

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное научное учреждение  
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР  
«КРАСНОЯРСКИЙ НАУЧНЫЙ ЦЕНТР СИБИРСКОГО ОТДЕЛЕНИЯ  
РОССИЙСКОЙ АКАДЕМИИ НАУК»  
(КНЦ СО РАН, ФИЦ КНЦ СО РАН)

УДК 519.63:519.64  
Пер. № НИОКТР 121081900115-1  
Пер. № ИКРБС

УТВЕРЖДАЮ

Директор ФИЦ КНЦ СО РАН  
член-корреспондент РАН

\_\_\_\_\_ А.А. Шпедт

«\_\_\_\_» \_\_\_\_\_ 2023 г.

ОТЧЕТ  
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

КОМПЛЕКСНОЕ МАТЕМАТИЧЕСКОЕ, ИНФОРМАЦИОННОЕ, ВЫЧИСЛИТЕЛЬНОЕ  
МОДЕЛИРОВАНИЕ СЛОЖНЫХ ТЕХНИЧЕСКИХ, ФИЗИЧЕСКИХ,  
ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ И СИСТЕМ  
(промежуточный, этап 2)

Направление фундаментальных и поисковых исследований  
1.1.3. Математическое моделирование  
(№FWES-2021-0003)

Руководитель НИР  
Директор ИВМ СО РАН

\_\_\_\_\_ В.В. Шайдуров

Красноярск 2023

## СПИСОК ИСПОЛНИТЕЛЕЙ

Руководитель темы член-корреспондент РАН	_____	В.В. Шайдуров (Раздел 1)
	<i>(подпись, дата)</i>	
Исполнители:		
зам. директора по научной работе ИВМ СО РАН, к.т.н.	_____	С.В. Исаев (Раздел 4)
	<i>(подпись, дата)</i>	
зам. директора по научной работе ИВМ СО РАН, к.ф.-м.н.	_____	Д.А. Нестеров (Раздел 1, 2)
	<i>(подпись, дата)</i>	
г.н.с., д.ф.-м.н.	_____	Н.В. Еркаев (Раздел 1)
	<i>(подпись, дата)</i>	
зав. отд., к.ф.-м.н.	_____	Е.Д. Карепова (Раздел 1)
	<i>(подпись, дата)</i>	
в.н.с., д.т.н.	_____	Л.Ф. Ноженкова (Раздел 3)
	<i>(подпись, дата)</i>	
г.н.с., д.т.н.	_____	А.В. Лапко (Раздел 3)
	<i>(подпись, дата)</i>	
в.н.с., д.ф.-м.н.	_____	В.В. Денисенко (Раздел 1)
	<i>(подпись, дата)</i>	
в.н.с., д.ф.-м.н.	_____	И.В. Краснов (Раздел 2)
	<i>(подпись, дата)</i>	
в.н.с., д.т.н.	_____	В.А. Лапко (Раздел 3)
	<i>(подпись, дата)</i>	
в.н.с., д.ф.-м.н.	_____	И.И. Рыжков (Раздел 3)
	<i>(подпись, дата)</i>	
уч. секр., к.ф.-м.н.	_____	А.В. Вяткин (Раздел 1)
	<i>(подпись, дата)</i>	
с.н.с., к.ф.-м.н.	_____	Е.Н. Васильев (Раздел 1, 2)
	<i>(подпись, дата)</i>	
с.н.с., к.ф.-м.н.	_____	В.С. Герасимов (Раздел 2)
	<i>(подпись, дата)</i>	
с.н.с., к.ф.-м.н.	_____	В.А. Деревянко (Раздел 1, 2)
	<i>(подпись, дата)</i>	

с.н.с., к.ф.-м.н.	_____	А.Е. Ершов (Раздел 2)
	(подпись, дата)	
с.н.с., к.т.н.	_____	О.С. Исаева (Раздел 3, 4)
	(подпись, дата)	
с.н.с., к.т.н.	_____	А.В. Коробко (Раздел 3)
	(подпись, дата)	
зав. отделом, к.т.н.	_____	Т.Г. Пенькова (Раздел 3)
	(подпись, дата)	
с.н.с., к.ф.-м.н.	_____	М.Ю. Сенашова (Раздел 3)
	(подпись, дата)	
н.с., к.ф.-м.н.	_____	Л.В. Гилева (Раздел 1)
	(подпись, дата)	
н.с., к.ф.-м.н.	_____	И.Л. Исаев (Раздел 2)
	(подпись, дата)	
н.с., к.ф.-м.н.	_____	Л.П. Каменщиков (Раздел 2)
	(подпись, дата)	
н.с., к.т.н.	_____	В.В. Морозов (Раздел 3)
	(подпись, дата)	
н.с., к.т.н.	_____	А.И. Ноженков (Раздел 3)
	(подпись, дата)	
н.с.	_____	Д.Д. Кононов (Раздел 4)
	(подпись, дата)	
м.н.с.	_____	В.И. Закомирный (Раздел 2)
	(подпись, дата)	
м.н.с., к.ф.-м.н.	_____	С.В. Козлова (Раздел 2)
	(подпись, дата)	
м.н.с., к.ф.-м.н.	_____	В.С. Петракова (Раздел 1)
	(подпись, дата)	
м.н.с.	_____	Н.П. Фадеева (Раздел 2)
	(подпись, дата)	
м.н.с.	_____	М.В. Якубович (Раздел 1)
	(подпись, дата)	

прогр. 1-ой кат.	_____	А.А. Коробко (Раздел 3)
	(подпись, дата)	
прогр. 1-ой кат.	_____	Н.В. Кулясов (Раздел 4)
	(подпись, дата)	
прогр. 1-ой кат.	_____	И.А. Ларионова (Раздел 4)
	(подпись, дата)	
прогр. 1-ой кат.	_____	А.В. Малышев (Раздел 4)
	(подпись, дата)	
прогр. 1-ой кат.	_____	А.С. Михалев (Раздел 3)
	(подпись, дата)	
прогр. 1-ой кат.	_____	А.М. Метус (Раздел 3)
	(подпись, дата)	
ведущий инженер	_____	А.В. Макуха (Раздел 1, 2)
	(подпись, дата)	
старший инженер	_____	Д.П. Емельянов (Раздел 1, 2)
	(подпись, дата)	
старший инженер	_____	С.В. Кукушкин (Раздел 1, 2)
	(подпись, дата)	
старший инженер	_____	А.А. Сиротинин (Раздел 1, 2)
	(подпись, дата)	
старший инженер	_____	И.В. Тимошина (Раздел 1)
	(подпись, дата)	
инженер-исследователь	_____	И.Р. Адаев (Раздел 1)
	(подпись, дата)	
инженер-исследователь	_____	И.А. Харченко (Раздел 1)
	(подпись, дата)	
инженер	_____	Р.А. Голубев (Раздел 1)
	(подпись, дата)	
нормоконтроль	_____	А.В. Вяткин
	(подпись, дата)	

## РЕФЕРАТ

Отчёт 73 с., 27 рис., 3 табл., 3 прил.

### МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ЭЛЕКТРОМАГНИТНЫЕ ПОЛЯ, НАНОЧАСТИЦЫ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ, МОДЕЛИ БЕЗОПАСНОСТИ КОРПОРАТИВНОЙ СЕТИ

В соответствии с планом работ на 2022 год исследования проводились по четырем разделам.

В разделе «Математическое моделирование процессов и устройств в условиях ближнего космоса» представлена модель электрических полей и токов Глобальной электрической цепи Земли, проведено математическое моделирование диссипативных процессов, вызывающих генерацию электрического поля на границе и в хвосте магнитосферы. Приведены также алгоритмы, реализующие длительное движение спутника по возмущенной орбите с выбором симметричных методов интегрирования уравнений движения.

В разделе «Математическое моделирование физических процессов и устройств» выявлены условия максимального усиления локального поля за счет настройки поверхностного решеточного резонанса путем варьирования геометрических параметров периодических структур из резонансных частиц, а также представлены математические модели ионного транспорта в наноструктурированных мембранных материалах под действием электрического поля.

В разделе «Исследование больших данных, обнаружение новых знаний и прогнозирование состояния сложных систем и объектов» представлены результаты популяционно-эволюционного анализа геномов вирусов семейства Coronaviridae, а также модификации непараметрических алгоритмов автоматической классификации в условиях больших объёмов статистических данных. Предложены метамодель и алгоритмы детализации интегральных оценок для определения причин состояния природно-техногенной безопасности территорий.

В разделе «Технологии создания информационно-аналитических систем на основе интегрированных платформ» представлены методические и программные средства поддержки формирования аналитических отчетов в модельно-ориентированной платформе. Представлено программное обеспечение трансформации данных различных форматов для построения аналитической модели. Предложена модель безопасности системы облачных вычислений для различных способов обслуживания, проведено исследование безопасности сервисов корпоративной сети Красноярского научного центра (ФИЦ КНЦ СО РАН).

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	8
1. Математическое моделирование процессов и устройств в условиях ближнего космоса.....	10
1.1 Модели электрических полей и токов Глобальной электрической цепи Земли.....	10
1.2 Математическое моделирование диссипативных процессов, вызывающих генерацию электрического поля на границе и в хвосте магнитосферы.....	13
1.3 Алгоритмы, реализующие длительное движение спутника по возмущенной орбите с выбором методов интегрирования уравнений движения.....	17
2 Математическое моделирование физических процессов и устройств.....	22
2.1 Условия максимального усиления локального поля за счет настройки поверхностного решеточного резонанса путем варьирования геометрических параметров структуры.....	22
2.2 Математические модели ионного транспорта в наноструктурированных мембранных материалах под действием электрического поля.....	26
3 Исследование больших данных, обнаружение новых знаний и прогнозирование состояния сложных систем и объектов.....	31
3.1 Популяционно-эволюционный анализ геномов вирусов семейства Coronaviridae.....	31
3.2 Модификации непараметрических алгоритмов автоматической классификации в условиях больших объёмов статистических данных .....	36
3.3 Мета модель и алгоритмы детализации интегральных оценок для определения причин состояния природно-техногенной безопасности территорий.....	42
4 Технологии создания информационно-аналитических систем на основе интегрированных платформ.....	46
4.1 Методические и программные средства поддержки оперативной аналитической обработки данных научных исследований в рамках модельно-ориентированных систем сбора данных.....	46
4.2 Программное обеспечение трансформации данных различных форматов путем сопоставления структурных элементов и выполнения преобразований по установленным взаимным отношениям .....	49
4.3 Исследование кибербезопасности систем облачных вычислений для различных способов обслуживания .....	53
4.4 Исследование безопасности сервисов корпоративной сети Красноярского научного центра (ФИЦ КНЦ СО РАН).....	56
ЗАКЛЮЧЕНИЕ.....	63

Приложение А. Научные публикации, выполненные в рамках проекта, в журналах, индексируемых в российских и международных информационно-аналитических системах научного цитирования.....	67
Приложение Б. Научные публикации, опубликованные вне рамок выполнения проекта.....	72
Приложение В. Свидетельства о регистрации программ.....	73

## ВВЕДЕНИЕ

В соответствии с планом работ на 2022 год по теме "Комплексное математическое, информационное, вычислительное моделирование сложных технических, физических, производственных процессов и систем" исследования проводились по следующим разделам.

### 1. Математическое моделирование процессов и устройств в условиях ближнего космоса.

В этом разделе рассмотрены модели электрических полей и токов Глобальной электрической цепи Земли, проведено математическое моделирование диссипативных процессов, вызывающих генерацию электрического поля на границе и в хвосте магнитосферы, а также построены алгоритмы, реализующие длительное движение спутника по возмущенной орбите с выбором симметричных методов интегрирования уравнений движения. В первом направлении построены глобальные карты грозовой активности на основе данных наземной Всемирной сети определения местоположения молний (World Wide Lightning Location Network, WWLLN) за 2007-2009 годы. Во втором направлении в рамках Холл-МГД-модели изучена проблема магнитного пересоединения магнитосферы, исследовано влияние неоднородной проводимости и параметра Холла на процесс магнитного пересоединения при различных начальных магнитных конфигурациях. В третьем направлении разработан ряд алгоритмов получения и исследования устойчивости (периодичности) симметричных методов и получаемых на их основе схем предиктор-корректор.

### 2. Математическое моделирование физических процессов и устройств.

Здесь выявлены условия максимального усиления локального поля за счет настройки поверхностного решеточного резонанса путем варьирования геометрических параметров периодических структур из резонансных частиц, а также представлены математические модели ионного транспорта в наноструктурированных мембранных материалах под действием электрического поля. В первом направлении для решения задач рассеяния периодическими структурами наночастиц использован метод конечных разностей во временной области (FDTD) для расчета спектров экстинкции, отражения и поглощения в наноструктурах. А во втором направлении в рамках сотрудничества с коллегами из Европейского центра передовых технологий водоподготовки (Wetsus) и Университета Твенте (Нидерланды) разработана модель транспорта ионов через полимерную мембрану обратного осмоса.



### 3. Исследование больших данных, обнаружение новых знаний и прогнозирование состояния сложных систем и объектов.

В разделе представлены результаты популяционно-эволюционного анализа геномов вирусов семейства Coronaviridae, модификации непараметрических алгоритмов автоматической классификации в условиях больших объёмов статистических данных, а также когнитивная модель и алгоритмы формирования управляющих решений на основе анализа оперативных наблюдений и комплексного оценивания состояния природно-техногенной безопасности территорий. В анализе геномов проведено сравнение внутреннего структурирования 69 геномов коронавируса, выявленных разными методами сравнения: 1) традиционное выравнивание с последующим построением филогении, 2) неконтролируемая классификация на основе k-средних и 3) современный метод нелинейной статистики на основе упругих карт. Эффективность разработанной методики автоматической классификации подтверждена результатами её применения при анализе данных дистанционного зондирования лесных массивов, повреждённых сибирским шелкопрядом. А разработанная модель состояния природно-техногенной безопасности является основой информационной поддержки по планированию мероприятий долгосрочного развития территорий и формированию стратегий снижения рисков чрезвычайных ситуаций.

### 4. Технологии создания информационно-аналитических систем на основе интегрированных платформ.

Здесь представлены методические и программные средства поддержки формирования аналитических отчетов в модельно-ориентированной платформе. Представлено программное обеспечение трансформации данных различных форматов для построения аналитической модели. В этом же разделе представлена модель безопасности системы облачных вычислений для различных способов обслуживания. Апробация и внедрение метода адаптивных рассылок позволила сократить количество передаваемых данных и периодов активности устройств IoT в пять раз. Проведено исследование безопасности сервисов корпоративной сети Красноярского научного центра (ФИЦ КНЦ СО РАН). С этой целью для обработки данных разработан программный комплекс автоматизированной классификации кибератак на облачные веб-сервисы на основе сигнатурного метода.

## **1 Математическое моделирование процессов и устройств в условиях ближнего космоса**

### **1.1 Модели электрических полей и токов Глобальной электрической цепи Земли**

Построены глобальные карты грозовой активности на основе данных наземной Всемирной сети определения местоположения молний (World Wide Lightning Location Network, WWLLN) за 2007-2009 годы. Это годы с низкой солнечной активностью, включая абсолютный минимум в 2009 году. Такие карты созданы для разных сезонов и периодов универсального времени. Общее количество вспышек молний в данных WWLLN на порядок меньше, чем в данных спутниковой системы OTD/LIS. Отметим, что детальные данные системы OTD/LIS нам недоступны, как и данные WWLLN за другие годы. Тем не менее, сравнение с опубликованными результатами позволяет утверждать, что ключевые особенности пространственного распределения и сезонные тенденции, полученные наземными и спутниковыми методами, находятся в удовлетворительном соответствии. Основное различие заключается в отсутствии суточных вариаций, аналогичных кривой Карнеги, в данных WWLLN при их наличии в спутниковых данных. Это касается количества молний как на всей Земле, так и отдельно в основных грозовых областях. Зависимость от местного времени также слабее выражена в данных WWLLN. Показано, что в 2007-2009 годах средняя широта обнаружения молний смещается в летнее полушарие на величину до 10 градусов от среднегодового значения. В 2007-2009 годах среднемесячное число вспышек молний в мире по данным WWLLN увеличилось в три раза. Мы объясняем этот факт в первую очередь совершенствованием методов наблюдения в системе WWLLN. Типичная построенная нами карта глобального распределения плотности молний представлена на Рис. 1.1. Построенные карты использованы для численного моделирования Глобальной электрической цепи.

Глобальная электрическая цепь (ГЭЦ) состоит из токов, идущих из грозовых областей атмосферы в ионосферу, растекающихся по ионосфере, стекающих на землю по всему глобусу, и возвращающихся в грозовые области по морям и земной коре. Ионосферная часть ГЭЦ гораздо менее изучена, чем ее атмосферная часть. Мы построили ее новую, более совершенную, математическую модель.

Физико-математический подход к моделированию электрического потенциала, который управляет токами в ионосфере, был описан в предыдущих статьях исполнителей проекта. Использовано двумерное приближение при моделировании ионосферного

проводника. Понижение размерности связано с высокой проводимости в направлении геомагнитного поля. В этом приближении электрический потенциал в E- и F-слоях ионосферы считается постоянными вдоль каждой силовой линии магнитного поля. В математическом отношении задача состоит в решении трех краевых задач для уравнения электропроводности, которые описывают основную часть ионосферы и отдельно – высокоширотные части. Уравнение имеет вид

$$-div(\sigma grad V) = Q,$$

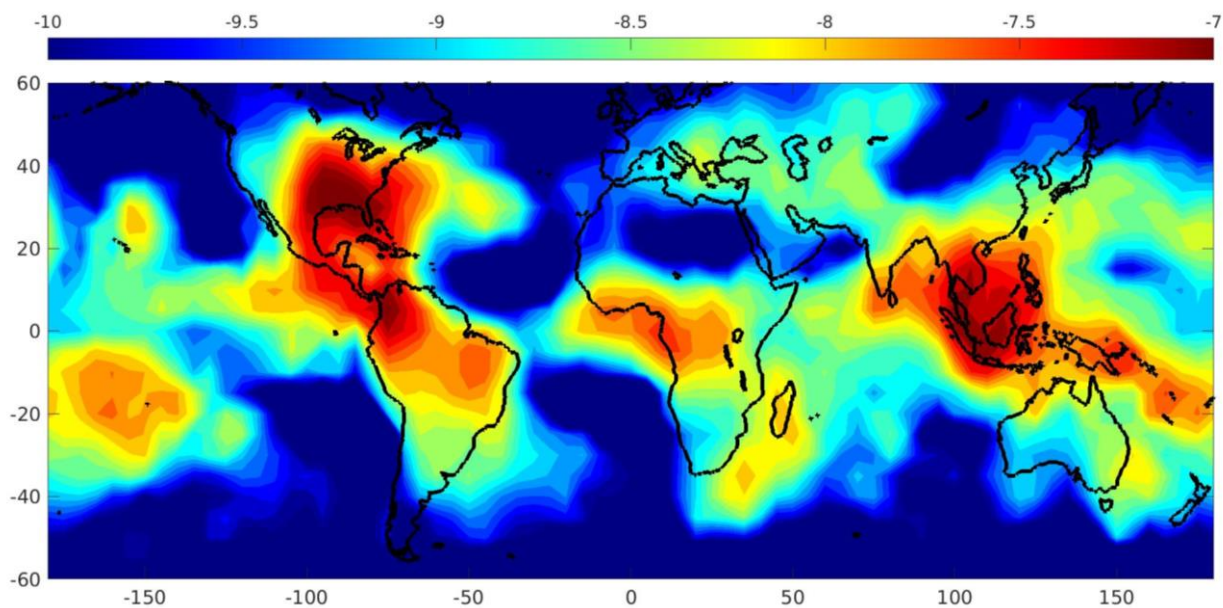


Рисунок 1.1 – Глобальное распределение плотности молний, усредненное за период март-май 2009 по данным WWLLN. Логарифмическая шкала, единица измерения  $1/(\text{км}^2\text{сек})$ .

где  $V$  – искомый электрический потенциал,  $\sigma$  – заданный тензор проводимости,  $Q$  – заданная плотность тока из атмосферы. Компоненты тензора проводимости  $\sigma$  получаются в результате преобразования координат из проводимостей Педерсена и Холла. Глобальные распределения последних рассчитаны с использованием эмпирических моделей ионосферных параметров (IRI, MSIS) и модели геомагнитного поля (IGRF). Правой частью  $Q$  является плотность тока из атмосферы, которая нами получена на основе карты грозовой активности, построенной по данным WWLLN.

Две задачи, для северной и южной полярных шапок, решаются в кругах с краевым условием Дирихле, то есть с потенциалом, равным нулю на границе. Это обусловлено связью авроральных областей ионосферы с магнитосферными объектами, обладающими высокой проводимостью. Задача для основной части ионосферы решается в кольце со смешанными краевыми условиями: Дирихле на внешней границе, и с заданной косо

производной на внутренней границе. Второе условие соответствует заданию тока из атмосферы в ионосферу на геомагнитном экваторе. Оператор каждой из этих трех краевых задач в традиционной постановке является эллиптическим и несимметричным. Мы используем оригинальную формулировку задач для новых неизвестных функций. Операторы становятся симметричными, для них справедлив принцип минимума квадратичного функционала энергии. На этой основе нами ранее создан многосеточный метод численного решения.

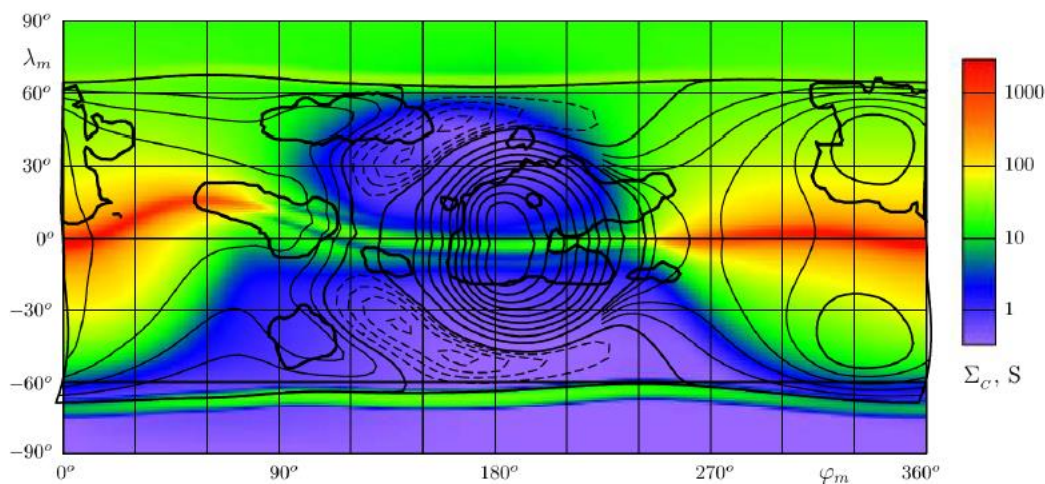


Рисунок 1.2 – Распределение электрического потенциала  $V$  в ионосфере на высоте 120 км в геомагнитных координатах. Эквипотенциалы построены с шагом 5 В, и дополнительно для малых значений  $|V| < 5$  В с шагом 1 В – тонкие линии. Штриховые линии соответствуют отрицательным значениям потенциала. Цветовая шкала показывает распределение проводимости Каулинга. Жирные контуры ограничивают грозовые области, в которых ток течет из атмосферы в ионосферу. Карта построена для 18 часов универсального времени при типичных условиях в июле при низкой солнечной активности.

