

Федеральное агентство научных организаций

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ
УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«КРАСНОЯРСКИЙ НАУЧНЫЙ ЦЕНТР СИБИРСКОГО ОТДЕЛЕНИЯ
РОССИЙСКОЙ АКАДЕМИИ НАУК»

Институт вычислительного моделирования СО РАН – обособленное
подразделение ФИЦ КНЦ СО РАН

УДК 007 + 004.9
N гос. регистрации 01201356264
Инв. № 0203/2016

УТВЕРЖДАЮ
Директор ФИЦ КНЦ СО РАН

_____ Шабанов В.Ф.
_____ 2017 г.

ОТЧЕТ О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ
БАЗОВЫЙ ПРОЕКТ IV.35.1.2 "МЕТОДЫ АНАЛИТИЧЕСКОЙ ОБРАБОТКИ
ДАННЫХ И ТЕХНОЛОГИИ СОЗДАНИЯ ИНТЕГРИРОВАННЫХ
ИНФОРМАЦИОННО-УПРАВЛЯЮЩИХ СИСТЕМ" ПРОГРАММЫ СО РАН
IV.35.1 "ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И ТЕХНОЛОГИИ СОЗДАНИЯ И
ПРИМЕНЕНИЯ ИНТЕГРИРОВАННЫХ ИНФОРМАЦИОННО-
ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ ДЛЯ РЕШЕНИЯ ЗАДАЧ ПОДДЕРЖКИ
ПРИНЯТИЯ РЕШЕНИЙ". № ГОС. РЕГИСТРАЦИИ 01201356264

(заключительный)

Номер проекта в ИСГЗ ФАНО 356-2014-0203

Приоритетное направление (номер и наименование) Информационно-телекоммуникационные системы

Программа ФНИ (номер и наименование) 35. Когнитивные системы и технологии, нейроинформатика и биоинформатика, системный анализ, искусственный интеллект, системы распознавания образов, принятие решений при многих критериях

Протокол Ученого совета _____
№ __ от «__» _____ 2017 г.

Руководитель проекта _____ Л.Ф. Ноженкова
д.т.н., профессор " " _____ 2017 г.

Красноярск, 2017

СПИСОК ИСПОЛНИТЕЛЕЙ

Руководитель темы д.т.н., профессор	_____	Л.Ф. Ноженкова (Введение, Разделы 2, 3, Заключение)
	(подпись, дата)	
Исполнители: зам. дир., к.т.н.	_____	С.В. Исаев (Раздел 4)
	(подпись, дата)	
Г.н.с., д.т.н.	_____	А.В. Лапко (Раздел 5)
	(подпись, дата)	
В.н.с., д.т.н.	_____	В.А. Лапко (Раздел 5)
	(подпись, дата)	
В.н.с., д.ф.-м.н.	_____	М.Г. Садовский (Раздел 6)
	(подпись, дата)	
С.н.с., к.т.н.	_____	А.А. Евсюков (Раздел 3)
	(подпись, дата)	
С.н.с., к.т.н.	_____	О.С. Исаева (Раздел 2)
	(подпись, дата)	
С.н.с., к.т.н.	_____	В.В. Ничепорчук (Раздел 3)
	(подпись, дата)	
С.н.с., к.т.н.	_____	Т.Г. Пенькова (Раздел 3)
	(подпись, дата)	
С.н.с., к.ф.-м.н.	_____	М.Ю. Сенашова (Раздел 6)
	(подпись, дата)	
Н.с., к.т.н.	_____	Д.В. Жучков (Раздел 3)
	(подпись, дата)	
Н.с., к.т.н.	_____	Е.В. Ковязина (Раздел 3)
	(подпись, дата)	
Н.с.	_____	Д.Д. Кононов (Раздел 4)
	(подпись, дата)	
Н.с., к.т.н.	_____	А.В. Коробко (Разделы 3)
	(подпись, дата)	
Н.с., к.т.н.	_____	А.И. Ноженков (Раздел 3)
	(подпись, дата)	
М.н.с.	_____	А.И. Белорусов (Раздел 3)
	(подпись, дата)	

М.Н.С.	_____	Е.А. Грузенко (Раздел 2)
	<i>(подпись, дата)</i>	
М.Н.С.	_____	А.А. Коробко (Разделы 3)
	<i>(подпись, дата)</i>	
М.Н.С.	_____	А.А. Марков (Раздел 3)
	<i>(подпись, дата)</i>	
М.Н.С.	_____	В.В. Морозов (Раздел 3)
	<i>(подпись, дата)</i>	
прогр. 1-ой кат.	_____	С.Н. Кочетков (Раздел 3)
	<i>(подпись, дата)</i>	
прогр. 1-ой кат.	_____	И.А. Ларионова (Раздел 4)
	<i>(подпись, дата)</i>	
прогр. 1-ой кат.	_____	А.В. Малышев (Раздел 4)
	<i>(подпись, дата)</i>	
программист	_____	А.А. Кулясов (Раздел 4)
	<i>(подпись, дата)</i>	
инженер	_____	А.А. Сиротинин (Раздел 4)
	<i>(подпись, дата)</i>	
Нормоконтролер	_____	А.В. Вяткин
	<i>(подпись, дата)</i>	

РЕФЕРАТ

Отчёт 36 с., 13 рис., 1 табл., 3 прил.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ, КОНСОЛИДАЦИЯ ДАННЫХ, ОБРАБОТКА И АНАЛИЗ ДАННЫХ, ИНТЕГРИРОВАННЫЕ СИСТЕМЫ.

