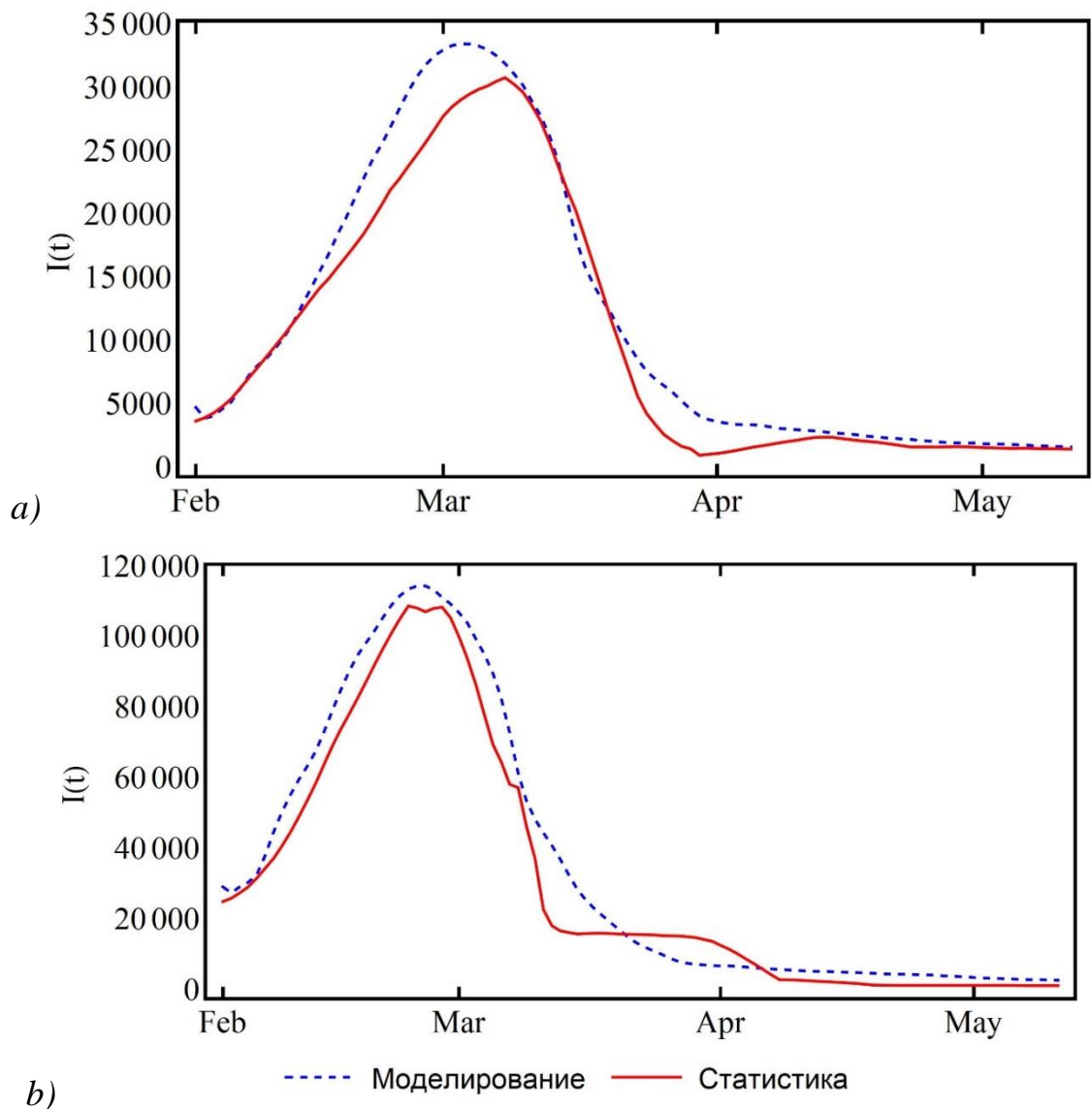


Figure 6. Comparison of the simulation results (taking into account restrictive measures) and statistics on the number of infected people in Tomsk (a) and Novosibirsk (b)

Conclusion

During the work on this master's thesis, a network-based model was developed to describe the spread of coronavirus infection. The model was implemented as a program code in Python. During the verification of the model, the qualitative and quantitative convergence of the simulation results and statistical data was obtained, which confirms the ability of the model to describe the real processes of the spread of infections.