Основной прогресс по сравнению с предыдущими версиями нашей модели достигнут за счет использования глобального распределения гроз, полученного из наземной Всемирной сети определения местоположения молнии (WWLLN). На Рис. 1.2 показано полученное в результате численного решения задачи электропроводности распределение электрического потенциала  $V$  в ионосфере на высоте 120 км в типичных условиях июля, при низкой солнечной активности в 2008 году, в 18:00 универсального времени. Полученная в модели максимальная разность потенциалов в ионосфере составляет 54 В. Эта новейшая версия нашей модели содержит экваториальные электроструи. Существуют дневные электроструи, сила которых достигает 65 А, и ночные (до 40 А), в то время как общий ток, протекающий в ГЭЦ, в нашей модели задан равным

1.43 кА. Эта сила тока взята для соответствия кривой Карнеги, то есть суточному изменению вертикального электрического поля в атмосфере вблизи земли. В связи с изменением модели грозового генератора в нашей новой модели максимум электрического потенциала сместился из Африки в Юго-Восточную Азию. Экваториальные электроструи также изменили свое положение, направление и интенсивность.

## **1.2 Математическое моделирование диссипативных процессов, вызывающих генерацию электрического поля на границе и в хвосте магнитосферы**

Общепризнано, что процессы магнитного пересоединения играют ключевую роль в динамике магнитосферы. Оно рассматривается как физический процесс быстрого преобразования магнитной энергии в кинетическую и тепловую энергии плазмы. Это явление, связанное с изменением топологии магнитного поля, имеет широкий спектр применений в лабораторной и космической физике плазмы. В частности, оно имеет большое значение для солнечных вспышек и магнитосферных суббурь. Вследствие магнитного пересоединения на границе магнитосферы межпланетное магнитное поле становится топологически связанным с геомагнитным полем. В результате возникает поток энергии из солнечного ветра внутрь магнитосферы. Одним из важных следствий является накопление магнитной энергии в долях магнитосферного хвоста. Магнитные силовые линии из солнечного ветра пересекают границу магнитосферы и проходят внутрь до ионосферы и поверхности Земли. При этом вдоль магнитных силовых линий переносится электрическое поле и втекают электрические токи, создающие магнитные возмущения на уровне ионосферы. Магнитное пересоединение в хвосте магнитосферы обеспечивает преобразование запасенной в долях хвоста магнитной энергии в тепловую и кинетическую энергию частиц плазменного слоя хвоста и является причиной так называемых магнитосферных суббурь.

В космической плазме эффективная проводимость может быть существенно неоднородной. Это происходит из-за влияния плазменной турбулентности, обеспечивающей высокую диссипацию в протяженных и тонких токовых слоях, разделяющих разнонаправленные магнитные поля. В магнитосфере Земли такие слои располагаются в хвостовой части и на границе, обтекаемой солнечным ветром. В этих токовых слоях эффективная проводимость может уменьшаться на несколько порядков в зонах с очень высокой плотностью тока вследствие развития плазменных неустойчивостей. При этом в зоне пониженной проводимости плазмы происходит быстрая

локальная диссипация электрического тока и формируется диффузионная область, пересекаемая магнитными силовыми линиями. В результате генерируется электрическое поле, а плазма ускоряется до альфвеновской скорости. При этом увеличение кинетической энергии плазмы происходит за счет уменьшения энергии магнитного поля. Этот процесс называют быстрым магнитным пересоединением, поскольку в результате разделяемые токовым слоем магнитные поля становятся топологически связанными.

В качестве основной причины неоднородного аномального удельного сопротивления рассматривается ионно-акустическая турбулентность, возникающая в тонких токовых слоях, когда плотность электрического тока превышает критический порог. В магнитосферной плазме электропроводность обычно очень велика, и, следовательно, для эффективного магнитного пересоединения требуется очень малый масштаб длины диффузии. Когда этот масштаб длины становится порядка длины инерции протона, следует учитывать магнитогидродинамические (МГД) эффекты Холла, которые играют важную роль в магнитосферных условиях. В таком случае проблема магнитного пересоединения должна изучаться в рамках Холл-МГД-модели. В нашей работе исследовано влияние неоднородной проводимости и параметра Холла на процесс магнитного пересоединения при различных начальных магнитных конфигурациях.

Нестационарные Холл-МГД уравнения решаются с помощью конечно-разностной численной схемы Годунова на равномерной прямоугольной сетке. Для численного решения нестационарных Холл-МГД уравнений использовалась консервативная конечно-разностная схема с расщеплением, в которой пересчет на следующий временной шаг выполнялся в три этапа. На первом этапе определялись потоки на границах раздела ячеек сетки на основе решения одномерной задачи о распаде разрыва в линеаризованном варианте. На втором этапе пересчитывались все физические величины в ячейках сетки с использованием консервативных уравнений в форме конечных разностей и найденных потоков через границы ячеек без учета диффузии магнитного поля. На третьем этапе включается эллиптический оператор, отвечающий за диффузию магнитного поля, который аппроксимирован неявной разностной схемой с расщеплением по направлениям. В этом случае пересчет на следующий временной слой выполнялся с использованием алгоритма прогонки матрицы.

Процесс пересоединения характеризуется двумя безразмерными параметрами подобия: локальным числом Лундквиста  $S = B L \sigma (\mu_0/\rho)^{1/2}$  в области диффузии и параметром Холла  $\varepsilon = l_p / L$ , где  $B$  – магнитная индукция,  $L$  – пространственный масштаб

неоднородности проводимости,  $\rho$  – аномальная проводимость,  $\sigma$  – массовая плотность плазмы,  $\mu_0$  – магнитная проницаемость вакуума,  $l_p$  – инерционная длина протона.

Нестационарные диссипативные Холл-МГД-уравнения, проинтегрированные численно, дают эволюцию области магнитного пересоединения в зависимости от времени. На Рис. 1.3 показано полученное распределение электрического тока  $J_y$ , соответствующее безразмерному времени  $t = 16$ . Желтым цветом показаны зоны максимальной плотности электрического тока. Фронты интенсивного электрического тока напоминают клешни краба. Они соответствуют медленным ударным фронтам, генерируемым в области магнитной диффузии.

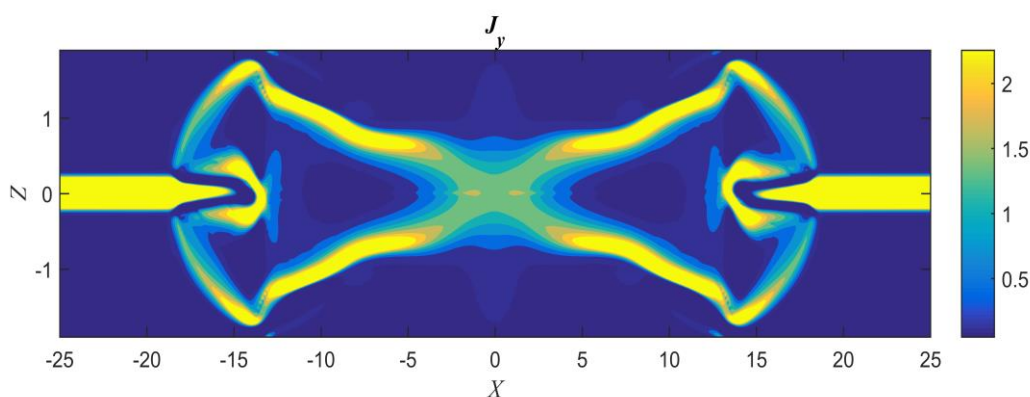


Рисунок 1.3 – Распределение интенсивности электрического тока.

На Рис. 1.4 представлена конфигурация магнитных силовых линий в тот же момент времени, что и на Рис. 1.3. Можно увидеть сжатие силовых линий перед фронтами «диполяризации», распространяющимися вправо и влево вдоль токового слоя. Эти фронты отделяют изначально покоящийся токовый слой от области ускоренной плазмы, пронизанной пересоединившимися магнитными силовыми линиями.

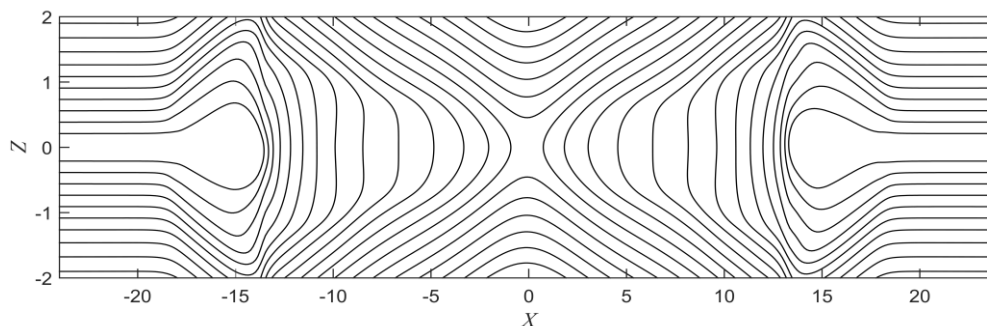


Рисунок 1.4 – Картина магнитных силовых линий, соответствующая распределению тока, показанному на Рис.1.3

Данные фронты аналогичны высокоскоростным потокам, экспериментально наблюдаемым в центральной части токового слоя геомагнитного хвоста. Полученное численное решение указывает на то, что нестационарный процесс магнитного пересоединения асимптотически эволюционирует к стационарному режиму с медленными ударными волнами в случае локально неоднородного удельного сопротивления с выраженным максимумом.

Процесс пересоединения исследован при антипараллельных и скрещенных магнитных полях, как для симметричных, так и несимметричных конфигураций плазменного слоя. Установлено, что генерируемое электрическое поле в области пересоединения пропорционально квадратному корню из обратного локального числа Лундквиста, определенного для максимального удельного сопротивления и его пространственного масштаба. В случае скрещенных магнитных полей электрическое поле пропорционально синусу половинного угла между векторами магнитных полей, разделяемых токовым слоем. Электрическое поле максимально в случае антипараллельных магнитных полей. Такая конфигурация имеет место в плазменном слое магнитосферного хвоста.

На Рис. 1.5 показаны графики изменения электрического поля вдоль оси  $X$  в различные моменты времени. Пики электрического поля соответствуют фронтам «диполяризации», распространяющимся из области диффузии в обоих направлениях. При достаточно большом времени ( $t > 30$ ) фронты «диполяризации» достигают границ и удаляются из области расчета. После этого решение переходит в установившийся режим, соответствующий квазистационарному магнитному пересоединению. При этом максимальная скорость плазмы достигается на расстоянии  $x = 5 L$  от точки максимума аномального сопротивления.

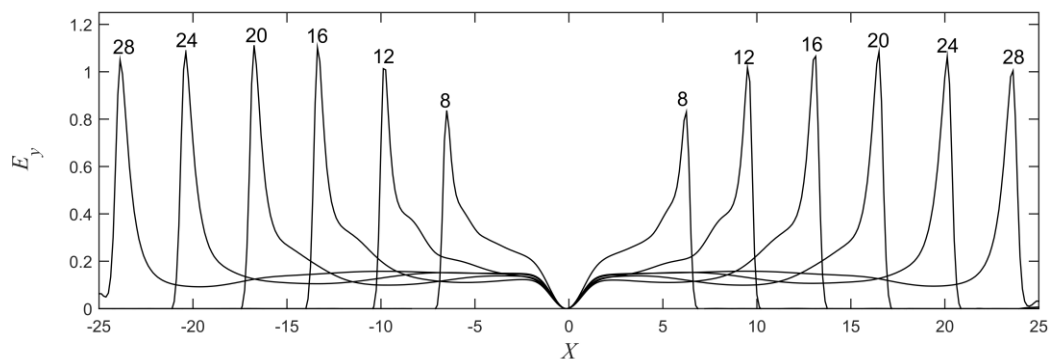


Рисунок 1.5 – Графики изменения электрического поля вдоль оси  $X$  в различные моменты времени



Важно отметить, что решение, полученное в рамках модели пересоединения МГД-Холла, зависит только от двух параметров – числа Лундквиста и параметра Холла. Первый определяет величину генерируемого электрического поля, а второй отвечает за генерацию особых волн Альфвена-Холла, распространяющихся быстрее, чем обычные волны Альфвена в идеальной МГД-модели. Эти волны, возникающие в области диффузии, имеют групповую скорость, превышающую обычную скорость Альфвена, и распространяются далеко от диффузионной области пересоединения, сохраняя свою амплитуду и перенося магнитные возмущения. Групповая скорость волн Холла-Альфвена возрастает с увеличением параметра Холла. На фронте таких волн возникают интенсивные токи вдоль магнитных силовых линий, а также вариации соответствующих холловских компонент магнитного поля и скорости, пропорциональных параметру Холла. Толщина этих волновых фронтов имеет порядок длины диффузионной области ( $L$ ). Волны Холла-Альфвена, распространяющиеся быстрее, чем обычные МГД-волны, могут служить предвестниками импульсного магнитного пересоединения.

### **1.3 Алгоритмы, реализующие длительное движение спутника по возмущенной орбите с выбором методов интегрирования уравнений движения**

В 2022 году разработано несколько приемов конструкции и исследования устойчивости (периодичности) симметричных методов и получаемых на их основе схем предиктор-корректор.

Орбитальное движение описывается системой уравнений 2-го порядка

$$\mathbf{x}'' = \mathbf{f}(t, \mathbf{x}), \quad (1.1)$$

где  $\mathbf{x}(t)$  – позиция спутника на орбите в момент времени  $t$ , а  $\mathbf{f}(t, \mathbf{x})$  – силы, действующие на спутник, причем их расчет требует значительных вычислительных ресурсов. Именно поэтому количество оценок правой части на шаг интегрирования рассматривается как мера вычислительной мощности численного метода интегрирования систем вида (1.1). Линейные многошаговые методы, несмотря на конечный интервал устойчивости, являются конкурентноспособными при длительном интегрировании по сравнению с неявными методами Рунге-Кутты (интегратором Эверхарта и коллокационными), которые широко применяются при интегрировании орбит на небольшом промежутке времени. Для почти круговых орбит (такowymi являются орбиты спутников ГЛОНАСС) особый интерес представляют симметричные линейные многошаговые методы

$$\sum_{j=0}^k \alpha_j \mathbf{x}_{n+j} = h^2 \sum_{j=0}^k \beta_j \mathbf{f}_{n+j}, \quad k \geq 2, \quad \alpha_j = \alpha_{k-j}, \beta_j = \beta_{k-j}, j = 0, \dots, k. \quad (1.2)$$

Здесь на множестве  $\{t_n: t_n = t_0 + nh, h > 0, n = 0, 1, \dots\}$  определена аппроксимация  $\mathbf{x}_n$  точного решения  $\mathbf{x}(t_n)$ ,  $\mathbf{f}_n = \mathbf{f}(t_n, \mathbf{x}_n)$ . Методы (1.2), во-первых, дают только линейный рост ошибки по долготе и почти постоянную ошибку в энергии, что значительно повышает точность расчетов по сравнению с методами семейства Адамса. Во-вторых, методы обратимы во времени для обратимых во времени задач. Более того, многие симметричные методы, примененные к решению периодических задач при разумных ограничениях на шаг интегрирования имеют периодические решения. В этом случае говорят о методах с ненулевым интервалом периодичности.

Симметричный метод может быть только четного порядка  $p$ , причем  $k$ -шаговый явный метод ( $\beta_k = 0$ ) достигает порядка аппроксимации  $p = k$ , а неявный ( $\beta_k \neq 0$ ) – порядка аппроксимации  $p = k + 2$ . Таким образом, нуль-устойчивый неявный симметричный линейный  $k$ -шаговый метод имеет максимальный порядок среди всех линейных методов и является оптимальным в смысле первого барьера Далквиста.

Симметричный метод не определяется однозначно, даже если указаны его порядок и явность/неявность. В литературе описано множество подходов для получения коэффициентов симметричных методов, в том числе высокого порядка. В проекте развит подход, описанный Г. Куинланом и С. Тремано [G. Quinlan, S. Tremaine. Symmetric multistep methods for the numerical integration of planetary orbits // Astron. J., 100(5), 1694–1700 (1990)]. В этом случае метод однозначно определяется указанием его явности/неявности и набором коэффициентов  $\{\alpha_j\}_{j=0}^{k/2}$ , где  $\alpha_j \in \{0, \pm 1, \pm 2, \pm 3, \pm 4\}$ .

Неявные линейные многошаговые методы более точны и имеют больший интервал устойчивости, поэтому более предпочтительны для приложений. Чтобы разрешить неявность, обычно используется режим предиктора-корректора в форме  $P(EC)^m$  или  $P(EC)^m E$ . Здесь  $P$  и  $C$  используются для обозначения одного применения предиктора или корректора соответственно, а  $E$  – для обозначения одного вычисления  $\mathbf{f}(t, \mathbf{x})$ ;  $m$  – количество шагов корректора (т.е. применения пары  $EC$ ). Отметим, что Ламберт [J.D. Lambert, Numerical Methods for Ordinary Differential Systems. The Initial Value Problem. – John Wiley and Sons, 1991] показал, что каждой схеме  $\hat{P}(E\hat{C})^{m+1}$  с предиктором  $\hat{P}$  и корректором  $\hat{C}$  можно сопоставить схему  $P(EC)^m E$  с (другими) предиктором  $P$  и корректором  $C$ , которая имеет такой же главный член локальной ошибки

и такой же полиномом устойчивости, но не наоборот. Поэтому эффективные схемы «предиктор-корректор» следует искать среди схем в форме  $P(EC)^mE$ . Действительно, схему с самым длинным интервалом устойчивости, требующую  $(m + 1)$ -ую оценку правой части за шаг интегрирования, можно найти среди  $P(EC)^mE$ , а соответствующая схема  $\hat{P}(E\hat{C})^{m+1}$  может и не существовать.

В отчетный период разработан ряд алгоритмов получения и исследования устойчивости (периодичности) симметричных методов и получаемых на их основе схем предиктор-корректор. В частности, доказан ряд утверждений о главном члене локальной ошибки схемы  $P(EC)^mE$ ; разработана теория линейной устойчивости режима предиктор-корректор.

Доказано, что если предиктор и корректор имеют один порядок, то 1) главные члены локальной ошибки схемы  $P(EC)^mE$  и корректора совпадают; 2) с помощью механизма Милна [Milne W.E., Numerical solution of differential equations. – John Wiley and Sons, 1953] можно апостериорно оценивать локальную ошибку, что удобно для практических расчетов.

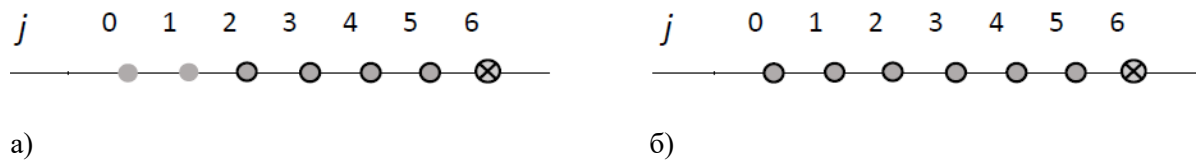


Рисунок 1.6 – Примеры шаблонов схемы  $P(EC)^mE$ : а) предиктор и корректор 6-го порядка (симметричность нарушается); б) предиктор 6-го порядка и корректор 8-го порядка (симметричность сохраняется). Точки, соответствующие предиктору, закрашены серым, корректору – имеют черную границу, а помеченная крестиком точка оценивается на текущем шаге

Поскольку в этом случае шаблоны предиктора и корректора не совпадают (рис. 1.6.а), схема  $P(EC)^mE$  не имеет интервала периодичности и теряет часть преимуществ симметричных методов. Был доказан ряд утверждений о корнях полинома устойчивости схем  $P(EC)^mE$  и показано, что для некоторых  $m$  не существует устойчивых схем (табл. 1.1).

Чтобы схема  $P(EC)^mE$  сохранила преимущества симметричных методов, шаблоны предиктора и корректора должны совпадать друг с другом (рис. 1.6.б). Доказано, что в этом случае 1) порядок предиктора на две единицы меньше порядка корректора; 2) порядок схемы  $PECE$  и корректора совпадают, а главные члены локальной ошибки нет; 3)

порядок и главные члены локальной ошибки схемы  $P(EC)^mE$  и корректора совпадают  $\forall m \geq 2$ ; 4) если предиктор и корректор имеют ненулевые интервалы периодичности, то схема  $P(EC)^mE$  тоже имеет ненулевой интервал периодичности.

Таблица 1.1. Существование устойчивых схем  $P(EC)^mE$  при совпадении порядка предиктора и корректора.

Порядок	$PECE$	$P(EC)^2E$	$P(EC)^3E$	$P(EC)^4E$
6	+	-	+	-
8	-	+	-	+
10	+	-	+	-
12	-	+	-	+

Нами предложена технология выбора пар предиктор-корректор из широкого семейства симметричных методов высокого порядка. Технология учитывает как длину интервала периодичности, так и точность численного решения. В результате работы построены, исследованы и описаны предикторы и корректоры схем  $P(EC)^mE$  от 6-го до 14-го порядка точности, лучшие по точности среди имеющих наибольшие интервалы периодичности как в случаях совпадения порядков предиктора и корректора, так и различия их порядков на 2.

Ранее участники проекта показали, что симметричные методы превосходят методы семейства Адамса по точности. В этом году сравнили вычислительные затраты режима предиктор-корректор для выбранных пар симметричных методов высокого порядка и интегратора Эверхарта того же порядка. Численные эксперименты проведены для задачи возмущенного движения спутника по орбите. Модельная задача построена из ограниченной задачи трех тел (Земля, Луна и спутник Земли пренебрежимо малой массы) таким образом, что имеется дополнительная сила, действующая на спутник, зависящая только от времени и описывающая влияние Луны, движущейся вокруг Земли. В результате модельная задача имеет точное решение, и можно явно вычислить ошибки численного решения. Поскольку матрица Якоби этой задачи совпадает с матрицей Якоби ограниченной задачи трех тел, то и свойства устойчивости численных методов для этих задач совпадают и отличаются от свойств традиционной задачи Кеплера для невозмущенной орбиты вокруг Земли. Численные эксперименты проведены на реальных данных, а кеплеровы элементы соответствуют положению спутника COSMOS 2413 (ГЛОНАСС) в один из моментов 12 июля 2015 г.

Таблица 1.2. Сравнение вычислительной стоимости построенных методов ( $P_{p_1}C_{p_2}$ ) и интегратора Эверхарда ( $Ev-p$ ), где  $p_1$ ,  $p_2$  и  $p$  – порядки предиктора, корректора и интегратора Эверхарта.

Ref.	$P_4C_6$	$Ev-6$	$P_6C_8$	$Ev-8$	$P_8C_{10}$	$Ev-10$	$P_{10}C_{12}$	$Ev-12$	$P_{12}C_{14}$	$Ev-14$
$N$	1718	487	445	86	192	29	109	14	73	9
$h$ (с)	23,60	83,25	91,11	471,44	211,17	1398,07	371,96	895,99	555,40	4504,88
$R1$		3,53		5,17		6,62		7,79		8,11
$E$	2676642	4931849	693306	1406874	299130	700321	169814	468958	113724	399627
$R2$		1,84		2,03		2,34		2,76		3,51
$\rho^h$	2,0E-03	2,0E-03	2,0E-03	1,8E-03	1,9E-03	1,5E-03	1,9E-03	8,6E-03	1,7E-03	3,1E-03

В экспериментах определялся шаг, при котором максимальное отклонение  $\rho^h$  численного решения от точного не превышает 2 мм при движении спутника в течение примерно одного года (780 периодов). Результаты численных экспериментов представлены в табл. 1.2. Количество шагов за период и размер шага в секундах обозначаются  $N$  и  $h$ . Строка, помеченная « $E$ », показывает количество вычислений правых частей, необходимых для достижения точности, и характеризует вычислительную стоимость методов. Отношения  $R1$  и  $R2$  показывают, во сколько раз размер шага и количество вычислений правой части интегратора Эверхарта больше соответствующих параметров схемы  $PECE$  того же порядка. Из табл. 1.2 видно, что хотя размер шага, при котором достигается  $\rho^h \approx 2$  мм, для метода Эверхарда больше по сравнению с режимом  $PECE$ , но этот метод более вычислительно трудоемкий по сравнению с предложенными.

## **2 Математическое моделирование физических процессов и устройств**

### **2.1 Условия максимального усиления локального поля за счет настройки поверхностного решеточного резонанса путем варьирования геометрических параметров структуры**

Периодические структуры из резонансных частиц привлекают внимание в связи с проявлением в них нетривиальных оптических свойств и возможностью широкого применения в задачах усиления локального поля, сенсорики, генерации лазерного излучения и других прикладных областях. В таких системах возбуждение электромагнитных мод различных порядков возможно за счет взаимодействия мод отдельных частиц структуры с коллективными модами, обусловленными взаимодействием частиц друг с другом. Для сферических частиц модами первого порядка являются электро- и магнитодипольные. За счет увеличения размеров частиц возможно достичь возбуждения мод более высоких порядков. В случае периодических структур (например, решеток наночастиц) взаимодействие локальных мод каждой отдельной частицы с коллективными резонансами также приводит к возбуждению мод различных порядков. При этом модовый состав таких резонансов зависит от ряда геометрических параметров как самих частиц, так и структуры в целом. Особый интерес представляет решение задачи поиска геометрии системы, которая обеспечивает максимальное усиление локального поля. В настоящем разделе проекта нами исследовались периодические структуры из альтернативного плазмонного материала – алюминия. Этот материал обладает уникальными оптическими свойствами и позволяет достичь резонансный отклик в ближней ультрафиолетовой области спектра, что крайне актуально для биомедицинских приложений.

Для решения задач рассеяния периодическими структурами наночастиц использовался метод конечных разностей во временной области (FDTD). Он является одним из ключевых методов для описания оптических свойств наноструктур и использовался для расчета спектров экстинкции, отражения и поглощения в наноструктурах. Этот метод разрабатывался с середины XX века и за это время стал одним из самых популярных методов численной электродинамики. Метод основан на дискретизации уравнений Максвелла, записанных в дифференциальной форме. В отличие от других методов, например, метода связанных диполей, математическая модель, лежащая в основе FDTD, не использует никаких приближений и является точной для классической электродинамики.

Уравнения решаются методом конечных разностей на двух вложенных структурированных прямоугольных сетках, одна из которых используется для расчета

электрических полей, а другая – для магнитных полей. Метод позволяет за один расчет получить спектры пропускания и отражения исследуемой системы, а также распределение поля в расчетной области с учетом как сложной композитной структуры исследуемой системы, так и оптических свойств используемых материалов.

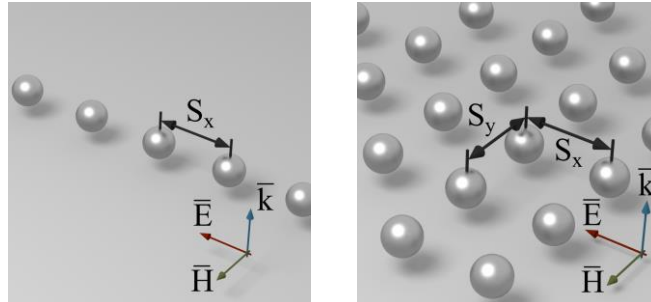


Рисунок 2.1 – Схематичное изображение исследуемых структур: одномерной и двумерной.  $S_x$  и  $S_y$  – периоды структуры в соответствующем направлении,  $\mathbf{E}$ ,  $\mathbf{H}$  и  $\mathbf{k}$  – векторы электрического, магнитного поля и волновой вектор падающей волны

Возможность применения периодических граничных условий, а также оптимизация за счет учета симметрии элементарной ячейки периодической структуры позволяют значительно ускорить процесс расчета. Кроме того, метод позволяет проводить параллельные вычисления с использованием высокопроизводительных кластерных систем. На Рис. 2.1 представлен общий вид исследуемой структуры. Моделирование осуществлялось для элементарной ячейки с периодическими граничными условиями в направлениях  $Y$  и  $Z$ , и PML-слоями в направлении  $X$  для имитации бесконечного пространства. В наших расчетах периодическая структура освещалась сверху плоской волной с нормальным падением вдоль оси  $X$  и поляризацией вдоль оси  $Y$ . Структура электромагнитного поля записывалась с помощью трехмерных DFT-мониторов, полностью вмещающих в себя исследуемые частицы. Данные с пространственным распределением комплексных векторов электрического и магнитного поля записывались в файлы в формате VTK для последующей обработки и разложения поля по мультиполям.

Дальнее поле, рассеянное произвольным объектом и полученное численным методом (в нашем случае методом FDTD, как описано выше), может быть разложено на дипольную и квадрупольную компоненты.

В нашем конкретном случае, исходя из общих соображений, получим

$$\mathbf{E}_{sc} = -\frac{\omega^2}{c^2 r_0} \left( [[\mathbf{d} \times \mathbf{n}] \times \mathbf{n}] - [\mathbf{m} \times \mathbf{n}] - \frac{i\omega}{6c} ([[ \mathbf{D} \times \mathbf{n}] \times \mathbf{n}] - [\mathbf{M} \times \mathbf{n}]) \right) e^{ikr_0}$$

Здесь  $r_0$  – расстояние до точки наблюдения, а  $\mathbf{n}$  – единичный вектор, направленный к точке наблюдения;  $\omega$  – это частота,  $k = \omega/c$  – волновой вектор,  $\mathbf{d}$  и  $\mathbf{m}$  – соответственно ED и MD. Декартовы компоненты EQ и MQ имеют вид  $\mathbf{D} = \hat{D}\mathbf{n}$ ,  $\mathbf{M} = \hat{M}\mathbf{n}$ . Электрический и магнитный дипольный и квадрупольный моменты, возбуждаемые на частицах, определяются как

$$\begin{aligned}\mathbf{d} &= \iiint_{V_p} \left( \mathbf{P} j_0(kr) + \frac{k^2}{2(kr)^2} (3(\mathbf{P} \cdot \mathbf{r})\mathbf{r} - r^2\mathbf{P}) j_2(kr) \right) d^3\mathbf{r}, \\ \mathbf{m} &= i\frac{k}{2} \iiint_{V_p} [\mathbf{P} \times \mathbf{r}] \frac{j_1(kr)}{kr} d^3\mathbf{r}, \\ \hat{\mathbf{D}} &= \iiint_{V_p} \left( (3(\mathbf{P} \otimes \mathbf{r} + \mathbf{r} \otimes \mathbf{P}) - 2(\mathbf{P} \cdot \mathbf{r})\hat{\mathbf{I}}) \frac{j_1(kr)}{kr} + \right. \\ & \left. 2k^2 (5\mathbf{r} \otimes \mathbf{r}(\mathbf{P} \cdot \mathbf{r}) - (\mathbf{P} \otimes \mathbf{r} + \mathbf{r} \otimes \mathbf{P})r^2 - (\mathbf{P} \cdot \mathbf{r})\hat{\mathbf{I}}) \right) \frac{j_3(kr)}{(kr)^3} d^3\mathbf{r}, \\ \hat{\mathbf{M}} &= \frac{1}{3}ik \iiint_{V_p} (\mathbf{r} \otimes [\mathbf{P} \times \mathbf{r}] + [\mathbf{P} \times \mathbf{r}] \otimes \mathbf{r}) \frac{j_2(kr)}{(kr)^2} d^3\mathbf{r},\end{aligned}$$

где  $\varepsilon$  – проницаемость материала наночастиц,

$\mathbf{P} = (\varepsilon - 1)/4\pi\mathbf{E}$  – поляризация,

$\mathbf{J} = -i\omega\mathbf{P}$  – плотность тока,

$j_n$  – сферические функции Бесселя первого рода,

$\hat{\mathbf{I}}$  – единичный тензор,

$r = |\mathbf{r}|$ ,  $V_p$  – предел интегрирования по всем частицам.

На Рис. 2.2 представлены спектры экстинкции одномерной и двумерной структуры, а также его компоненты ED, MD, EQ, MQ. Период структур  $S = 280$  нм, радиус частиц 60 нм. Как видно из представленных зависимостей, амплитуда экстинкции в случае одномерной структуры значительно меньше, чем в случае двумерной. Кроме того, в спектре двумерной структуры присутствуют две обособленные линии экстинкции, которые обусловлены возбуждением магнитодипольного и электрквадрупольного резонанса в случае коротковолнового пика, а в случае длинноволнового пика – электродипольного и магнитоквадрупольного резонансов. Интересно отметить, что положение каждого из максимумов экстинкции связано со спектральным положением аномалий Рэлея для взаимно ортогональных направлений в плоскости решетки.



Коротковолновый пик связан с периодом решетки в направлении, лежащем в плоскости поляризации падающего излучения, а длинноволновый пик – в перпендикулярном направлении. Изменяя шаг решетки в одном из направлений, можно добиться совпадения длин волн всех четырех мультипольных компонент.

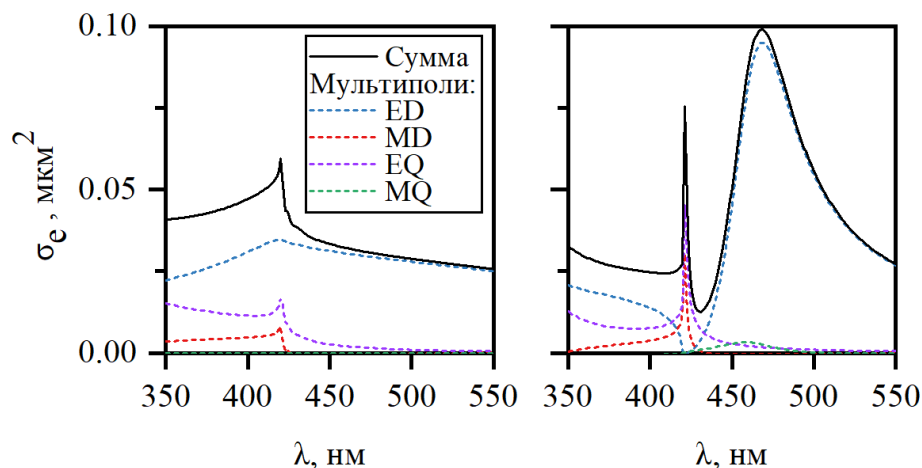


Рисунок 2.2 – Спектры экстинкции одномерной (слева) и двумерной (справа) периодической структуры из частиц алюминия, а также их разложение на компоненты по мультиполям.

Период структур 280 нм, радиус частиц 60 нм.

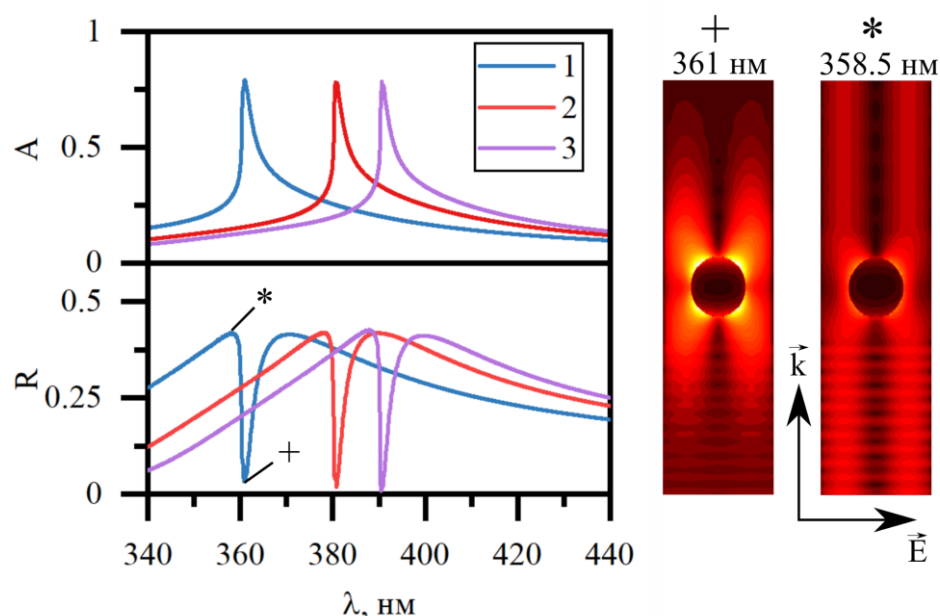


Рисунок 2.3 – Спектры поглощения (вверху), отражения (внизу) двумерной периодической структуры из алюминиевых частиц с различными геометрическими параметрами 1:  $R_p=50$  нм  $S_x=240$  нм,  $S_y=197$  нм; 2:  $R_p=53.3$  нм,  $S_x=253.3$  нм,  $S_y=211$  нм; 3:  $R_p=55$  нм,  $S_x=260$  нм,  $S_y=217.5$  нм, где  $R_p$  – радиус сферической наночастицы,  $S_x$  – шаг решетки в направлении  $\mathbf{E}$ ,  $S_y$  – шаг решетки в направлении  $\mathbf{H}$ . Распределение электрического поля (справа) в элементарной ячейке на длине волны максимума и минимума отражения.

Как показано на Рис. 2.3, в случае совпадения спектрального положения всех четырех компонент в разложении электромагнитного поля по мультиполям, в отражении наблюдается минимум, а в поглощении максимум. Кроме того, как видно из распределения локального электрического поля, такая синхронизация приводит к максимальному усилению поля вблизи частицы.

## 2.2 Математические модели ионного транспорта в наноструктурированных мембранных материалах под действием электрического поля

По прогнозам в ближайшие десятилетия под действием изменения климата и из-за ухудшения экологической ситуации кризис водных ресурсов превратится в одну из ведущих угроз для жизни миллиардов людей. Возможным вариантом решения этой проблемы может стать опреснение соленой воды и очистка сточных вод. Наиболее популярной и удобной технологией опреснения воды является мембранная технология обратного осмоса, в которой вода при определенном давлении проходит через полупроницаемую мембрану из более концентрированного в менее концентрированный раствор (мембрана пропускает воду и задерживает растворенные вещества). Обратный осмос используется для обессоливания морской воды, получения питьевой, а также особо чистой воды. Основу установок обратного осмоса составляют тонкопленочные композитные мембраны, очищающие воду от солей. На селективные свойства таких фильтров помимо технологических характеристик также влияют pH и состав разделяемого раствора, однако сложный характер этого воздействия еще недостаточно изучен.

В рамках сотрудничества с коллегами из Европейского центра передовых технологий водоподготовки (Wetsus) и Университета Твенте (Нидерланды) была разработана модель транспорта ионов через полимерную мембрану обратного осмоса. Такие мембраны являются композитными и состоят из подложки из полиэстера (150 мкм), промежуточного слоя из полисульфона (40 мкм) и селективного слоя из полиамида (200 нм). В структуре полиамида присутствуют карбоксильные группы и аминогруппы, которые диссоциируют согласно уравнениям



При низких pH преобладает реакция (2.1), в результате чего мембрана заряжается положительно, а при высоких pH – реакция (2.2), которая приводит к отрицательному

заряду мембраны. Электрическое поле, создаваемое поверхностными зарядами мембраны, существенно влияет на транспорт ионов через поры, характерный размер которых составляет порядка 1 нм, см. Рис. 2.4.

Для описания влияния рН на перенос ионов через селективный слой и их задержание была разработана математическая модель, которая основывается на следующих предположениях.

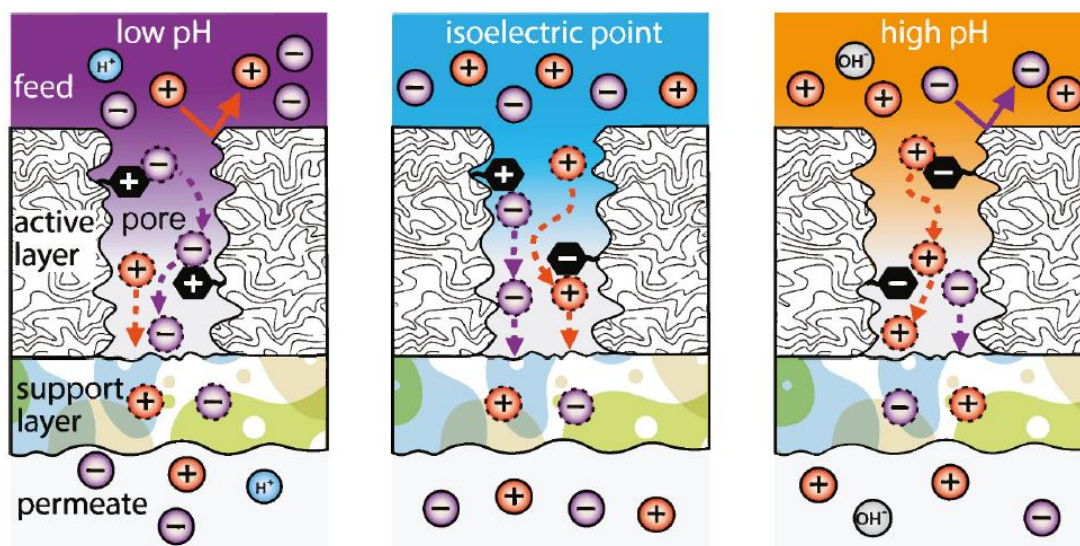


Рисунок 2.4 – Влияние рН на транспорт ионов через селективный слой обратноосмотической мембраны.

1. Концентрации ионов и электрический потенциал изменяются только в направлении  $Z$ , перпендикулярном поверхности мембраны.

2. Перенос катиона, аниона,  $H^+$  и  $OH^-$  ионов происходит за счет конвекции, диффузии и электромиграции и описывается уравнением Нернста-Планка:

$$J_i = KC_i J_V - K\varepsilon D_i \left( \frac{dC_i}{dZ} + \frac{F}{RT} z_i C_i \frac{d\Phi}{dZ} \right), \quad i = 1, 2, 3, 4. \quad (2.3)$$

Здесь  $J_i$  – поток ионов типа  $i$ ,  $C_i$  – молярная концентрация ионов,  $z_i$  – зарядовое число ионов,  $D_i$  – коэффициент диффузии ионов,  $J_V$  – объемный поток (скорость) воды,  $K$  – коэффициент трения между ионами и мембраной,  $\varepsilon$  – коэффициент пористости мембраны,  $\Phi$  – электрический потенциал,  $T$  – температура,  $F$  – постоянная Фарадея,  $R$  – универсальная газовая постоянная.

3. Учитывается диссоциация воды, при этом ионное произведение воды является постоянным в соответствии с уравнением

$$C_{\text{H}^+} C_{\text{OH}^-} = K_w. \quad (2.4)$$

4. Заряд поверхности мембраны определяется диссоциацией карбоксильных групп и аминогрупп, в результате чего плотность заряда  $X$  (приведенная к мол / м<sup>3</sup>) дается формулой

$$X = \xi \left( \frac{C_{\text{NH}_2}^t}{1 + K_{\text{NH}_2} / C_{\text{H}^+}} - \frac{C_{\text{COOH}}^{1t}}{1 + C_{\text{H}^+} / K_{\text{COOH}}^1} - \frac{C_{\text{COOH}}^{2t}}{1 + C_{\text{H}^+} / K_{\text{COOH}}^2} \right) \quad (2.5)$$

Здесь  $C_{\text{H}^+}$  – концентрация ионов водорода,  $K_j$  – константы равновесия реакций (2.1), (2.2) (при этом вторая реакция характеризуется двумя константами),  $C_j^t$  – плотность поверхностных групп,  $\xi$  – коэффициент, учитывающий долю активных поверхностных групп.

5. Выполняется условие электронейтральности

$$z_+ C_+ + z_- C_- + C_{\text{H}^+} - C_{\text{OH}^-} + X = 0. \quad (2.6)$$

6. Ионный ток отсутствует:

$$z_+ J_+ + z_- J_- + J_{\text{H}^+} - J_{\text{OH}^-} = 0. \quad (2.7)$$

7. Выполняются условия Доннана на границе мембрана/раствор:

$$C_i = C_i^\infty \phi_i \exp(-z_i \Delta\Phi). \quad (2.8)$$

Здесь  $C_i$  – концентрация ионов внутри мембраны,  $C_i^\infty$  – концентрация ионов в растворе,  $\phi_i$  – стерический коэффициент, описывающий соотношения размера пор мембраны и размера ионов,  $\Delta\Phi$  – скачок потенциала на границе.

8. Потоки катиона и аниона определяются формулами

$$J_\pm = C^P J_V, \quad (2.9)$$

где  $C^P$  – концентрация соли в выходном потоке (пермеате).

Предложенная модель сводится к системе обыкновенных дифференциальных уравнений для концентраций ионов и электрического потенциала с краевыми условиями, которая решалась численно. В модели задается объемный расход (скорость течения) и концентрации ионов в сырьевом потоке  $C_i^f$ . В результате решения необходимо определить концентрацию ионов в пермеате  $C_i^p$  и задержание ионов согласно формуле

$$R_i = 1 - C_i^p / C_i^f. \quad (2.10)$$

Расчетные данные сравнивались с данными экспериментов, которые проводились в Университете Твенте. Использовалась установка тангенциальной фильтрации (Convergence Inspector Colossus, The Netherlands) и мембрана BW30HR-440 (DuPont™ FilmTec™, США), вырезанная из мембранного модуля. В качестве рабочего был выбран водный раствор NaCl (30 мМ); поток через мембрану составлял 20 л/м<sup>2</sup> час, давление 6.2 – 6.8 бар; сырьевой поток 50 л/час; поток пермеата 0.5 л/час. Эксперименты проводились для различных значений pH в диапазоне pH = 3.5–10.5 (посредством добавления кислоты / основания, а также прокачкой смеси газов CO<sub>2</sub> / N<sub>2</sub> в определенном соотношении). Для анализа состава пермеата использовался ионный хроматограф Metrohm 930 со встроенным датчиком проводимости для определения концентраций ионов.

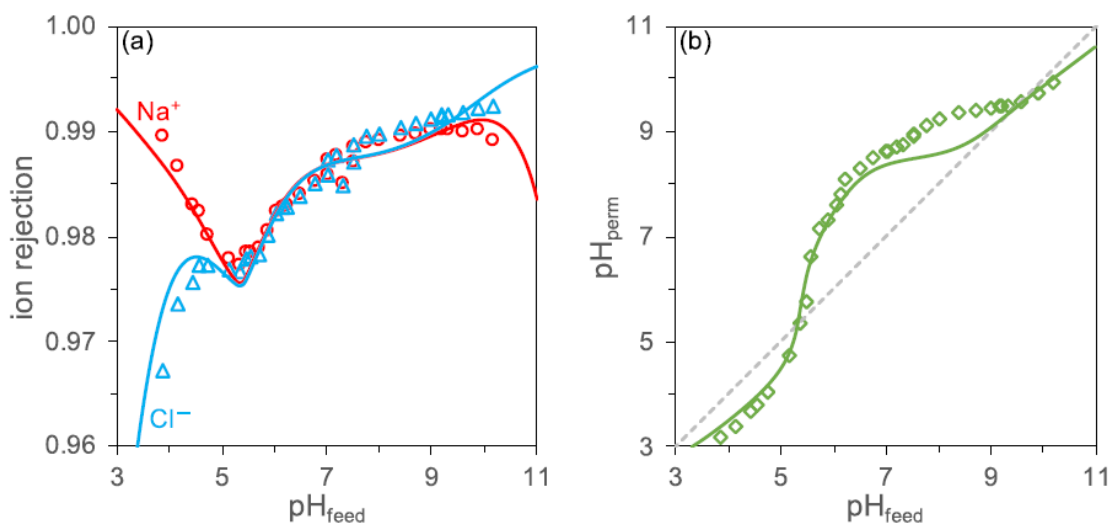


Рисунок 2.5 – Влияние pH сырьевого потока на задержание ионов (a) и pH пермеата (b). Экспериментальные данные показаны точками, расчетные – непрерывными кривыми.

Результаты расчетных и экспериментальных исследований показаны на Рис. 2.5 для задержания ионов натрия и хлора, а также pH пермеата в зависимости от pH сырьевого

потока. В работе была осуществлена подгонка параметров модели (коэффициентов трения, пористости и заряда) для описания экспериментальных данных. На Рис. 2.6 представлена расчетная зависимость заряда поверхности мембраны от рН сырьевого потока. Видно, что при  $\text{pH} = 5.2$  мембрана является незаряженной, что приводит к локальному минимуму задержания соли (Рис. 2.5.а). При  $\text{pH} < 5.2$  заряд мембраны является положительным за счет диссоциации аминогрупп (реакция (2.1)). В этом случае задержание ионов натрия больше, чем задержание ионов хлора (рис. 2.5.а) за счет электростатического взаимодействия. При  $\text{pH} > 5.2$  заряд мембраны является отрицательным за счет диссоциации карбоксильных групп (реакция (2.2)) и его величина растет с ростом рН. Это приводит к увеличению задержания соли. Из Рис. 2.5.б видно, что фильтрация соли приводит к увеличению рН пермеата по сравнению с рН сырьевого потока.

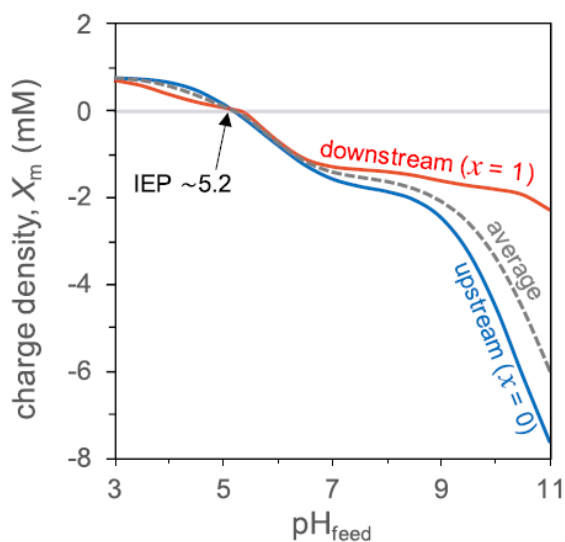


Рисунок 2.6 – Влияние рН сырьевого потока на заряд мембраны.

Результаты работы показывают ключевую роль локального рН в ионизации функциональных групп и их взаимодействия с ионами  $\text{H}^+$  и  $\text{OH}^-$ . Установлено, что степень очистки воды зависит от ее водородного показателя рН и ионного заряда мембраны. Модель позволяет предсказывать эффективность опреснения, минеральный состав полученной воды и может использоваться для разработки улучшенных фильтров и мембранных элементов.

### **3 Исследование больших данных, обнаружение новых знаний и прогнозирование состояния сложных систем и объектов**

#### **3.1 Популяционно-эволюционный анализ геномов вирусов семейства Coronaviridae**

SARS-CoV-2 является третьим коронавирусом в дополнение к SARS-CoV и MERS-CoV, который вызывает тяжелый респираторный синдром у людей. Все они, вероятно, пересекли межвидовой барьер между животными и человеком и соответственно имеют зоонозное происхождение. Происхождение и эволюция вирусов и их филогенетические взаимоотношения имеют большое значение для изучения их патогенности и разработки противовирусных препаратов и вакцин. Основная цель исследования состояла в сравнении двух методов выявления взаимосвязей между геномами коронавирусов: филогенетического, основанного на выравнивании всего генома с последующим выводом молекулярного филогенетического дерева, и кластеризации частот триплетов без выравнивания, соответственно, с использованием 69 геномов коронавирусов, выбранных из двух общедоступных баз данных. Оба подхода привели к хорошо разделяемым надежным классификациям. В целом кластеры, выделенные первым подходом, хорошо согласовывались с классами, выделенными вторым подходом методами K-средних и упругих карт. Оба подхода продемонстрировали также значительное расхождение геномов на таксономическом уровне, но меньшее соответствие между геномами по типам вызываемых ими заболеваний, что может быть связано с индивидуальными особенностями организмов.

Было отобрано 69 геномных последовательностей коронавирусов (семейство Coronaviridae, подсемейство Orthocoronavirinae), представляющих все четыре рода: 19 альфакоронавирусов, в том числе два (NL63 и 229E), поражающих человека и восемь, поражающих летучих мышей, 24 бетакоронавируса, в том числе пять (OC43, HKU1, SARS-CoV, MERS-CoV и SARS-CoV-2), поражающие людей, два поражающих панголинов и три поражающих летучих мышей, 6 гаммакоронавирусов, 10 дельтакоронавирусов и 10 неклассифицированных. Большинство последовательностей были взяты из NCBI GenBank. Три последовательности генома коронавирусов SARS-CoV-2, появившихся в июне 2020 года в Пекине, Китай, и последовательности вируса RaTG13 были взяты из базы данных GISAID.

Для первого подхода выравнивание последовательностей проводили с использованием MAFFT (версия 7) и универсального итеративного алгоритма уточнения L-INS-i, подходящего для последовательностей, насыщенных расширенными вставками.

Поиск оптимальной модели молекулярной эволюции и филогенетических построений осуществлялся с использованием программы IQ-TREE версии 1.6.12 и версии модели молекулярной эволюции GTR, в которой некоторые сайты инвариантны, а остальные следуют гамма-распределению. Надежность топологии дерева оценивалась с помощью метода приближенной оценки LRT IQ-TREE — aLRT (приближенный тест отношения правдоподобия), основанного на идее LRT с гипотезой о нулевой длине предполагаемой ветви. Анализ и визуализация результатов проводились с использованием оригинальных программ на языках программирования R и Python.

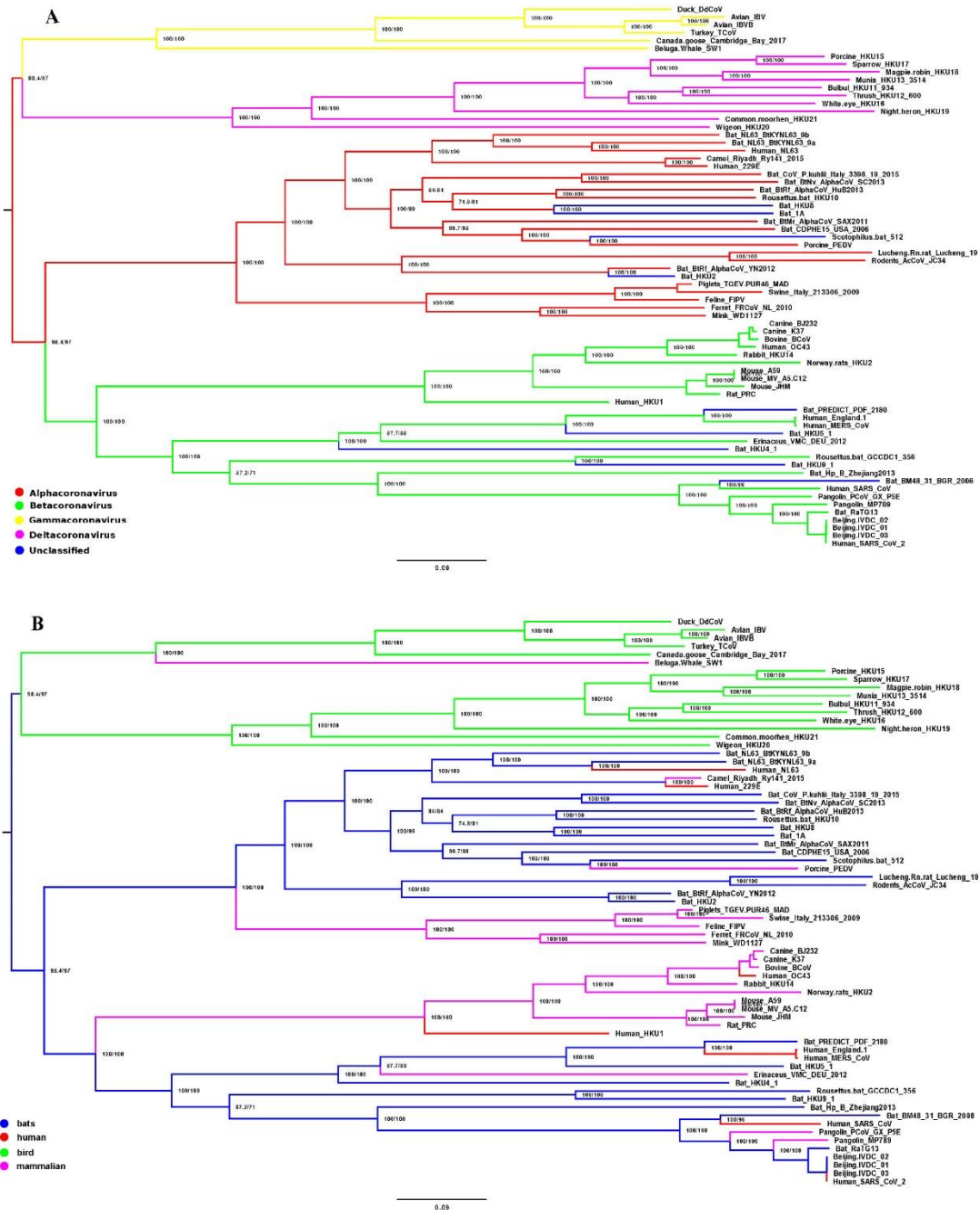


Рисунок 3.1 – Филогенетическое дерево, основанное на выравнивании 69 полногеномных последовательностей коронавируса.



Предполагаемые филогении были визуализированы с помощью программы FigTree версии 1.4.4 и Splitstree версии 5.0.20. На полученном филогенетическом дереве идентичности геномов были выделены разными цветами в соответствии с таксономической принадлежностью к одному из четырех родов коронавирусов (Рис. 3.1.А) и типом вызываемого ими заболевания (Рис. 3.1.Б). Таксономическую принадлежность и тип заболевания определяли по аннотации в NCBI GenBank с учетом доступных публикаций и первичных описаний.

Во втором подходе каждый геном рассматривался как последовательность из четырех букв алфавита  $\mathcal{K} = \{A, C, G, T\}$ ;  $L$  – длина последовательности. Словарь частот триплетов представляет собой список всех возможных триплетов  $\omega = v_1v_2v_3$ , от  $\omega = AAA$  до  $\omega = TTT$ . Количество каждого триплета в последовательности вычислялось путем перемещения рамки считывания с окном в три нуклеотида и шагом в три нуклеотида от первого до последнего нуклеотида в последовательности генома. Всего возможно 64 триплета. Частота  $f_\omega$  – это отношение числа копий  $n_\omega$  данного триплета к общему числу всех триплетов  $N$ , где  $N$  – сумма всех  $n_\omega$ :  $f_\omega = n_\omega / N$ . Таким образом, каждый геном можно представить как точку в 64-мерном (64D) пространстве.

Два генома считаются настолько близкими друг к другу, насколько мало евклидово расстояние между ними в 64D-пространстве. Это расстояние можно использовать для сравнения с другими мерами расстояния, такими как наиболее вероятное мутационное расстояние между двумя геномами или патристическое расстояние.

Метод  $k$ -средних — простейший и наиболее широко используемый алгоритм итерационной классификации (кластеризации), который делит набор данных на  $k$  классов, расположенных на определенном расстоянии друг от друга. Алгоритм метода принципиально делит множество точек, представленных в  $n$ -мерном пространстве (в нашем случае 63-мерном, поскольку сумма всех частот составляет единицу) на  $k$  классов, причем каждая точка изначально попадает в случайно определенный класс. Далее для каждого класса вычисляется среднее арифметическое, и точки проверяются на близость к классу с помощью евклидова расстояния. Тогда, если точка, изначально принадлежащая классу  $j$ , становится близкой к центру (среднему арифметическому) класса  $m$ , то точка меняет принадлежность к классу. Такая перестановка повторяется для всех точек. Как только новое распределение получено, центры пересчитываются, и процедура выполняется снова. Алгоритм повторяется до тех пор, пока не прекратится переход точек между классами. Для реализации метода  $k$ -средних использовалось свободно

распространяемое программное обеспечение ViDaExpert. В работе было протестировано различное количество кластеров, перебирались значения  $k$  от 2 до 11, но поиск максимально возможного оптимального количества кластеров целью не ставился. Ключевым вопросом исследования была связь между составами классов, наблюдаемых для различных значений  $k$ .

Для нелинейного сокращения размерности и визуализации многомерных данных использовались упругие карты. Суть метода состоит в следующем. На первом шаге вычисляются первая и вторая главные компоненты и по ним, как по осям, строится плоскость. Затем все точки данных проецируются на эту плоскость и определяется минимальный квадрат, содержащий все точки. Квадрат делится на определенное количество меньших квадратов (в нашем исследовании используются 16 и 25). На втором этапе каждая точка данных соединяется математической пружиной с ближайшим к проекции узлом решетки. Тогда жесткая плоскость (точнее, та ее часть, которая соответствует большему квадрату) заменяется эластичной мембраной с равномерной упругостью. На третьем этапе вся система освобождается, а пружины сокращаются (или растягиваются под воздействием мембраны и соседних пружин), так что мембрана деформируется. В этом случае деформация мембраны и ее конечное состояние определяются минимумом полной энергии деформации. На четвертом шаге переопределяется положение каждой точки на деформированной карте и находится новая ортогональная проекция (точка на деформированной карте, ближайшая к исходной). Наконец, все математические пружины убираются, и деформированная мембрана возвращается в исходное плоское состояние; проекции точек также меняют свое положение на эластичной карте. Это преобразование называется переходом во внутренние координаты. Определение кластера на упругой карте, представленной во внутренних координатах, вычислялось по локальной плотности. Для этого каждой точке упругой карты была поставлена в соответствие функция Гаусса. Затем значения функции для каждой точки суммировались по всем точкам, и значение суммарной функции определяло локальную плотность. Для отображения использовалась схема с 15 уровнями значения функции локальной плотности. Кластером считалась область с локальной плотностью, превышающей 9-й уровень.

Попарные патристические расстояния между всеми геномами, участвующими в анализе, рассчитывали из филогенеза, описанного выше, с использованием библиотеки Python ete3. Использовались две различные метрики патристического расстояния, далее обозначенные как «1» и «2», основанные на сумме длин всех ребер, соединяющих

вершины («1»), и количестве внутренних вершин на пути от начала графа к концу («2») соответственно. Евклидово расстояние между точками на «упругой карте» использовалось для сравнения с патристическим расстоянием. Данные в 63-мерном евклидовом пространстве были дважды нелинейно преобразованы и спроецированы на упругую карту с сохранением реальной связи между геномами. 2D-координаты проекций этих точек на карту сравнивались с патристическими расстояниями. Для построения попарных разностей, а также для расчета корреляций между расстояниями использовались специальные скрипты на языке R. Коэффициент  $\tau$ -Кендалла использовался для минимизации возможных эффектов автокорреляции. Для группировки последовательностей использовались два признака: таксономическая идентичность вирусов была выбрана как предположительно наиболее филогенетически информативная, а вид вызываемого заболевания, являющийся вторым признаком, считался наиболее практически важным.

Полученные результаты, основанные на очень специфичном наборе геномных данных, показали, что методы сравнения без выравнивания, свободные от неформального получения знаний, обладают высокой вычислительной производительностью и дают альтернативную информацию, которая потенциально может быть очень важной для филогенетического анализа. Применение эластичных карт приводит к кластеризации последовательностей ДНК, которые могут отражать не только эволюцию клад, но и некоторые другие скрытые особенности и, таким образом, предоставлять дополнительную информацию для эволюционного анализа, основанного на модели выравнивания и молекулярной эволюции. Следует отметить, что методы без выравнивания имеют значительное преимущество в вычислительной сложности и не зависят от качества выравнивания. По-видимому, использование методов сравнения без выравнивания наряду с молекулярно-филогенетическим анализом может дать дополнительную информацию о последовательностях и помочь в решении различных биологических задач.

Геномы коронавирусов являются очень хорошим объектом для сравнительного анализа методов сравнения последовательностей на основе выравнивания и без выравнивания. В данном исследовании предпринята попытка сравнить внутреннее структурирование 69 геномов коронавирусов, выявленных разными методами сравнения: 1) традиционное выравнивание с последующим построением филогении, 2) неконтролируемая классификация на основе k-средних и 3) современный метод нелинейной статистики на основе упругих карт. Сравнение показало высокую эффективность каждого из этих методов. В целом кластеры, выделенные по

филогенетическому дереву на основе множественного выравнивания последовательностей, хорошо согласовывались с классами, выделенными по  $k$ -средним, а слоистый граф, полученный из набора классификаций с увеличением числа классов, представлял собой структура, которая является полностью естественной с биологической точки зрения. Следует подчеркнуть, что классификация с использованием  $K$ -средних и кластеризация методом эластичных карт проводились в пространстве частот триплетов, определенных по всему геному.

Обнаружено, что некоторые геномы, далекие друг от друга по филогенезу, принадлежат к одному и тому же кластеру на упругой карте, при этом сходные состояния признака часто проявляются в разных кладах, а в одну и ту же эволюционную кладу входят представители разных «упругих» кластеров. Например, HCoV человека NL63 (альфакоронавирус) и HKU1 (бетакоронавирус), камышницы обыкновенной HKU21 (дельтакоронавирус) и норки WD1127 (альфакоронавирус) сформировали кластер 5; три птичьих гаммакоронавируса (IBV, IBVB и TCoV) и бетакоронавирус ежа *Echinaceus\_VMC* сформировали кластер 6. Геномы двух птичьих дельтакоронавирусов белоглазки HKU16 и луковичного HKU11, бетакоронавируса человека SARS-CoV и коронавируса летучих мышей BM48\_31 сформировали кластер 2, но этот кластер вел себя адекватно в многомерном пространстве, в том смысле, что он был близок к кластерам, образованным бета-коронавирусами человека SARS-CoV-2 и MERS-CoV, а также был достаточно близок к кластеру, образованному дельтакоронавирусами. Таким образом, одновременное применение обоих подходов к одному и тому же набору данных может привести к обнаружению деталей, которые иначе остались бы незамеченными. Важно, что комбинированный подход может позволить предположить, что в случае эволюционно далеких геномов, но все же принадлежащих к одному и тому же кластеру на эластичной карте, некоторые из них, вероятно, скоро изменят состояние этого признака. В случае группы заболеваний это означает, что этот вирус может быть «готов» к изменению типа вызываемого заболевания и/или к смене хозяина.

### **3.2 Модификации непараметрических алгоритмов автоматической классификации в условиях больших объёмов статистических данных**

В отчёте предлагается модифицированный непараметрический алгоритм автоматической классификации статистических данных большого объёма. Его применение на основе сжатия исходной информации позволяет обнаруживать классы, соответствующие одномерным фрагментам плотности вероятности многомерной

случайной величины. По сжатым данным формируется регрессионная оценка плотности вероятности, которая является основой алгоритмизации процедуры автоматической классификации. Эффективность разработанной методики автоматической классификации подтверждается результатами её применения при анализе данных дистанционного зондирования лесных массивов, повреждённых сибирским шелкопрядом.

Пусть имеется выборка многомерных статистических данных  $V = (x_v^i, v = \overline{1, k}, i = \overline{1, n})$  большого объёма  $n$ , распределённая с неизвестной плотностью вероятности  $p(x) = p(x_1, \dots, x_k)$ . Необходимо разбить выборку  $V$  на группы компактных наблюдений  $V_j, j = \overline{1, M}$ , количество  $M$  которых неизвестно. Под компактной группой наблюдений считается область их значений, соответствующая одномодальному фрагменту плотности вероятности  $p(x)$ .

Следуя методике синтеза регрессионной оценки плотности вероятности, исходная выборка  $V$  преобразуется в массив данных  $\bar{V} = (z^i, \bar{P}^i, i = \overline{1, \bar{N}})$ . Здесь  $\bar{N}$  – количество многомерных интервалов дискретизации области определения плотности вероятности  $p(x)$ ,  $z^i$  – их центры, а  $\bar{P}^i$  – частоты попадания наблюдений из выборки  $V$  в интервалы дискретизации. Оценивание оптимального количества интервалов дискретизации значений случайных величин ранее рассмотрено в работах А.В. Лапко и В.А. Лапко. Для этого используются результаты анализа асимптотических свойств регрессионной оценки плотности вероятности.

Для обнаружения первого класса по информации массива данных  $\bar{V}$  выберем элемент  $(z^j, \bar{P}^j)$ , который определяется условием

$$\bar{P}^j = \max_{t=1, \bar{N}} \bar{P}^t.$$

Этот элемент  $(z^j, \bar{P}^j)$  соответствует максимальному значению регрессионной оценки плотности вероятности  $\bar{p}(x)$ . Тогда начальный этап формирования элементов выборки  $\bar{V}$ , принадлежащих первому классу  $\Omega_1$ , определяется правилом

$$z^i \in \Omega_1, \text{ если } \prod_{v=1}^k \Phi\left(\frac{z_v^i - z_v^j}{c_v}\right) 1(\bar{P}^i, \bar{P}^j) > 0, i = \overline{1, \bar{N}}, i \neq j.$$

Ядерные функции  $\Phi(\cdot)$  обладают свойствами нормированности и симметричности относительно  $z_v^j, v = \overline{1, k}$ . При  $c_v > 2\beta_v$  ядерные функции принимают нулевые

значения. В этих условиях определяются смежные к  $j$ -му интервалы дискретизации области значений случайной величины. Индикаторная функция  $1(\bar{P}^i, \bar{P}^j)$  принимает значение 1, если  $\bar{P}^i \leq \bar{P}^j$  и значение 0, если  $\bar{P}^i > \bar{P}^j$ . В этих условиях из смежных интервалов дискретизации определяются интервалы, принадлежащие классу  $\Omega_1$ .

Обозначим через  $I_1$  - множество номеров элементов массива данных  $\bar{V}$ , отнесённых на начальном этапе автоматической классификации к первому классу  $\Omega_1$ .

Множество  $I_1$  включает номер  $j$  элемента  $(z^j, \bar{P}^j)$  массива данных  $\bar{V}$ .

Второй этап классификации осуществляется решающим правилом

$$z^i \in \Omega_1, \text{ если } \sum_{j \in I_1} \prod_{v=1}^k \Phi\left(\frac{z_v^i - z_v^j}{c_v}\right) 1(\bar{P}^i, \bar{P}^j) > 0, i \in I \setminus I_1, i \neq j, I = \overline{1, N}.$$

По аналогии на  $t$ -ом этапе автоматической классификации элементы выборки  $\bar{V}$  принадлежащие первому классу  $\Omega_1$ , определяются правилом

$$z^i \in \Omega_1, \text{ если } \sum_{j \in I_{t-1}} \prod_{v=1}^k \Phi\left(\frac{z_v^i - z_v^j}{c_v}\right) 1(\bar{P}^i, \bar{P}^j) > 0, i \in I \setminus (I_1 \cup I_2 \cup \dots \cup I_{t-1}).$$

Эта последовательная процедура автоматической классификации продолжается до выполнения условия  $I_t = \emptyset$ , когда отсутствует факт обнаружения новых элементов массива данных  $\bar{V}$  из класса  $\Omega_1$ .

Для обнаружения элементов массива  $\bar{V}_2 \in \bar{V}$ , принадлежащих второму классу  $\Omega_2$ , из оставшихся данных  $\bar{V} \setminus \bar{V}_1$  выбирается элемент  $(z^j, \bar{P}^j)$  с максимальным значением  $\bar{P}^j$  и предложенный выше алгоритм автоматической классификации используется для определения класса  $\Omega_2$ . По аналогии формируются оставшиеся классы, количество которых априори не определено.

Непараметрический алгоритм автоматической классификации и его программная реализация НАС v2.0 имеют государственную регистрацию (свидетельство о государственной регистрации программы для ЭВМ № 2022619023 от 18.05.2022). В программе реализованы методы многопоточных вычислений. Её применение позволяет загружать изображения в формате GeoTIFF, производить классификацию, получать изображение для дальнейшей геообработки.

Разработанные алгоритмические и программные средства использовались при исследовании состояния лесных массивов, повреждённых сибирским шелкопрядом.

Территория исследования определялась восточной частью Ибейского района Красноярского края. На этой территории преобладают пихтовые, кедровые и еловые древостои. Высоты изменяются в диапазоне от 500 до 900 метров над уровнем моря.

Исходная информация сформирована 1 июня 2022 года по данным дистанционного зондирования Земли аппаратом Европейского космического агентства Sentinel-2A. Снимок получен с геопортала Earth Explorer, из которого вырезан тестовый участок в 13,8 тысяч гектар (рис. 1а). Он определялся 1286776 пикселями. Каждый пиксель характеризовался шестью спектральными признаками  $x = (x_1, x_2, x_3, x_4, x_5, x_6)$ , которым соответствуют длины волн (нанометры): 492,7 ( $x_1$ ), 559,8 ( $x_2$ ), 664,6 ( $x_3$ ), 740,5 ( $x_4$ ), 832,8 ( $x_5$ ), 2202,4 ( $x_6$ ). Полученные данные подвергались атмосферной коррекции с помощью алгоритмов Dark Object Subtraction (DOS), реализованных в модуле свободной кроссплатформенной геоинформационной системы QGIS. На рис. 3.2.а представлено RGB-изображение исходных данных, где каналам R, G, B соответствуют признаки  $x_6$ ,  $x_5$ ,  $x_3$ .

Применение программы NAC v.2.0. позволило в исходном изображении обнаружить 320 компактных групп точек (Рис. 3.2.в). Из них 45 классов содержали более 0.1% от общего объема пикселей. Обнаруженные классы соответствуют лесным массивам с различной степенью поражения сибирским шелкопрядом, усохшим и лиственным древостоям, травяно-кустарниковым сообществам и заболоченным участкам исследуемой территории.

Результаты автоматической классификации в указанных условиях методом ISODATA средствами программного пакета ArcGIS ArcMap приведены на Рис. 3.2.г и 3.2.д. При настройке этого инструмента было указано количество классов равное обнаруженным непараметрическим алгоритмом автоматической классификации. В результате было получено классифицированное изображение с 237 компактными группами точек (Рис. 3.2.г). Пиксели между ними были распределены равномерно и в каждом классе содержалось более 0.1% от общего объема наблюдений. На Рис. 3.2.д приставлены результаты работы алгоритма ISODATA при количестве классов 45.

В центральной части исследуемой территории располагались лесные массивы усохших темнохвойных древостоев, поврежденных сибирским шелкопрядом. Экспертный анализ показал, что предлагаемый метод автоматической классификации обладает преимуществом по сравнению с алгоритмом ISODATA, что следует из анализа Рис. 3.1.а, 3.1.в, 3.1.г, 3.1.д. Исходный снимок (Рис. 3.1.а) более близок к Рис. 3.1.в, чем к Рис. 3.1.г. Рассматриваемые алгоритмы автоматической классификации достаточно полно выделяют зоны поврежденных древостоев, разделяют темнохвойные и светлохвойные породы,

определяют заболоченные территории и участки с открытой почвой. По результатам автоматической классификации алгоритмом ISODATA и непараметрическим классификатором были рассчитаны площади территорий поврежденных древостоев сибирским шелкопрядом (Рис. 3.1.б). Они составили 1825 и 1934,5 гектаров лесных массивов соответственно, что указывает на их различие в 6%.

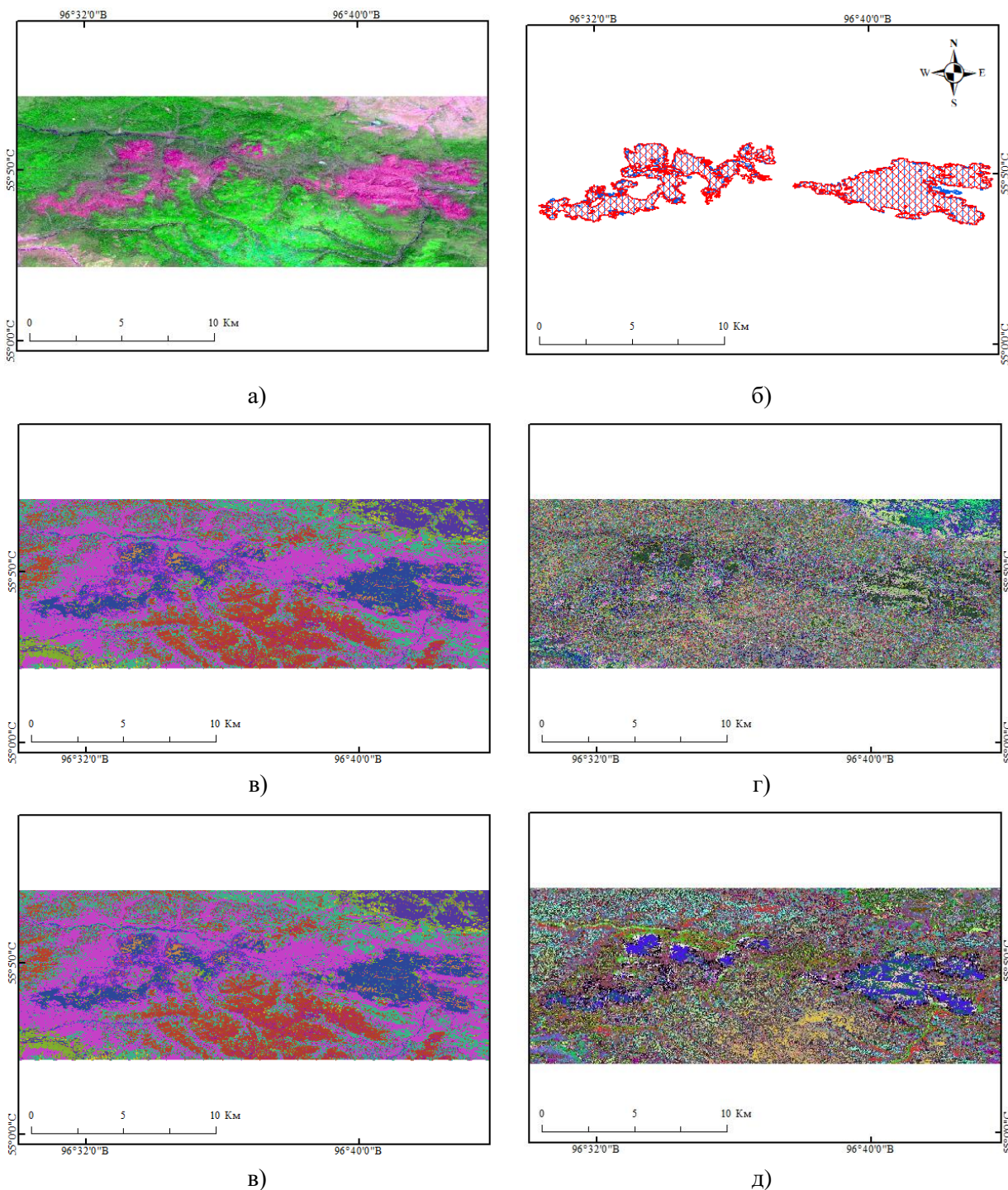


Рисунок 3.2 – Сопоставление исходного снимка (а) с результатами классификации непараметрическим алгоритмом (в), методом ISODATA при 237 классах (г), при 45 классах (д) и отображение векторного представления поврежденных лесных насаждений сибирским шелкопрядом (б)



### **3.3 Мета модель и алгоритмы детализации интегральных оценок для определения причин состояния природно-техногенной безопасности территорий**

Разработана метамодель детализации интегральных оценок для определения причин состояния природно-техногенной безопасности территорий. Модель является основой информационной поддержки по планированию мероприятий долгосрочного развития территорий и формированию стратегий снижения рисков чрезвычайных ситуаций. Исследование представляет развитие метода интегрального аналитического оценивания состояния территорий. Обеспечена возможность детализации комплексного показателя природно-техногенной безопасности до причин, которые могут обуславливать его выход за пределы нормативных значений. Выполнен анализ связей «причина отклонения от нормы – рекомендуемые превентивные мероприятия». Это позволяет завершить цикл получения и обработки данных комплексного мониторинга до формирования управляющих воздействий с учётом особенностей территорий.

Комплексный показатель природно-техногенной безопасности территорий формируется путем построения иерархического представления оценок показателей, характеризующих природные и техногенные факторы риска возникновения чрезвычайных ситуаций (ЧС). Расчёт оценок выполняется на основе территориально-ориентированной нормативной модели, которая определяет «желаемый» уровень безопасности с учетом индивидуальных особенностей территорий и реальных возможностей его достижения. Нормативная модель включает иерархическую систему показателей, коэффициенты значимости показателей, функции агрегирования оценок, нормативные значения показателей и коэффициенты чувствительности оценок.

Определение причин текущего состояния природно-техногенной безопасности территорий основано на анализе сложившейся ситуации по результатам интегрального оценивания. Поскольку возникновение проблемных ситуаций отражается в оценках комплексных показателей, метамодель определения причин состояния безопасности территорий строится на основе иерархии оценок. Мета модель представляет собой когнитивную карту, позволяющую анализировать сложившуюся ситуацию с выделением конкретных проблем и причин (Рис. 3.3).

Под «ситуацией» здесь понимается сочетание условий и обстоятельств, создающих определённую обстановку, в которой возникла «проблема». Проблема – несоответствие желаемого и фактического уровней достижения целей, возникшее в результате определенных «причин». Для определенного сочетания проблем и причин формируются управляющие рекомендации, направленные на достижение желаемого уровня.

Особенность метамодели состоит в применении иерархических зависимостей между причинами возникновения проблемных ситуаций.

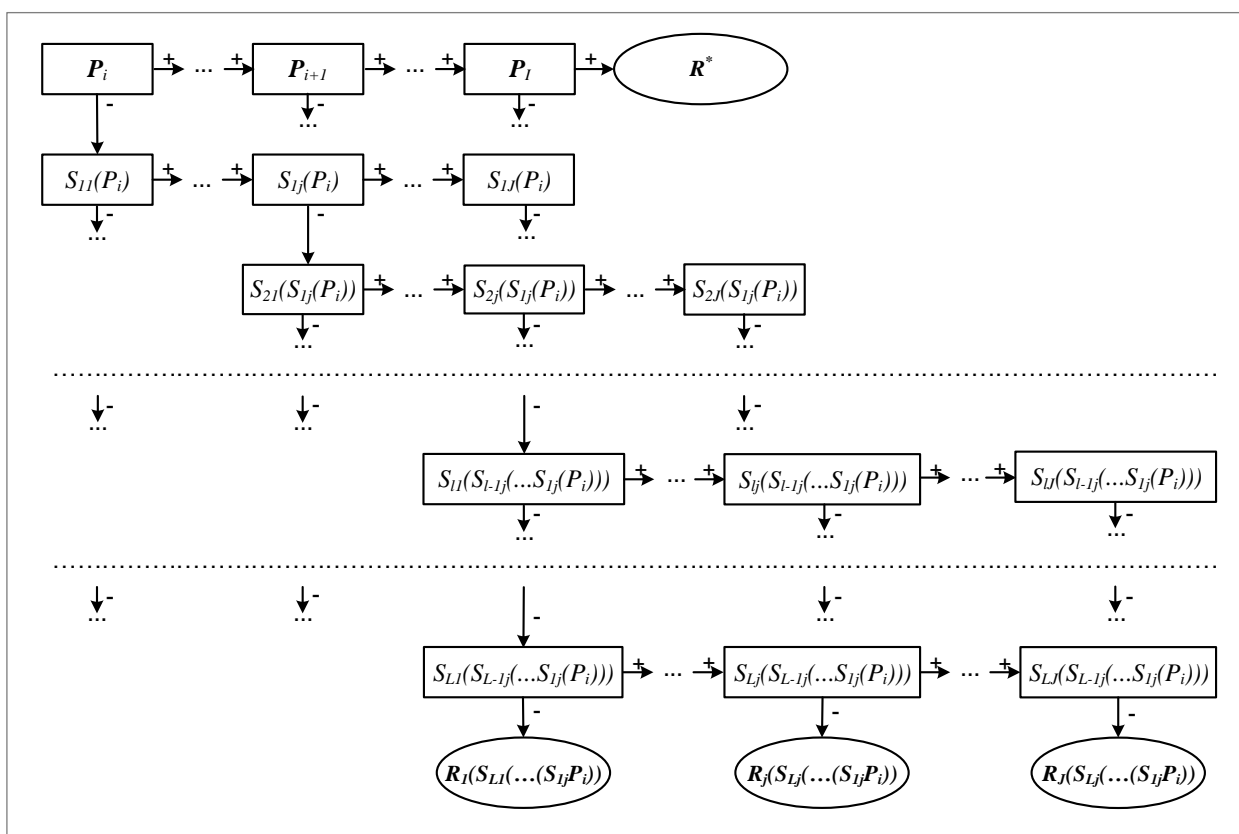


Рисунок 3.3 – Метамодель определения причин состояния природно-техногенной безопасности территорий

В ходе моделирования ситуации определяется показатель «проблема»  $P_i$ , характеризующий текущее состояние, где  $i = \overline{1, I}$  – количество возникших «проблем». Для каждого показателя «проблема» определяется совокупность иерархически подчиненных показателей «причина»  $S_{lj}(S_{l-1j}(\dots S_{1j}(P_i)))$ , которые характеризуют возникшую проблемную ситуацию, где  $l = \overline{1, L}$  – количество уровней иерархии «причин»,  $j = \overline{1, J}$  – количество показателей «причина» на уровне иерархии. Например, определена «проблема»  $P_1$ : оценка комплексного показателя «Природно-техногенная безопасность» имеет значение «Пониженный». Согласно иерархии комплексных показателей, данная «проблема» связана со следующими причинами: «Техногенная безопасность»  $S_{11}(P_1)$  и «Природная безопасность»  $S_{12}(P_1)$ . В свою очередь, каждый из указанных показателей «причина», содержит дочерние показатели «причина». Так, для показателя  $S_{11}(P_1)$  определяются «причины»:  $S_{21}(S_{11}(P_1))$  «Обстановка на объектах техносферы»,  $S_{22}(S_{11}(P_1))$  «Транспортная обстановка»,  $S_{23}(S_{11}(P_1))$  «Обстановка на объектах ЖКХ»,  $S_{24}(S_{11}(P_1))$  «Пожарная обстановка» и  $S_{25}(S_{11}(P_1))$  «Радиационная обстановка». Для

показателя  $S_{12}(P_1)$  определяются «причины»:  $S_{21}(S_{12}(P_1))$  «Метеорологическая обстановка»,  $S_{22}(S_{12}(P_1))$  «Гидрологическая обстановка»,  $S_{23}(S_{12}(P_1))$  «Геофизическая обстановка» и  $S_{24}(S_{12}(P_1))$  «Лесопожарная обстановка». И так далее, спускаясь вниз по иерархии до базовых показателей. При этом показатель «проблема» может задаваться на любом уровне иерархии комплексных показателей.

В когнитивной модели используется два типа причинно-следственных связей: положительные (+) и отрицательные (-), которые соответствуют состоянию показателей: «норма» и «не норма». Состояние «норма» и «не норма», обеспечивающее направленный переход между концептами метамодели, определяется исходя из значений интегральных оценок. Например, в качестве состояния «не норма» могут рассматриваться значения оценок «Пониженный», «Низкий» или «Критический», либо ухудшение оценки показателя по сравнению с предыдущим отчетным периодом. В случае, когда все показатели «проблема»  $P_i$  в «норме», формируется общая рекомендация  $R^*$ . Если «проблема»  $P_i$  в «не норме», то на каждом уровне иерархии существует показатель «причина» в состоянии «не норма», что обеспечивает движение вниз по иерархии. Рекомендация  $R_j(S_{Lj}(\dots S_{1j}(P_i)))$  формируется для случая, когда показатель «причина»  $S_{Lj}(S_{l-1j}(\dots S_{1j}(P_i)))$  уровня иерархии  $L$  в состоянии «не норма». Управляющие рекомендации могут меняться в зависимости от конкретного значения оценки.

В основу алгоритма детализации интегральных оценок состояния природно-техногенной безопасности территорий положен «дедуктивный» принцип реализации стратегического управления безопасностью территорий. На начальном уровне определяются территориальные приоритеты на основе расчета оценок интегральных показателей и составления картограмм снижения рисков. Формирование рекомендаций по планированию управляющих мероприятий выполняется с привлечением специалистов муниципальных образований. Затем осуществляется распределение опасностей по видам ситуаций. Такая детализация позволяет определить зону ответственности ведомств и отраслей, полномочия которых включают обеспечение безопасности по видам обстановок. Далее анализируются факторы, являющиеся причинами опасных событий или возникновения ЧС. Для каждого вида обстановки причины определяются исходя из разных способов группировки факторов рисков (управляемые/неуправляемые; природные/антропогенные/социальные; детерминированные/вероятностные и т.п.). На основе выявленных факторов риска проводится обоснование объёмов планируемых мер и управляющих мероприятий (Рис. 3.4).

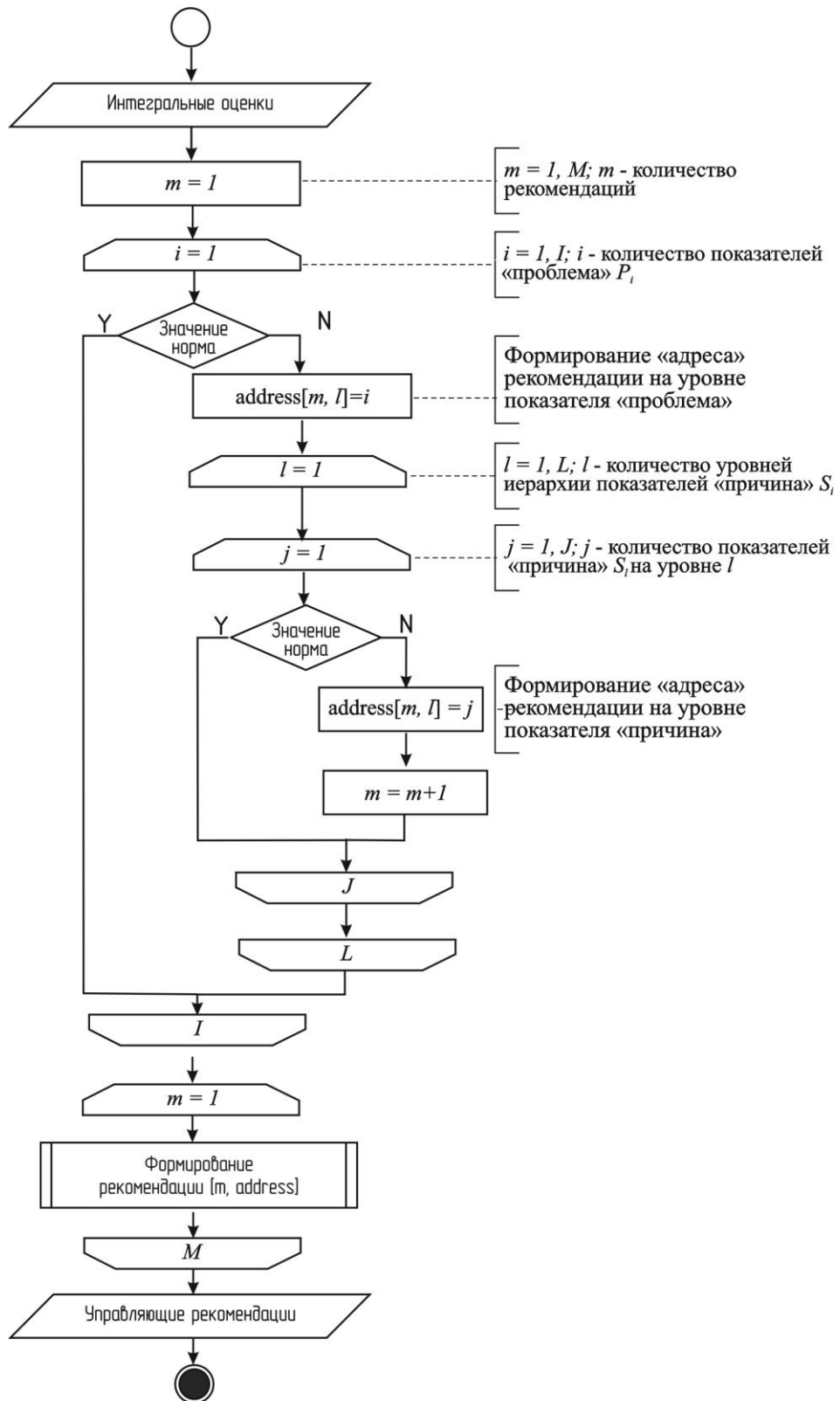


Рисунок 3.4 – Алгоритм детализации интегральных оценок природно-техногенной безопасности территорий

Работа алгоритма заключается в определении причин возникновения проблемы путем проверки значений интегральных оценок на соответствие «норме» и

формировании рекомендаций по нивелированию негативных факторов. Движение по иерархии показателей обеспечивают три цикла. Сначала проверяется соответствие значений интегральных оценок для показателей «проблема». При несоответствии одного или более показателей «проблема» запускается «маршрут» движения по уровням иерархии показателей «причина» с проверкой значений их оценок на каждом уровне. Как правило, показатель «проблема» имеет несколько показателей «причина», у которых оценка не соответствует норме. После проверки всех показателей «причина», выполняется формирование управляющих рекомендаций. Число рекомендаций равно числу выявленных «причин» последнего уровня. Процесс формирования управляющих рекомендаций основывается на анализе факторов риска для базовых показателей. Накопленные данные мониторинга опасных событий и паспорта безопасности территорий позволяют обосновать применение рекомендаций для конкретного объекта, оценить необходимые ресурсы и ожидаемый эффект.

Предложенные метамодель и алгоритм на основе иерархии интегральных оценок позволяют проанализировать текущее состояние природно-техногенной безопасности территорий, выявить причины возникновения проблемных ситуаций и, двигаясь от общих обстановок до конкретных факторов риска, сформировать обоснованные управляющие рекомендации. С использованием предложенной модели сформированы рекомендации по корректировке региональной программы обеспечения пожарной безопасности и мероприятий подготовки к лесопожарному сезону 2023 года. Показана возможность автоматизации стратегического планирования на основе инструментов расчёта интегральных оценок и визуализации результатов. Кроме того, оценка рисков пожаров на основе представленного информационного обеспечения используются при обосновании размещения вновь создаваемых противопожарных формирований, контроля работоспособности и обустройства источников водоснабжения, оборудования защищаемых объектов противопожарной сигнализацией, проведения обучения населения.

## **4 Технологии создания информационно-аналитических систем на основе интегрированных платформ**

### **4.1 Методические и программные средства поддержки оперативной аналитической обработки данных научных исследований в рамках модельно-ориентированных систем сбора данных**

В условиях обилия информационных потоков, возрастающих вычислительных мощностей и повышения требований к обоснованности принимаемых решений, приобретают актуальность отчетные формы, позволяющие пользователям оперативно получать аналитические данные, не беспокоясь о возможном изменении программного кода и манипуляциях с данными. Это особенно необходимо в критически важных областях деятельности. Несмотря на повсеместное внедрение оперативных информационных панелей и дашбордов, автоматически обновляемых в момент поступления новых данных, важной частью тактического и стратегического управления являются регламентированные срочные аналитические отчетные формы.

Современное решение данной проблемы – это low code/no code системы (LC/NC), которые позволяют вносить изменения в клиентское приложение (системы) без программирования совсем или с незначительным объемом программного кода. Суть современной концепции заключается в тезисе: «если участие человека необходимо, то пусть это будет аналитик, а не программист». Технологической основой разработки «без программирования» является модельно-ориентированный подход, который позволяет создавать системы за счет отделения бизнес-логики из программного кода в метаданные и формирования алгоритмов построения прикладных систем под управлением этих метаданных. Структура метаданных задает потенциальную функциональность системы, а данные, хранящиеся в этой структуре, определяют тематику системы, ее объем и информационное наполнение.

С целью создания «no-code» платформы для создания систем сбора данных на основе модельно-ориентированного подхода ранее предложена авторская реализация классического процесса MDD, в которой объединены процессы формирования управляющей и прикладной моделей. Объединенный процесс предполагает, что управляющая модель системы формируется пользователем с помощью интерфейса самой системы. При этом автоматически изменяется структура базы данных и сохраняется формальное описание моделей. Процесс автоматической трансформации прикладных моделей добавлен в пользовательский интерфейс модельно-ориентированной системы, что позволяет выполнять динамическую интерпретацию метамодели. Благодаря этому,

интерфейс системы в части ввода данных перестраивается автоматически, в соответствии с изменениями, внесенными пользователем в управляющую модель. В основе авторской реализации, как и в основе классического подхода, лежит концептуальная мета-мета-модель, определяющая основные классы сущностей и связи между ними для построения управляющих моделей.

Для решения задачи разработки типовых модельно-ориентированных систем консолидации данных предложена мета-мета-модель сбора данных, которая является спецификацией для создания управляющих и прикладных моделей проектируемых систем, обеспечивает динамическую интерпретацию мета-модели и генерацию форм сбора данных. Мета-мета-модель содержит три класса объектов: «Объект», «Атрибут» и «Группа». Объекты класса «Атрибут» описываются свойствами: наименование, тип, и флаг темпоральности атрибута. В рамках мета-мета-модели задан ряд отношений, определяющих связи между экземплярами классов модели.

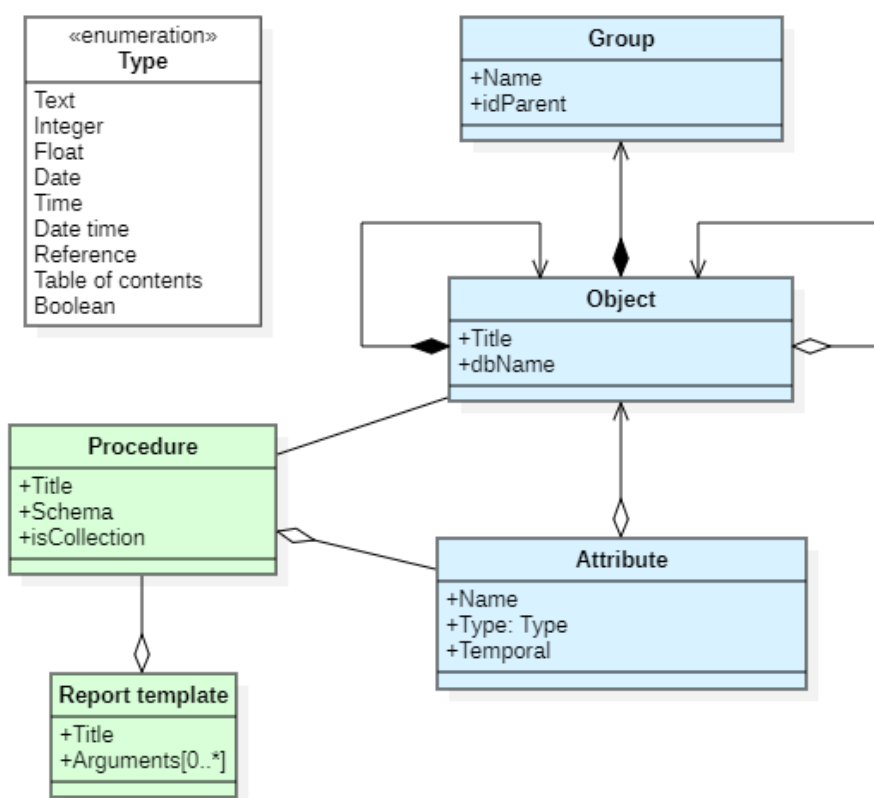


Рисунок 4.1 – Модифицированная мета-мета-модель модельно-ориентированной платформы (новые классы мета-модели окрашены зеленым)

Реализованный в платформе модельно-ориентированный подход применен к генерации аналитических отчетов. Информация о структуре и наполнении отчетов представляется как метаданные, а в программный код, отвечающий за формирование

отчетов на основе метамодели, включается унифицированный модуль генерации отчета. Применение модельно-ориентированного подхода при создании модуля генерации аналитических отчетов подразумевает модификацию мета-метамодели для полноценной интеграции новой функциональности в систему. Модификация заключается во внесении в структуру метаданных сущностей, без которых процесс генерации отчетов невозможен. Отчет принято разделять на статическую и динамическую составляющие. Статическая часть – шаблон отчета – определяет индивидуальную структуру отчета в соответствии с назначением документа: заголовок и параметры генерации, блоки, составляющие отчет, и их наполнение. Динамическая часть формируется в момент генерации отчета по заданному шаблону. В мета-метамодель введены такие сущности, как «шаблон» и «процедура», отвечающие за статическую и динамическую составляющие отчета соответственно. Процедуры позволяют перенести сбор и группировку данных из программного кода в базу данных. При необходимости изменить набор данных, добавить новые данные в отчет или удалить их (благодаря процедурам) не нужно редактировать программный код, а достаточно указать новую процедуру в шаблоне или удалить её оттуда. Модифицированная мета-метамодель представлена на Рис. 4.1.

С целью реализации механизма динамической генерации печатных отчетных форм модель дополнена двумя классами: «Procedure» и «Report template». Новые классы описывают организацию доступа к консолидированным данным в виде хранимых процедур и их связь с шаблоном печатной формы, использующим эти данные. Класс «Report template» в настоящей модификации содержит название шаблона, по которому программный модуль найдет выбранный шаблон на сервере. Класс «Procedure» содержит имя процедуры для получения оперативных данных, схему в базе данных, разграничивающую процедуры из разных разделов системы, и логическое поле «Коллекция». Поле «Коллекция» принимает значения: «Истина» или «Ложь». Значения определяют, какого рода данные будут получены и как их следует обрабатывать. Значение «Истина» означает, что результатом выполнения хранимой процедуры будет множество записей, которые могут быть представлены в отчете как таблица или могут требовать дополнительной обработки. Значение «Ложь» соответствует ситуации, когда процедура возвращает одну запись или скалярное выражение. Такие данные могут быть представлены в отчете «как есть» или могут использоваться в обработке данных, полученных от процедур со значением параметра «Истина».

Апробация выполнена на примере оперативной сводки, формируемой Центром обеспечения реализации полномочий в областях гражданской обороны и чрезвычайных



ситуаций Красноярского края (Центра ГО и ЧС). В зависимости от сезона в аналитическом отчете меняется перечень контролируемых обстановок: подтопления и лесные пожары характерны для весенне-летнего периода, а зимой необходим мониторинг ледовых переправ. С окончанием отчетного года в некоторые разделы добавляются новые столбцы для сопоставления показателей в предшествующие отчетные периоды. Ввиду развития средств инструментального контроля добавляются новые разделы. Применение новых инструментов не требует непосредственного участия человека для адаптации и настройки новой версии аналитического отчета.

#### **4.2 Программное обеспечение трансформации данных различных форматов путем сопоставления структурных элементов и выполнения преобразований по установленным взаимным отношениям**

Пусть реляционный набор данных объединяет  $n$  атрибутов:  $\{A_1, A_2, \dots, A_n\}$ . У каждого из атрибутов есть множество принимаемых значений – домен. Обозначим домен  $i$ -го атрибута как  $D_i$ , где  $i$  изменяется от 1 до  $n$ . Декартово произведение доменов записывается как  $D_1 \times D_2 \times \dots \times D_n$  и представляет все возможные комбинации значений атрибутов.

Название атрибута и множество принимаемых значений представляют собой свойство объекта, обладающего разными структурными особенностями с точки зрения аналитической обработки данных. Для анализа данных важно деление атрибутов по типам шкал: номинальный (категориальный), порядковый, интервальный (дискретный) и нормативный (непрерывный). Номинальные данные по определению не упорядочены и, как правило, представлены строковыми данными. К строковым данным невозможно применить стандартные методы агрегирования (сумма, среднее, минимум, максимум), но можно посчитать количество категорий и проанализировать процентное соотношение, если это соотносится с поставленной аналитической задачей. Можно заметить, что число принимаемых значений одних атрибутов близко или равно числу записей в исследуемом наборе данных, а для других атрибутов характерно существенно меньшее число принимаемых значений. В первом случае мы можем говорить о наличии сопроводительной неструктурированной информации, которая без дополнительного исследования и обогащения не пригодна для анализа. Во втором случае атрибут можно интерпретировать как некие категории, которые могут быть использованы в ходе аналитической обработки.

Данные порядковых типов шкал более информативны по сравнению с номинальными: значения таких атрибутов можно упорядочить. И порядковые, и номинальные данные относятся к качественным характеристикам и, как правило,

представлены в наборах в виде строковых значений, однако не редки случаи, когда атрибут может быть числовым даже с возможностью упорядочивания значений, но семантической нагрузки эти числа не несут и никак не характеризуют объект исследования. Такие данные могут быть использованы только для определения количества уникальных записей в наборе.

Количественные атрибуты подразделяются на два типа шкал: «интервальный» и «нормативный». На пути к увеличению информативности данных за порядковыми атрибутами следуют «интервальные». Значения «интервальных» атрибутов различны, упорядочены и можно оценить разницу между значениями в числовом выражении. Основным критерием выступает степень влияния оценок на решения, принимаемые на основе предоставленной информации. Самая детальная информация содержится в «нормативных» данных. Их главная отличительная особенность – это то, что значения могут быть соотнесены с «0», и ноль в этом случае будет значить отсутствие значения. Для «нормативных» данных доступен уже полный спектр функций агрегирования, включая сумму. Приведенная классификация типов шкал определяет тактику аналитической обработки данных при построении многомерных моделей и визуализации результатов анализа.

Функциональные зависимости отражают связи между атрибутами набора данных, которые присущи реальным сущностям в исследуемой предметной области. Говорят, что  $A_1$  функционально зависит от  $A_2$ , если в любой момент времени каждому значению  $A_2$  соответствует единственное значение  $A_1$ , записывается как  $A_2 \rightarrow A_1$ . Атрибут  $A_2$  в функциональной зависимости  $A_2 \rightarrow A_1$  называется *детерминантом*. Фактически функциональная зависимость определяет направление аналитической агрегации данных при многомерном моделировании.

Предлагается следующая последовательность построения многомерной модели.

1. Поиск показателей – выбор числовых атрибутов с интервальным и нормативным типом шкал, к которым применимы большинство функций агрегирования. В теории многомерного моделирования показатели, характеризующие один процесс, объединяются в факт.

2. Из числа оставшихся атрибутов, номинальных и порядковых, выбрать те, которые могут быть категориями для анализа – измерениями. Отличительной особенностью измерений является то, что число принимаемых значений этих атрибутов существенно меньше числа строк в наборе данных.

3. Номинальные и порядковые атрибуты, число уникальных значений у которых близко к числу строк во фрейме, относятся к описательным атрибутам.

Считается, что они содержат описательную текстовую информацию, не пригодную для аналитической обработки без дополнительного обогащения.

4. Определить ключ исследуемого реляционного набора данных путем поиска функциональной зависимости с минимальным детерминантом (может содержать более одного атрибута), которая содержит все атрибуты фрейма. Детерминант функциональной зависимости и становится ключом. Измерения, входящие в ключ фрейма, определяют максимальный уровень гранулярности многомерной модели (максимальный уровень детализации).

5. Выстроить иерархию среди выявленных измерений путем выявления между ними функциональных зависимостей, начиная с ключевых атрибутов.

Построенная таким образом структура задает многомерную модель для реляционного набора данных. Формально многомерная модель данных можно представить в графическом виде как «модель факт-измерение» (fact-dimensional model), инфографически – в виде OLAP-куба (многомерного куба), и аналитически – как сводная таблица (кросс-таблица).

В графической нотации многомерной модели факт отображается в виде прямоугольника, объединяя показатели; заголовок элемента указывает на анализируемый процесс. Измерения представлены в виде кругов, соединенных с фактом и друг с другом, образуя иерархию. Измерения, соединенные непосредственно с фактом, соответствуют максимальному уровню гранулярности (grain). Ветки иерархии измерений указывают направление агрегации. На Рис. 4.2 представлен пример графического представления многомерной модели для задачи анализа подтоплений в Красноярском крае.



Рисунок 4.2 – Графическое представление многомерной модели для задачи анализа подтоплений в Красноярском крае

Информация о количестве подтопленных объектов – это то, на чем концентрируется многомерная модель (факт). Атрибуты, которые могут играть роль показателей: «Количество проживающих детей», «Общее количество проживающих детей», «Количество жилых домов», «Количество детей в ПВР», «Общее количество людей в ПВР», «Количество взрослых в ПВР», «Количество объектов подтопления» (представлены оранжевым цветом). Шкалы этих атрибутов относятся к нормативному типу, то есть представляют собой самый удобный вид данных, к которым применимы любые функции агрегирования без потери смысла.

Из оставшихся атрибутов, которые имеют номинальный и порядковый тип шкал выбраны те, которые могут стать измерениями для анализа: «Событие», «Дата», «Время», «Пункт временного размещения», «Муниципальное образование и округ», «Категория зоны подтопления», «Зона подтопления», «Категория объекта подтопления», «Река».

Библиотеки на языке python, совместимые с объектами pandas, предоставляют широкий спектр возможностей по отображению наборов данных. Принимая во внимание типы шкал атрибутов, функциональные зависимости и аналитические свойства элементов многомерной модели предложен следующий формальный подход к выбору средств выражения при построении диаграмм.

- Цвет может использоваться для отображения интервальных и нормативных данных. Цвет сам относится к номинальным шкалам, но благодаря большому числу значений и возможности задания легенды может служить для отображения практически любых данных.

- Заполнение – номинальная шкала, которая подходит для отображения любых данных, приведенных к номинальной шкале. Даже нормативные данные могут быть классифицированы или кластеризованы.

- Форма служит для отображения данных в номинальных шкалах.

- Размер используется для отображения нормативных данных. Здесь важна возможность числового выражения разницы между значениями и возможность интерпретации нуля как отсутствия значений.

- Оси являются одним из средств графического представления информации на диаграмме. Они позволяют отобразить интервальные и нормативные данные. Допустимо использовать оси для отображения данных с порядковыми шкалами.

- Карта используется для отображения пространственно распределенных данных.

В зависимости от аналитических свойств исходных данных можно формально задать процесс трансформации реляционного набора данных в многомерную модель, а затем в инфограмму.

### 4.3 Исследование кибербезопасности систем облачных вычислений для различных способов обслуживания

В отчетном периоде проводились исследования в области безопасности облачных вычислений для модели обслуживания SaaS (Software-as-a-Service). На основе телекоммуникационной инфраструктуры ФИЦ КНЦ СО РАН и реализованного технологического стенда инфраструктуры Интернета вещей (IoT) (Рис. 4.3) были налажены сбор и обработка данных взаимодействия различных сервисов, предложены методы анализа и выявления рисков киберугроз.

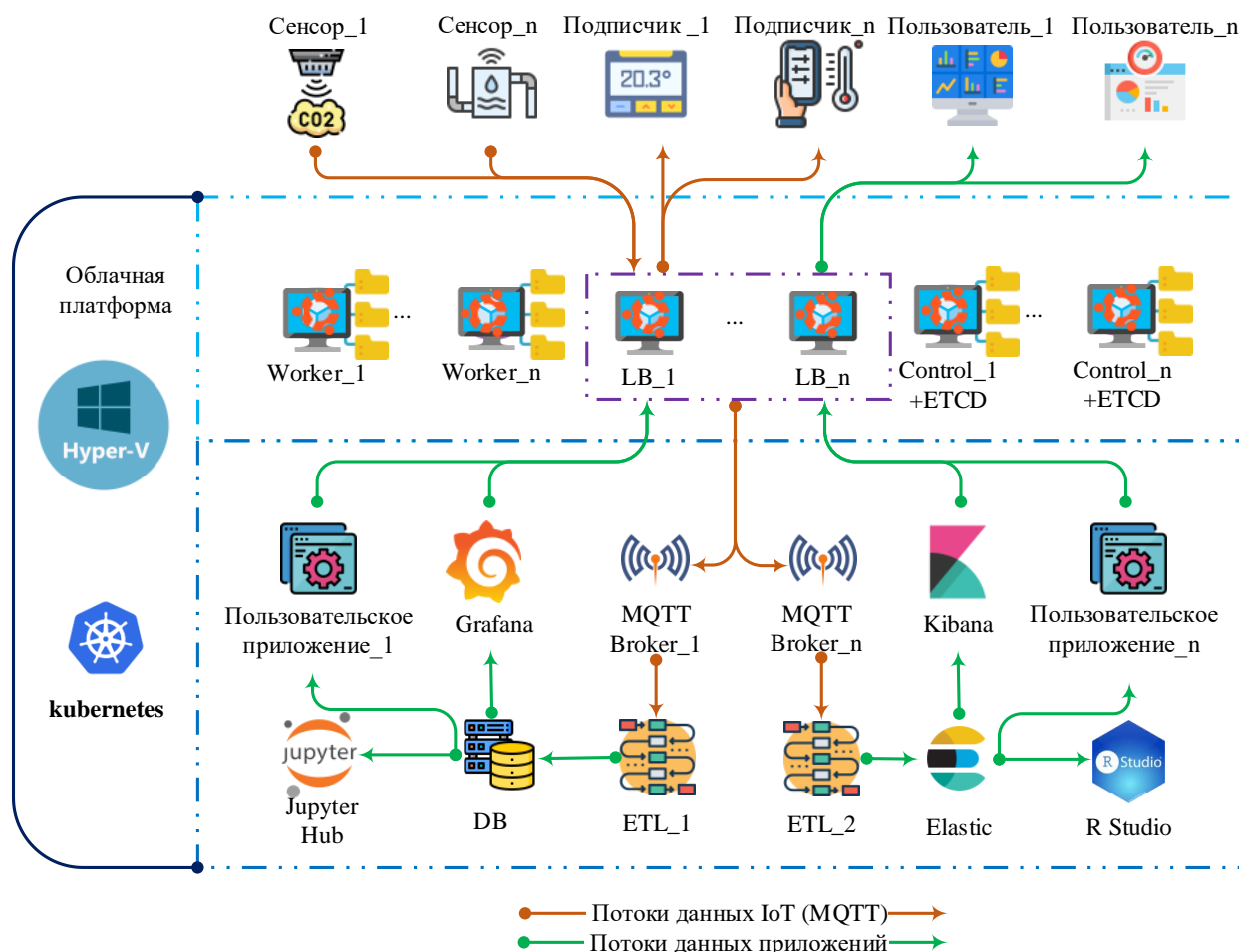


Рисунок 4.3 – Концептуальная схема реализованной инфраструктуры IoT

В частности, отмечено, что взаимодействие через сетевые инфраструктуры пространственно-распределённых узлов IoT на основе модели облачных вычислений может вызывать существенные нагрузки на сети и потребителей информации. Для

обеспечения безопасной работы облачных сервисов и снижения интенсивности сетевого трафика разработан *Метод адаптивных рассылок*, позволяющий выполнять синхронизацию режимов информационного взаимодействия между источниками данных и средствами мониторинга на основании частотных характеристик наблюдаемых процессов. Создание новых методов, позволяющих сокращать объёмы передаваемых данных без потери их информационной значимости, является актуальным в гетерогенных, распределённых и иерархических архитектурах.

Разработка и апробация метода адаптивных рассылок выполнена для настройки режимов выдачи данных в сетевой инфраструктуре Интернета вещей, в состав которой входят устройства сенсорного уровня, маршрутизаторы трафика из внешней сети, узлы кластера с контейнерами приложений, брокеры данных, базы данных, средства анализа и визуализации. Инфраструктура основана на инструментах, входящих в международную ассоциацию нативных облачных вычислений Cloud Nativ Computing Foundation (CNCf), что обеспечивает совместимость разработанных методов с другими облачными сервисами.

Для формирования критериев настройки адаптивных рассылок выполняется исследование данных, получаемых от устройств IoT. Они представляют собой функцию дискретного аргумента, принимающую произвольные значения в промежутки времени  $kT$ :  $x(t)=\{x(kT)\}$ , где  $x(kT)$  –  $k$ -тый результат наблюдений в момент  $kT$ . Выборка измерений в  $N$  отсчётах приведена на Рис. 4.4.

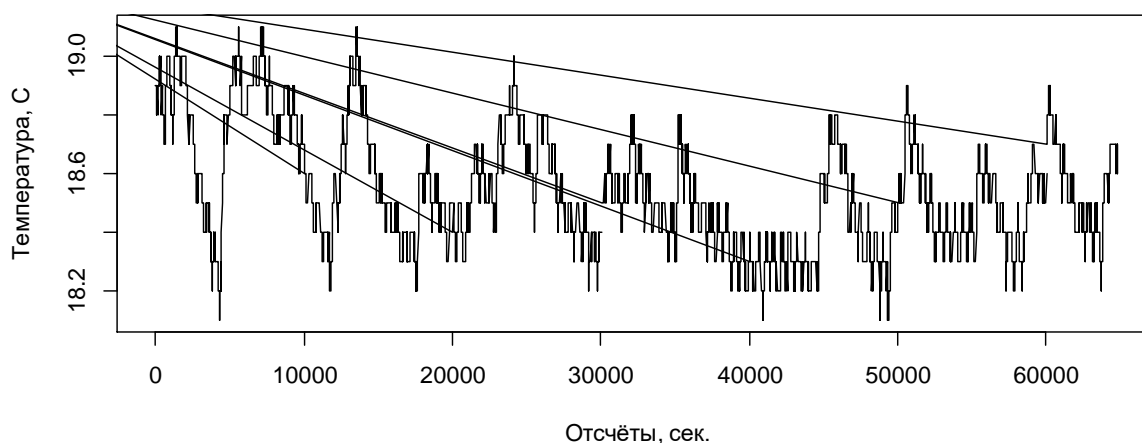


Рисунок 4.4 – Данные наблюдений

Время между наблюдениями  $\Delta t=(t_{n+1}-t_n)+\tau$ , где  $\tau$  – задержка передачи данных. Значение  $\tau$  достаточно мало, поэтому период дискретизации считается постоянным и равным  $T=(t_{n+1}-t_n)$  для всего периода наблюдений  $NT$ .

Алгоритм действий для определения периода дискретизации состоит в следующем.

- 1) Выполнить дискретное преобразование Фурье для функции  $x(t)$ , результатом которого является дискретный спектр  $X(k)$ , то есть по  $N$  измерениям значений  $x(t)$  построить  $N$  спектральных отсчётов на одном периоде повторения  $X(k)$ . Результат разложения приведён на Рис. 4.5.
- 2) Выбрать коэффициенты разложения  $X(k)$ , так что  $Re(X(k)) > \varepsilon$  и  $X(k-1) \leq X(k) \leq X(k+1)$ . Пороговое значение  $\varepsilon$  определяется точностью устройств.
- 3) Рассчитать частоту колебаний  $F(k) = k/NT$ , где  $k$  – номер коэффициента, соответствующий выбранным  $X(k)$  и выбрать из них частоту  $Fd = \max F(k)$ .
- 4) Определить период дискретизации по частоте изменения данных:  $P_F = 1/F$ , где  $F$  – частота дискретизации, по теореме Котельникова  $F \geq 2Fd$ .

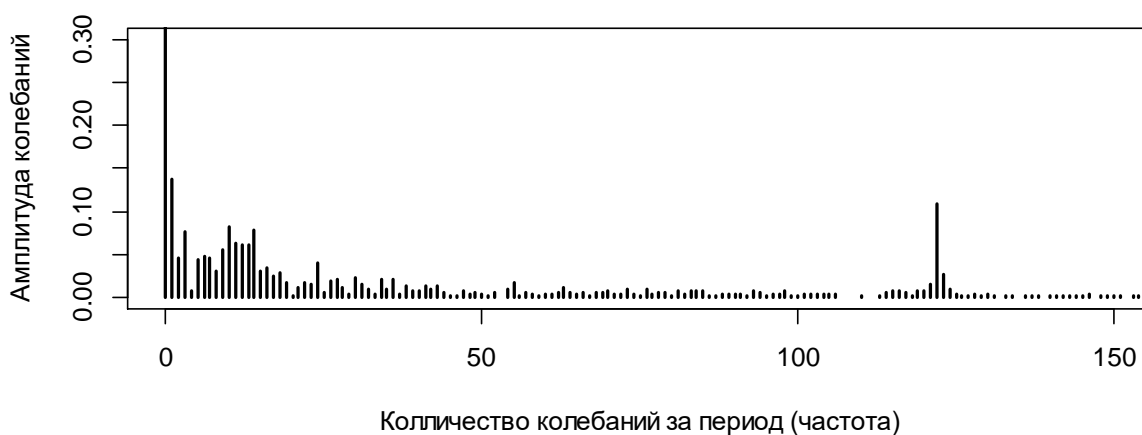


Рисунок 4.5 – Частотный спектр данных

5) Определить период дискретизации по скорости протекания процессов  $PS = X_{\min} / S_{\max}$ , где  $X_{\min}$  – допустимое изменение показателя в единицу времени, – скорость изменения данных, рассчитанная на основе собранной статистики при имитации критических состояний.

6) Выбрать период дискретизации для работы устройств IoT  $P = \min (PS, PF)$ , который отражает информацию о частоте изменения наблюдаемых событий.

7) Сформировать настройки брокера для выдачи данных: период рассылки  $P$  и обработчик поля «Качество обслуживания» (Quality of Service – QoS), для контроля критических ситуаций:

- стандартные рассылки в поле QoS помечаются флагом «0 – at most once»;
- сообщения, выходящие за границы диапазонов безопасного функционирования систем в QoS с флагом «2 – exactly once».

На основе проведенного исследования сделан вывод о том, что на облачных сервисах, выполняющих мониторинг данных, для снижения загрузки каналов связи без потери информативности и сохранении адекватности представления протекающих процессов достаточно выполнять рассылку данных с периодом  $P$  и условием QoS.

Апробация и внедрение метода адаптивных рассылок позволила сократить количество передаваемых данных и периодов активности устройств IoT в пять раз. Метод выполняется для каждого показателя, имеющего собственные спектральные характеристики. Перенастройка брокера для выдачи данных от устройств IoT происходит при изменении параметров физического состояния окружения, что позволяет оперативно соотносить время наблюдений со временем протекания реальных процессов.

#### **4 4 Исследование безопасности сервисов корпоративной сети Красноярского научного центра (ФИЦ КНЦ СО РАН)**

Проведено исследование безопасности сервисов корпоративной сети Красноярского научного центра (ФИЦ КНЦ СО РАН). Целью работы является анализ безопасности в динамике за последние два года, классификация кибератак по видам, выявление зависимостей между различными параметрами атак. В отличие от существующих работ, анализ выполняется на длительных временных интервалах, что позволяет выявить динамику поведения веб-сервисов по часам, дням, месяцам и годам.

В настоящее время активно используются веб-технологии для организации корпоративных облачных сервисов различного уровня (электронная почта, облачные диски, видеоконференции). Облачные сервисы используют технологии WWW и функционируют в открытой сети Интернет, поэтому они подвержены повышенным рискам информационной безопасности. Одним из важных компонентов по обеспечению безопасности является анализ различных журналов активности. В частности, представляет интерес журналы веб-серверов `nginx` и `apache`, анализ которых позволяет выявить кибератаки, совершаемые на облачные системы. В веб-системах объемы журналов могут иметь значительные размеры, что затрудняет их анализ в ручном режиме; в этом случае необходимо использовать автоматизированные инструменты для обработки и анализа данных.

При анализе журналов используются разные подходы. Один из самых популярных методов является сигнатурный анализ. При этом из элементов журнала могут извлекаться дополнительные параметры и характеристики, которые могут быть использованы для последующего анализа, например, кластеризации и обнаружения аномалий. Как правило,



сбои в процессе кибератак порождают записи журнала, которые отличаются от записей, представляющих штатное поведение системы. Поэтому целесообразно обращать внимание на отдельные записи журнала, которые не вписываются в общую картину. При кластеризации такие записи идентифицируются высокой степенью несходства со всеми существующими кластерами или не соответствуют ни каким сигнатурам. Неблагоприятные события системы проявляются не только в виде отдельных аномальных записей журналов, но и в виде динамических или последовательных аномалий. Поэтому необходимы подходы, которые позволяют группировать последовательности записей или выявлять временные закономерности и корреляции.

Для обработки данных разработан программный комплекс автоматизированной классификации кибератак на облачные веб-сервисы на основе сигнатурного метода. Проведенный анализ общего числа атак за 2020-2022 годы показывает, что в среднем их ежедневное количество меняется в небольших пределах: 3664 в 2020, 3481 в 2021 году и 3698 в 2022 году (3% отклонения по годам). Вместе с тем максимальное число атак изменяется в широких пределах от 8500 до 21000 за день, что свидетельствует об одновременном функционировании нескольких нескоординированных источников. На Рис. 4.6 представлена общая динамика обнаруженных атак по месяцам за 2020-2022 годы. Зафиксировано отсутствие выраженной периодичности как при ежедневном, так и при ежемесячном суммировании.

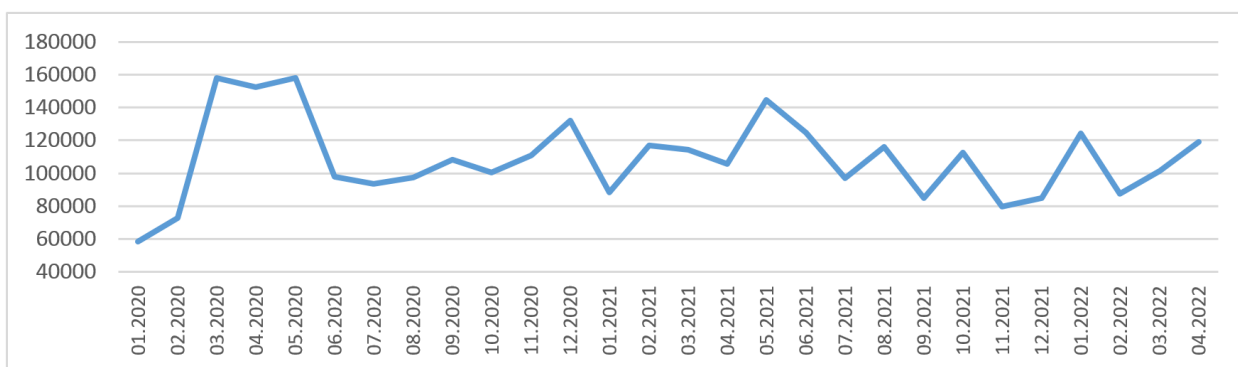


Рисунок 4.6 – Общая динамика атак по месяцам

При переходе к детальному анализу атак выделяются явные тренды на увеличение количества атак отдельных типов. На Рис. 4.7 представлено ежемесячное количество атак типов POLICY/EXT\_REST (запрещенное расширение) и WEB/FILE\_INJ (инъекция файла), на примере которых отчетливо видно увеличение интенсивности в 2 раза и более, незаметное при анализе общей динамики.

При анализе агрегированных атак по дням, обнаруженный ранее восходящий тренд не выявлен. Таким образом, наиболее эффективным способом обнаружения повышения риска киберугроз является анализ классифицированных угроз с агрегацией до месяца.

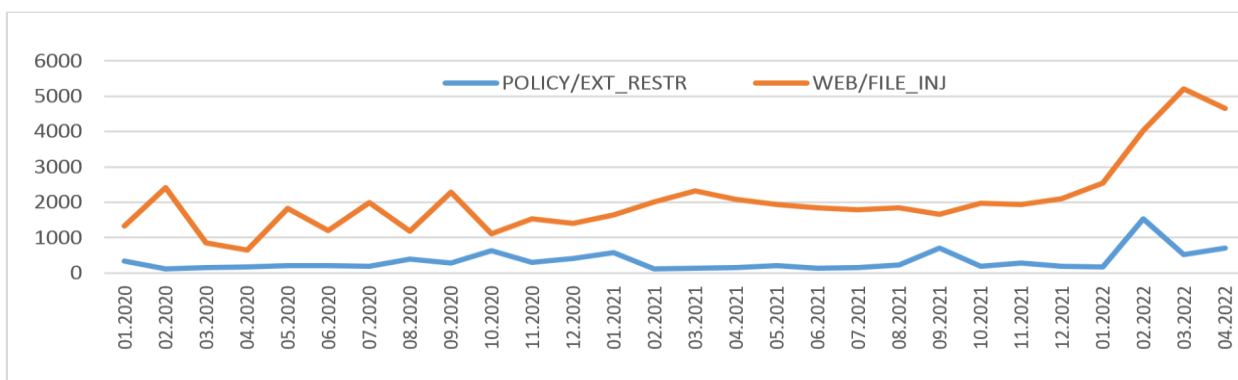
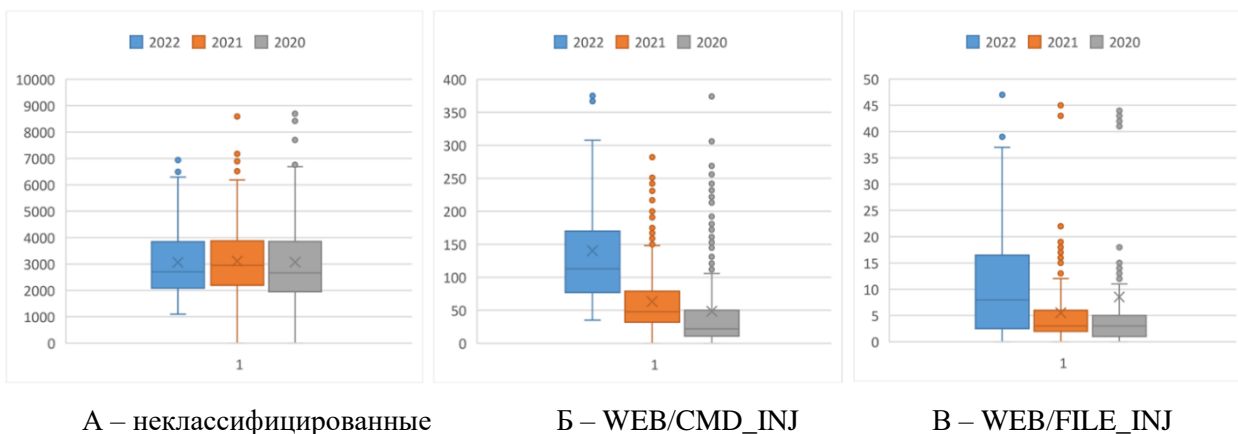


Рисунок 4.7 – Динамика классифицированных атак по месяцам

На рис. 4.8 приведены диаграммы размаха для распределений за 2020-2022 годы: (А) неклассифицированные атаки, (Б) атаки типа WEB/CMD\_INJ (инъекция команд) и (В) атаки типа WEB/FILE\_INJ (инъекция файла). Если при неклассифицированных атаках распределение имеет схожие параметры за разные годы, то в случае применения классификации параметры распределения изменяются существенно, особенно для 2022 года, который характеризуется увеличением риска киберугроз. Для выбора набора показателей была построена матрица корреляций их распределений по дням за весь наблюдаемый период 2020-2022 годы (рис. 4.9). Большинство показателей имеют слабую корреляцию, за исключением WEB/CMD\_INJ (инъекция команд), WEB/DIR\_TRAVERSAL (просмотр директории) и WEB/JAVA\_INJ (инъекция кода Ява), которые можно агрегировать.



А – неклассифицированные

Б – WEB/CMD\_INJ

В – WEB/FILE\_INJ

Рисунок 4.8 – Диаграммы размаха атак по годам

На основе проведенного анализа предложен метод для оценки изменения рисков киберугроз, основанный на сравнении параметров распределений выборок. Для выборок  $V_1$  и  $V_2$ , содержащих  $N$  независимых показателей введена следующая функция  $R$  оценки изменения рисков:

$$R(V_1, V_2) = \frac{1}{N} * \sum_{i=1}^N K_i,$$

$$\text{где } K_i = \begin{cases} 1, & \text{если } \mu_i > 0,6745 * \sigma_i, \\ 0, & \text{если } -0,6745 * \sigma_i \leq \mu_i \leq 0,6745 * \sigma_i, \\ -1, & \text{если } \mu_i < -0,6745 * \sigma_i, \end{cases}$$

$\mu_i$  – среднее значение выборки  $i$ -го признака выборки  $V_2$ ,

$\sigma_i$  – среднее квадратическое отклонение выборки  $i$ -го признака выборки  $V_1$ .

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. AUTO/CRAWLER		0,21	0,44	0,38	-0,34	0,10	-0,09	-0,10	-0,14	0,24	0,00	0,20	0,25	0,31	-0,20
2. AUTO/SCRIPT	0,21		0,10	-0,02	-0,06	-0,01	-0,15	-0,18	-0,34	0,13	-0,05	0,66	-0,34	0,28	-0,11
3. AUTO/SEC_SCAN	0,44	0,10		0,04	-0,24	-0,06	-0,30	-0,25	-0,29	-0,04	-0,22	0,09	0,17	0,29	-0,19
4. OTHERS	0,38	-0,02	0,04		-0,18	0,63	-0,08	-0,06	-0,09	0,23	-0,01	0,09	0,06	0,52	0,01
5. POLICY/EXT_RESTR	-0,34	-0,06	-0,24	-0,18		-0,09	0,55	0,60	0,46	0,12	0,50	-0,16	-0,09	-0,17	0,82
6. PROTOCOL	0,10	-0,01	-0,06	0,63	-0,09		0,20	0,18	0,06	0,09	0,28	0,41	-0,26	0,64	0,01
7. WEB/CMD_INJ	-0,09	-0,15	-0,30	-0,08	0,55	0,20		0,99	0,73	0,04	0,97	0,11	-0,11	-0,10	0,47
8. WEB/DIR_TRAVERSA	-0,10	-0,18	-0,25	-0,06	0,60	0,18	0,99		0,75	0,07	0,96	0,05	-0,06	-0,15	0,52
9. WEB/FILE_INJ	-0,14	-0,34	-0,29	-0,09	0,46	0,06	0,73	0,75		0,31	0,62	-0,11	0,07	-0,40	0,59
10. WEB/HEADER_INJ	0,24	0,13	-0,04	0,23	0,12	0,09	0,04	0,07	0,31		-0,01	-0,02	0,28	0,08	0,15
11. WEB/JAVA_INJ	0,00	-0,05	-0,22	-0,01	0,50	0,28	0,97	0,96	0,62	-0,01		0,22	-0,13	0,00	0,42
12. WEB/PHP_INJ	0,20	0,66	0,09	0,09	-0,16	0,41	0,11	0,05	-0,11	-0,02	0,22		-0,27	0,41	-0,13
13. WEB/RFI	0,25	-0,34	0,17	0,06	-0,09	-0,26	-0,11	-0,06	0,07	0,28	-0,13	-0,27		-0,26	-0,11
14. WEB/SQL_INJ	0,31	0,28	0,29	0,52	-0,17	0,64	-0,10	-0,15	-0,40	0,08	0,00	0,41	-0,26		-0,11
15. WEB/XSS	-0,20	-0,11	-0,19	0,01	0,82	0,01	0,47	0,52	0,59	0,15	0,42	-0,13	-0,11	-0,11	

Рисунок 4.9 – Матрица корреляций распределений классифицированных атак

Согласно предложенному методу, если среднее значение всех  $N$  признаков выборки  $V_2$  больше третьего квартиля выборки  $V_1$ , то значение изменения риска равно 1, которое можно интерпретировать как существенное увеличение риска по всем показателям. Если среднее значение всех  $N$  признаков выборки  $V_2$  меньше первого квартиля выборки  $V_1$ , то значение изменения риска равно  $-1$  (уменьшение риска по всем показателям). Значение  $R(V_1, V_2) \in [-1, 1]$  позволяет использовать этот показатель для анализа с помощью методов искусственного интеллекта, основанных на использовании ненадёжных данных. Метод не зависит от сравниваемых временных интервалов и объема выборок, поскольку основан на статистических показателях. Он может использоваться в других областях, в которых существует классификация показателей.

На основе расширенной ролевой модели разграничения доступа предложена новая модель безопасности, которая учитывает особенности функционирования облачных веб-

сервисов. Схема модели представлена на Рис 4.10. Модель описывается следующим образом:  $\langle U, R, P, S, Tk, C, D, Rq, RqM, RqP, UA(U), UC(U), PA(R), RqA(P), DA(P), RqMA(P), RqPA(P), cloud(D), user(S), roles(S), token(Tk), requests(S), domains(Rq), methods(Rq), params(Rq) \rangle$ , где заданы следующие множества:  $U$  – пользователи,  $R$  – роли,  $P$  – права доступа,  $S$  – сессии пользователей,  $Tk$  – токены пользователей,  $C$  – облачные системы,  $D$  – доменные имена,  $Rq$  – запросы,  $RqM$  – методы запросов,  $RqP$  – параметры запросов. Заданы также различные функции для связи элементов модели. Полученная модель безопасности позволяет разграничивать доступ с использованием пути запроса (URI), параметров запроса, методов запроса, доменов с привязкой к конкретной облачной системе.

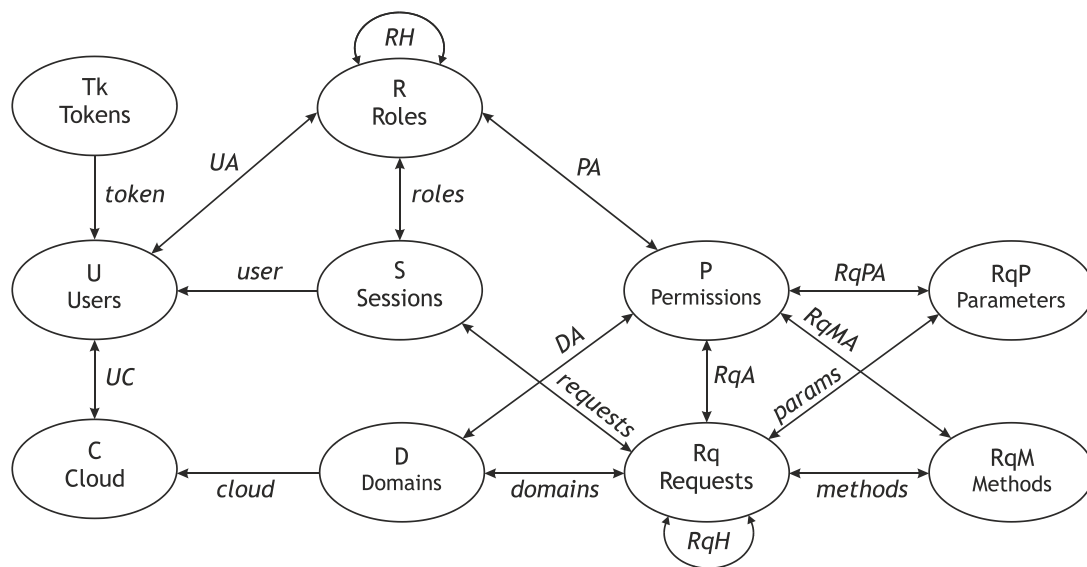


Рисунок 4.10 – Модель безопасности систем облачных вычислений

На основе собранных данных об использовании веб-сервисов сотрудниками ФИЦ КНЦ СО РАН предложен метод *кластеризации посещаемых интернет-ресурсов на группы со схожей тематикой*. В качестве предобработки выполняется выделение и установление связей между посещаемыми интернет-ресурсами через анализ их совместной встречаемости в пределах сеансов работы (сессий). Под сессией подразумевается совокупность интернет-ресурсов, посещенных за некоторый промежуток времени. В самом простом случае, сессией можно считать одни сутки, однако для повышения качества работы метода рассмотрены и другие варианты выделения сессий. Для анализа совместной встречаемости ресурсов в пределах одной сессии используется вероятностное тематическое моделирование. Основная цель предложенного метода в разделении ресурсов на группы с целью снижения объема анализируемых данных через исключение «безопасных» групп ресурсов. Анализ наиболее популярных интернет-ресурсов в каждой группе позволяет определить тему каждой группы, и решить, является

ли группа «безопасной». В табл. 4.1 представлен результат моделирования 5-ти групп. Чем выше ресурс расположен в группе, тем сильнее его принадлежность к этой группе.

Таблица 4.1. Результат моделирования данных за февраль 2020 г. на 5 групп.

1	2	3	4	5
newslab.ru	nowa.cc	update.eset.com	apps.webofknowledge.com	fitohobby.ru
4pda.ru	ugadalki.ru	law-college-sfu.ru	packages.linuxmint.com	ib.adnxs.com
sfkras.ru	scask.ru	kinoaction.ru	http.debian.net	allrefs.net
edu.sfu-kras.ru	forum.rcmir.com	kiwt.ru	urod.ru	ckp-rf.ru
worldcrisis.ru	2baksa.net	dostavka-krasnoyarsk.ru	fips.ru	teamodels.no
libgen.is	autoopt.ru	kinoaction.ru	mc.corel.com	profinance.ru

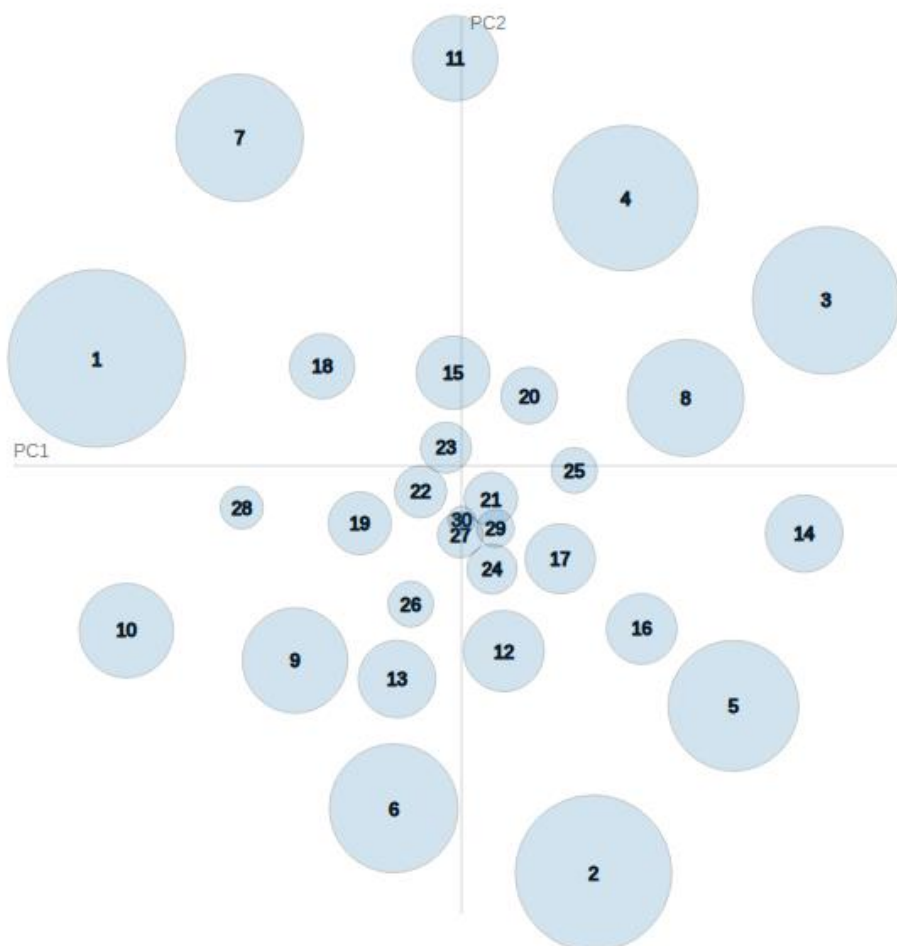


Рисунок 4.11 – Проекция 30-ти групп, полученных путем моделирования данных за февраль 2020 года, на две главные компоненты

Предложенный метод обладает большим количеством параметров, позволяющим точно настроить его под разные источники данных: от небольшой внутренней сети до множества высоконагруженных узлов в облачной архитектуре.

Группировка сервисов по схожей теме может быть использована в различных задачах, для определения:

- интересов пользователя;
- безопасных и опасных сайтов для снижения числа анализируемых данных;
- тематики интернет-ресурса.

## ЗАКЛЮЧЕНИЕ

В разделе 1.1 представлены модели электрических полей и токов Глобальной электрической цепи Земли. Построены глобальные карты грозовой активности на основе данных наземной Всемирной сети определения местоположения молний (World Wide Lightning Location Network, WWLLN) за 2007-2009 годы. Построенные карты использованы для численного моделирования Глобальной электрической цепи. Физико-математический подход к моделированию электрического потенциала, который управляет токами в ионосфере, был описан в предыдущих статьях исполнителей проекта. Использовано двумерное приближение при моделировании ионосферного проводника. Основной прогресс по сравнению с предыдущими версиями модели достигнут за счет использования глобального распределения гроз. Показано полученное в результате численного решения задачи электропроводности распределение электрического потенциала в ионосфере на высоте 120 км в типичных условиях июля.

В разделе 1.2 представлена математическая модель диссипативных процессов, вызывающих генерацию электрического поля на границе и в хвосте магнитосферы. В качестве основной причины неоднородного аномального удельного сопротивления магнитосферы рассматривается ионно-акустическая турбулентность, возникающая в тонких токовых слоях, когда плотность электрического тока превышает критический порог. Исследовано влияние неоднородной проводимости и параметра Холла на процесс магнитного пересоединения при различных начальных магнитных конфигурациях. Волны Холла-Альфвена, распространяющиеся быстрее, чем обычные МГД-волны, могут служить предвестниками импульсного магнитного пересоединения.

В разделе 1.3 представлены алгоритмы, реализующие длительное движение спутника по возмущенной орбите с выбором методов интегрирования уравнений движения. Линейные многошаговые методы, несмотря на конечный интервал устойчивости, являются конкурентноспособными по сравнению с неявными методами Рунге-Кутты (интегратором Эверхарта и коллокационными) при интегрировании орбит на небольшом промежутке времени. Для почти круговых орбит (какими являются орбиты спутников ГЛОНАСС) особый интерес представляют симметричные линейные многошаговые методы. В текущем году разработаны приемы построения и исследования устойчивости (периодичности) симметричных методов и получаемых на их основе схем предиктор-корректор. Предложена технология выбора пар предиктор-корректор из широкого семейства симметричных методов высокого порядка.

В разделе 2.1 выяснены условия максимального усиления локального поля за счет настройки поверхностного решеточного резонанса путем варьирования геометрических параметров периодических структур, привлекающих внимание в связи с проявлением в них нетривиальных оптических свойств и возможностью широкого применения. При решении задач рассеяния периодическими структурами наночастиц использовался метод конечных разностей во временной области (FDTD) для расчета спектров экстинкции, отражения и поглощения в наноструктурах. Отмечено, что положение максимумов экстинкции связано со спектральным положением аномалий Рэлея для взаимно ортогональных направлений в плоскости решетки: коротковолновый пик связан с периодом решетки в плоскости поляризации падающего излучения, а длинноволновый пик – в перпендикулярном направлении. Изменяя шаг решетки в одном из направлений, можно добиться совпадения длин волн всех четырех мультипольных компонент.

В разделе 2.2 представлены математические модели ионного транспорта в наноструктурированных мембранных материалах под действием электрического поля. В рамках сотрудничества с коллегами из Европейского центра передовых технологий водоподготовки (Wetsus) и Университета Твенте (Нидерланды) разработана модель транспорта ионов через полимерную мембрану обратного осмоса. Такие мембраны являются композитными и состоят из полиэферной подложки, промежуточного слоя из полисульфона и селективного слоя из полиамида. Расчетные данные сравнивались с данными экспериментов, которые проводились в Университете Твенте. Модель позволяет предсказывать эффективность опреснения, минеральный состав полученной воды и может использоваться для разработки улучшенных фильтров и мембранных элементов.

В разделе 3.1 представлены результаты популяционно-эволюционного анализа геномов вирусов семейства Coronaviridae. Проведено сравнение внутреннего структурирования 69 геномов коронавирусов, выявленных разными методами сравнения: 1) традиционное выравнивание с последующим построением филогении, 2) неконтролируемая классификация на основе  $k$ -средних и 3) современный метод нелинейной статистики на основе упругих карт. Одновременное применение нескольких подходов к одному набору данных может привести к обнаружению деталей, которые иначе остались бы незамеченными. Комбинированный подход позволяет предположить, что в случае эволюционно далеких геномов, но принадлежащих к одному и тому же кластеру на эластичной карте, некоторые из них, вероятно, скоро изменят состояние этого признака. В случае группы заболеваний это означает, что такой вирус может быть «готов» к изменению типа вызываемого заболевания и/или к смене хозяина.



В разделе 3.2 предложен модифицированный непараметрический алгоритм автоматической классификации статистических данных большого объёма, основанный на их сжатии путём декомпозиции многомерного пространства признаков исследуемых объектов, совершенствовании и алгоритмизации традиционной непараметрической процедуры автоматической классификации. Его применение позволяет обнаружить классы, соответствующие одномодальным фрагментам плотности вероятности многомерной случайной величины. Предлагаемый метод позволяет в 2 раза сократить время классификации при использовании многопоточной технологии обработки данных, что подтверждается результатами обработки спектральных данных дистанционного зондирования лесных массивов, повреждённых сибирским шелкопрядом. Топографическая нормализация исходного снимка позволяет повысить эффективность обнаружения состояний повреждённых лесных древостоев.

В разделе 3.3 разработана метамодель детализации интегральных оценок для определения причин состояния природно-техногенной безопасности территорий, которая является основой информационной поддержки по планированию мероприятий долгосрочного развития территорий и формированию стратегий снижения рисков чрезвычайных ситуаций. Развитие метода интегрального аналитического оценивания состояния территорий обеспечивает возможность детализации комплексного показателя природно-техногенной безопасности до причин, которые могут обуславливать его выход за пределы нормативных значений. Выполнен анализ связей «причина отклонения от нормы – рекомендуемые превентивные мероприятия». Это позволяет завершить цикл получения и обработки данных комплексного мониторинга до формирования управляющих воздействий с учётом особенностей территорий.

В разделе 4.1 представлены методические и программные средства поддержки оперативной аналитической обработки данных научных исследований в рамках модельно-ориентированных систем сбора данных. Реализованный в платформе модельно-ориентированный подход применен к генерации аналитических отчетов. Апробация выполнена на примере оперативной сводки, формируемой Центром обеспечения реализации полномочий в областях гражданской обороны и чрезвычайных ситуаций Красноярского края (Центра ГО и ЧС).

В разделе 4.2 представлено программное обеспечение для трансформации данных различных форматов путем сопоставления структурных элементов и выполнения преобразований по установленным взаимным отношениям. Программное обеспечение

применяется к построению аналитических моделей и для графической визуализации результатов аналитической обработки данных различных форматов.

В разделе 4.3 исследованы аспекты кибербезопасности систем облачных вычислений для различных способов обслуживания в области облачных вычислений для модели обслуживания SaaS (Software-as-a-Service). На основе реализованного технологического стенда инфраструктуры Интернета вещей (IoT), налажены сбор и обработка данных взаимодействия различных сервисов. Отмечено, что взаимодействие через сетевые инфраструктуры пространственно-распределённых узлов IoT на основе модели облачных вычислений может вызывать существенные нагрузки на сети и потребителей информации. Для обеспечения безопасной работы облачных сервисов и снижения интенсивности сетевого трафика разработан метод адаптивных рассылок, позволяющий выполнять синхронизацию режимов информационного взаимодействия между источниками данных и средствами мониторинга на основании частотных характеристик наблюдаемых процессов.

В разделе 4.4 на основе данных о взаимодействии пользователей и оборудования ФИЦ КНЦ СО РАН с различными интернет-сервисами (цифрового следа), предложены методы анализа и выявления рисков киберугроз. Рассмотрена динамика атак по видам на различных временных интервалах, выделены основные тренды и особенности киберугроз; показано преимущество классификации атак. Для сравнения киберугроз предложен метод количественной оценки изменения риска, основанный на сопоставлении параметров распределений выборок. На основе расширенной ролевой модели безопасности создана новая формальная модель безопасности, которая позволяет гибко разграничивать доступ с использованием элементов, специфичных для облачных веб-сервисов, и может быть использована в сложных гетерогенных системах. А на основе собранных данных об использовании веб-сервисов предложен метод *кластеризации посещаемых интернет-ресурсов на группы со схожей тематикой*, который обладает большим количеством параметров, позволяющих точно настроить его под разные источники данных: от небольшой внутренней сети до множества высоконагруженных узлов в облачной архитектуре. Группировка сервисов по схожей теме может быть использована в задачах для определения интересов пользователей, безопасных и опасных сайтов для снижения числа анализируемых данных, тематики интернет-ресурса.

## ПРИЛОЖЕНИЕ А

### Научные публикации, опубликованные в рамках выполнения проекта, в журналах, индексируемых в российских и международных информационно-аналитических системах научного цитирования

1. Denisenko V.V., Rycroft M.J. WWLLN Data Used to Model the Global Ionospheric Electric Field Generated by Thunderstorms // *Annals of geophysics*. – 2022. – Vol. 65. – Art. PA536. [Q3, web of science] DOI: 10.4401/ag-8821.  
<https://www.annalsofgeophysics.eu/index.php/annals/article/view/8821>
2. Erkaev N.V., Semenov V.S. About Petschek-Type Reconnection Driven by Inhomogeneous Plasma Resistivity // *Springer Proceedings in Earth and Environmental Sciences*. – 2022. – Vol. Problems of Geocosmos. – 2020. – P. 397-412. [Scopus]  
[https://doi.org/10.1007/978-3-030-91467-7\\_29](https://doi.org/10.1007/978-3-030-91467-7_29)
3. Schmid D., Lammer H., Plaschke F., Vorburger A., Erkaev N.V., Wurz P., Narita Y., Volwerk M., Baumjohann W., Anderson B.J. (2022). Magnetic evidence for an extended hydrogen exosphere at Mercury // *J. Geophysical Research: Planets*. – 2022. – Vol. 127. – Art. e2022JE007462. [Q1, web of science]  
<https://doi.org/10.1029/2022JE0074625>.
4. Lammer H., Scherf M., Ito Y., Mura Ф., Vorburger A., Guenther E., Wurz P., Erkaev N.V., Odert P. The Exosphere as a Boundary: Origin and Evolution of Airless Bodies in the Inner Solar System and Beyond Including Planets with Silicate Atmospheres // *Space Sci Rev*. – 2022. – Vol. 218. – Art. 15. [Q1, web of science]  
<https://doi.org/10.1007/s11214-022-00876-5>
5. Gorbunova K.D., Erkaev N.V. Compact MacCormac-type schemes applied for atmospheric escape problem // *J. Siberian Federal University. Mathematics & Physics*. – 2022. – Vol. 15, № 4. – P. 500-509. [Q, web of science] DOI: 10.17516/1997-1397-2022-15-4-500-509
6. Karepova E.D., Shan'ko Yu.V., Adaev I.R. Predictor-Corrector Mode for Symmetric Multistep Methods for the Numerical Modelling of Satellite Motion // *AIP Conference Proceedings*. – 2022. – V. 2522. – Art. 100006. [Scopus]  
<https://doi.org/10.1063/5.0100843>
7. Kimani E.M., Pranić M., Porada S., Kemperman A.J.B., Ryzhkov I.I., van der Meer W.G.J., Biesheuvel P.M. The influence of feedwater pH on membrane charge ionization and ion

- rejection by reverse osmosis: An experimental and theoretical study // *J. Membrane Science*. – 2022. – V. 660. – Art. 120800. [Q1, web of science]  
<https://doi.org/10.1016/j.memsci.2022.120800>
8. Jia T., Li D., Du J., Fang X., Gerasimov V., Ågren H., Chen, G. A bimodal type of AgPd Plasmonic Blackbody Nanozyme with boosted catalytic efficacy and synergized photothermal therapy for efficacious tumor treatment in the second biological window // *J. Nanobiotechnology*. – 2022. – Vol. 20, Issue 1. – Art. 424. [Q1, web of science]  
<https://doi.org/10.1186/s12951-022-01627-y>
9. Пенькова Т.Г., Метус А.М., Ноженкова Л.Ф., Морозов Р.В. Метод расчета интегральной оценки качества жизни муниципальных образований региона в разрезе реализации национальных проектов // *Современные наукоемкие технологии* – 2022 – Iss.7 – P.79-87  
<https://doi.org/10.17513/snt.39237>
10. Пенькова Т.Г., Ноженкова Л.Ф. Исследование территориальных кластеров макрорегиона на основе многомерного анализа показателей качества жизни (на примере Красноярского края) // *Информационные и математические технологии в науке и управлении* – 2022 – Iss.3(27) – P.15-24  
<https://doi.org/10.38028/ESI.2022.27.3.002>
11. Лапко А.В., Лапко В.А., Бахтина А.В. Быстрый выбор коэффициентов размытости ядерной оценки плотности вероятности для семейства одномерных логнормальных законов распределения // *Информатика и системы управления* – 2022 – Vol.71, Iss.1 – P.90-100  
[https://doi.org/10.22250/18142400\\_2022\\_71\\_1\\_90](https://doi.org/10.22250/18142400_2022_71_1_90)
12. Молявко А.А., Каропова Е.Д., Садовский М.Г. Сравнение последовательностей символов методом Шайдурова с применением технологии MPI для распараллеливания // *Информатизация и связь* – 2022 – Vol.4 – P.80-84  
<https://doi.org/10.34219/2078-8320-2022-13-4-80-84>
13. Kirichenko A.D., Poroshina A.A., Sherbakov D.Y., Sadovsky M.G., Krutovsky K.V. Comparative analysis of alignment-free genome clustering and whole genome alignment-based phylogenomic relationship of coronaviruses // *PLoS ONE*. – 2022. – Vol. 17(3): Art. e0264640. [Q1, web of science]  
<https://doi.org/10.1371/journal.pone.0264640>

14. Kalach A.V., Nicheporchuk V.V., Arifullin E.Z., Akulov A.Yu., Kulikova T.N. Systematization of the parameters of monitoring for situational modeling of response to emergency situations // *Journal of Physics: Conference Series*. – 2022. – Vol. 1902. – Art. 12072. [web of science] DOI: 10.1088/1742-6596/1902/1/012072
15. Исаева О.С., Исаев С. В., Кулясов Н. В. Формирование адаптивных рассылок брокера данных интернета вещей // *Информационно-управляющие системы*. – 2022. – № 5 (120). – С. 23-31. [ядро РИНЦ] DOI: 10.31799/1684-8853-2022-5-23-31.
16. Isaeva O.S., Kulyasov N.V. Isaev S.V. Creation of a simulation stand for studying of the internet of things' technologies // *AIP Conference Proceedings*. – 2022. – No. 2647. – Art. 040030. [Scopus] DOI: 10.1063/5.0104342.
17. A.V. Vyatkin, V.V. Shaydurov, and E.V. Kuchunova Two finite volume schemes for advection equation // *Lecture Notes in Computer Science and Engineering* – 2022 – Vol.141. – P.545-561. [Scopus] DOI: 10.1007/978-3-030-87809-2\_41
18. Лапко А.В., Лапко В.А., Бахтина А.В. Сравнение методики проверки гипотезы о независимости двумерных случайных величин, основанной на непараметрическом классификаторе // *Искусственный интеллект и принятие решений*. – 2022. – №1. – С. 45 – 56. [ядро РИНЦ] DOI: 10.14357/20718594220105
19. Лапко А.В., Лапко В.А. Быстрый выбор коэффициентов размытости ядерной оценки плотности вероятности независимых случайных величин // *Автометрия*. – 2022. – Т. 58, № 1. – С. 33-39. [ядро РИНЦ] DOI: 10.15372/AUT20220104
20. Лапко А.В., Лапко В.А. Оценивание нелинейного функционала от плотности вероятности трёхмерной случайной величины для повышения вычислительной эффективности непараметрических решающих правил // *Автометрия*. – 2022. – Т. 58, № 2. – С. 93-103. [ядро РИНЦ] DOI: 10.15372/AUT20220211
21. Сенашова М.Ю. Пространственная структура генов фотосинтетической системы хлоропластов водорослей с точки зрения биоинформатики // *Информационные и математические технологии в науке и управлении* – 2022 – Vol.3, Iss.27 – P.34-39  
<https://doi.org/10.38028/ESI.2022.27.3.004>
22. Исаев С. В., Кононов Д. Д. Исследование динамики и классификация атак на веб-сервисы корпоративной сети // *Сибирский аэрокосмический журнал* – 2022 – Vol.23, Iss.4 – P.593-601  
<https://doi.org/10.31772/2712-8970-2022-23-4-593-601>
23. Vasil'ev E.N. Calculation and optimization of heat exchangers for thermoelectric cooling unit // *Thermophysics and Aeromechanics*. – 2022. – Vol. 29(3). – P. 401–410. [Q4, web of science] DOI: 10.1134/S0869864322030088

24. Shaidurov V., Novikov A. Difference schemes for second-order ordinary differential equations with corrector and predictor properties // Russian J. Numerical Analysis and Mathematical Modelling. – 2022. – Vol. 37, Iss. 3. – P. 175-187 [Q3, web of science] DOI: 10.1515/rnam-2022-0015
25. Petrakova V., Tsipotan A., Slabko V. The mathematical model of quantum dots pair orientation under laser radiation field // Computer optics – 2022. – Vol. 46, Iss. 4. – P. 555-560 [Q, web of science] DOI: 10.18287/2412-6179-CO-1062
26. Захарова О.С., Брацун Д.А., Рыжков И.И. Конвективная неустойчивость в многокомпонентных смесях с эффектом Соре // Вычислительная механика сплошных сред. – 2022. – Т. 15, Вып. 1. – С. 67-82 [ядро РИНЦ] DOI: 10.7242/1999-6691/2022.15.1.6
27. Vasil'ev E.N. Numerical Simulation of Temperature and Thermal Stress Fields in a Carbon Block under External Thermal Effect // J. Siberian Federal University. Mathematics & Physics. – 2022. – Vol. 15, Iss. 3. – P. 267-272 [Q, web of science] DOI: 10.17516/1997-1397-2022-15-3-267-272  
<https://elib.sfu-kras.ru/handle/2311/145513>
28. Vasil'ev E.N. Modeling of the Ice Cover Dynamics of a Freshwater Reservoir // Journal of Siberian Federal University. Mathematics & Physics. – 2022. – Vol. 15, Iss. 6. – P. 753-762 [Q, web of science] DOI: 10.17516/1997-1397-2022-15-6-753-762  
<https://elibrary.ru/item.asp?id=49833998>
29. Васильев Е.Н., Емельянов Д.П., Нестеров Д.А. Экспериментальное исследование инееобразования в пластинчатом теплообменнике // Журнал Сибирского федерального университета. Техника и технологии. – 2022. – Т. 15, Вып. 1. – С. 24-34 [ядро РИНЦ] DOI: 10.17516/1999-494X-0371  
<https://elibrary.ru/item.asp?id=48184458>
30. Васильев Е.Н., Емельянов Д.П. Влияние угла наклона щелевого канала на стекание воды // Журнал Сибирского федерального университета. Техника и технологии. – 2022. – Т. 15, Вып. 5. – С. 526-533 [ядро РИНЦ] DOI: 10.17516/1999-494X-0413.  
<https://elibrary.ru/item.asp?id=49371587>
31. N. V. Erkaev, C. Weber, J.-M., Griemeier, H. Lammer, V.A. Ivanov, and P. Odert Can Radio Emission Escape from the Magnetosphere of Andromedae b - A New Method to Constrain the Minimum Mass of Hot Jupiters // MNRAS. – 2022. – Vol. 512, Iss. 4, – P. 4869-4876 [Q1, web of science] DOI: 10.1093/mnras/stac767

32. Isaeva O.S., Kulyasov N.V. Isaev S.V. Infographic method to support testing of spacecraft onboard systems based on knowledge base // Vestnik Tomskogo gosudarstvennogo universiteta-Upravlenie vychislitel'naja tehnika i informatika-Tomsk state university journal of control and computer science. – 2021. – Vol. 57, – P. 22-33 [Q, web of science] DOI: 10.17223/19988605/57/3  
<https://elibrary.ru/item.asp?id=48008259>
33. Krasnov I.V. Opto-thermophoretic superlattice for atoms // Laser Physics Letters. – 2022. – Vol. 19, Iss. 9. – Art. 95501 [Q3, web of science] DOI: 10.1088/1612-202X/ac81bc

## ПРИЛОЖЕНИЕ Б

### Научные публикации, опубликованные вне рамок выполнения проекта

1. Maksimov D.N., Gerasimov V.S., Bogdanov A.A., Polyutov S.P. Enhanced sensitivity of an all-dielectric refractive index sensor with an optical bound state in the continuum // *Physical Review A*. – 2022. – Vol. 105, Issue 3. – Art. 033518. [Q2, web of science]  
<https://doi.org/10.1103/physreva.105.033518>
2. Kostyukov A.S., Gerasimov V.S., Ershov A.E., Bulgakov E.N. Ring of bound states in the continuum in the reciprocal space of a monolayer of high-contrast dielectric spheres // *Physical Review B*. – 2022. – Vol. 105, Issue 7. – Art. 075404. [Q2, web of science]  
<https://doi.org/10.1103/physrevb.105.075404>
3. Kostyukov A.S., Isaev I.L., Ershov A.E., Gerasimov V.S., Polyutov S.P., Karpov S.V. Part I. Nanobubbles in pulsed laser fields for anticancer therapy: in search of adequate models and simulation approaches // *J. Physics D: Applied Physics*. – 2022. – Vol. 55, Issue 17. – Art. 175401. [Q2, web of science]  
<https://doi.org/10.1088/1361-6463/ac4c20>
4. Kostyukov A.S., Isaev I.L., Ershov A.E., Gerasimov V.S., Polyutov S.P., Karpov S.V. Part II. Nanobubbles around plasmonic nanoparticles in terms of modern simulation modeling: what makes them kill the malignant cells // *J. Physics D: Applied Physics*. – 2022. – Vol. 55, Issue 17. – Art. 175402. [Q2, web of science]  
<https://doi.org/10.1088/1361-6463/ac4c1f>
5. Ronzhin N.O., Posokhina E.D., Mikhлина E.V., Shestakov N.P., Nemtsev I.V., Bondar V.S., Ryzhkov I.I. The effect of copper and iron adsorption on the catalytic performance of alumina nanofiber/nanodiamond composite in sensing applications // *J. Physics and Chemistry of Solids*. – 2022. – V. 167. – Art. 110785. [Q2, web of science]  
<https://doi.org/10.1016/j.jpcs.2022.110785>
6. Fadeeva P., Pavlov M.V., Kharchenko I.A., Simunin M.M., Shabanova K.A., Pavlov V.F., Ryzhkov I.I. High strength ceramic substrates based on perlite and foam silicates for filtration membranes // *Membranes and membrane technologies*. – 2022. – Vol. 4 (3). – P. 170-176. [Q, web of science]  
<https://link.springer.com/article/10.1134/S2517751622030040>



## ПРИЛОЖЕНИЕ В

### Свидетельства о регистрации программ

1. Пенькова Т.Г., Метус А.М., Морозов Р.В., Ноженков А.И., Ничепорчук В.В. Модуль расчета рейтинговых оценок качества жизни территорий // Свидетельство о регистрации программы для ЭВМ № 2022665828 от 22.08.2022 г.
2. Лапко А.В., Лапко В.А., Им С.Т., Тубольцев В.П., Авдеенок В.Л. Программа автоматической классификации данных дистанционного зондирования Земли на основе непараметрических алгоритмов принятия решений (НАС v. 2.0) // Свидетельство о государственной регистрации программы для ЭВМ № 2022619023 от 18.05.2022.
3. Ноженкова Л.Ф., Пенькова Т.Г., Ноженков А.И., Метус А.М., Морозов Р.В. Веб-сервис анализа качества жизни территорий // Свидетельство о регистрации программ для ЭВМ № 2022666045 от 24.08.2022.