Объектом исследования являются технологии информационно-аналитической поддержки принятия решений.

Цель работы – создание новых методов обработки и анализа данных, интеллектуальных технологий имитационного моделирования и комплексной поддержки принятия решений как инструментов построения информационно-управляющих систем разного назначения.

Использованы оригинальные методы консолидации и оперативной аналитической обработки данных, методы интеллектуальной поддержки конструкторских задач, информационно-графического моделирования, методы анализа данных.

Основные результаты работы:

- Методы интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата.
- Модели консолидации данных гетерогенных источников в соответствии с форматами хранения.
- Модели и алгоритмы анализа структурных характеристик территорий.
- Алгоритмы построения управляющей модели для разработки модельно-ориентированного программного обеспечения.
- Технологии создания интегрированных систем территориального и корпоративного управления.
- Модели управления безопасностью в информационно-телекоммуникационных системах.
- Новые методы доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей.
- Анализ корреляции между кластерами, выделяемыми по статистическим и комбинаторным свойствам нуклеотидных последовательностей и соответствующими биологическими объектами.

Область применения – построение интегрированных информационных систем для комплексной поддержки принятия решений в разных прикладных областях.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	6
1 Важнейший результат.....	8
2 Разработка методов интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата	10
3 Разработка моделей и методов консолидации и анализа данных, технологий создания интегрированных систем территориального и корпоративного управления	15
4 Исследование моделей управления безопасностью в информационно-телекоммуникационных системах	23
5 Разработка методов доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей на основе регрессионной оценки плотности вероятности.....	25
6 Последовательный анализ созданных ранее баз генетических данных с целью выявления в них структурированности и порядка	28
ЗАКЛЮЧЕНИЕ.....	29
Приложение А.....	30
Приложение Б.....	35
Приложение В.....	36

ВВЕДЕНИЕ

Развитие методологических и технологических основ создания интегрированных систем на основе методов консолидации данных, аналитической обработки больших объемов многомерных данных, методов и технологий межсистемного взаимодействия, технологий инфографического и геоинформационного моделирования, технологий модельно-ориентированной разработки программного обеспечения позволяет создавать качественно новые технологии информационно-аналитической поддержки корпоративного и территориального управления. Результаты проекта могут рассматриваться как новый научный метод решения важной научно-технической задачи – повышения эффективности информационно-управляющих систем.

Цель работы – создание новых методов обработки и анализа данных, интеллектуальных технологий имитационного моделирования и комплексной поддержки принятия решений как инструментов построения информационно-управляющих систем разного назначения.

Основные задачи:

1. Разработка методов интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата.
2. Разработка моделей и методов консолидации и анализа данных, технологий создания интегрированных систем территориального и корпоративного управления.
3. Исследование моделей управления безопасностью в информационно-телекоммуникационных системах.
4. Разработка методов доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей на основе регрессионной оценки плотности вероятности.
5. Последовательный анализ созданных ранее баз генетических данных с целью выявления в них структурированности и порядка.

Использованы оригинальные методы консолидации и оперативной аналитической обработки данных, методы интеллектуальной поддержки конструкторских задач, информационно-графического моделирования.

Настоящий отчет является заключительным по теме: БАЗОВЫЙ ПРОЕКТ IV.35.1.2 "МЕТОДЫ АНАЛИТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ И ТЕХНОЛОГИИ СОЗДАНИЯ

ИНТЕГРИРОВАННЫХ ИНФОРМАЦИОННО-УПРАВЛЯЮЩИХ СИСТЕМ"
ПРОГРАММЫ СО РАН IV.35.1 "ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И ТЕХНОЛОГИИ
СОЗДАНИЯ И ПРИМЕНЕНИЯ ИНТЕГРИРОВАННЫХ ИНФОРМАЦИОННО-
ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ ДЛЯ РЕШЕНИЯ ЗАДАЧ ПОДДЕРЖКИ ПРИНЯТИЯ
РЕШЕНИЙ", по которой были представлены отчеты по этапам 2013, 2014, 2015 годов.

1 Важнейший результат

Оптимальный выбор количества интервалов дискретизации области определения одномерной и двумерной плотности вероятности.

Аннотация. На основе анализа асимптотических свойств регрессионной оценки плотности вероятности одномерной и двумерной случайных величин впервые обоснованы и определены оптимальные зависимости количества интервалов дискретизации области её определения от объёма исходных статистических данных. Установлена инвариантность полученной зависимости к изменению параметров восстанавливаемой плотности вероятности. Полученная аналитическая зависимость по виду близка к формуле Хайнхольда-Гаеде и совпадает с ней при оценивании одномерной плотности вероятности с равномерным законом распределения. Результаты исследований имеют важное значение в задачах проверки гипотез о распределениях случайных величин с использованием критерия Пирсона, оценивании плотностей вероятности и построении их доверительных границ.

Задача дискретизации области значений случайной величины является одной из основных в математической статистике при оценивании плотности вероятности, построении её доверительных границ и проверке гипотез об их распределениях. Начиная с работы Старджесса, предложен ряд методов дискретизации интервала значений одномерной случайной величины, к которым относятся, например, формулы Хайнхольда-Гаеде, Брукса-Каррузера, Скотта, Фридмана-Диакониса. Для многомерных случайных величин предлагаются эвристические рекомендации, основанные на пороговом значении минимального количества m наблюдений попадающих в один интервал ($m=3, 5, 10$).

На основе результатов исследования асимптотических свойств регрессионной оценки плотности вероятности авторами обоснована методика выбора оптимального количества интервалов дискретизации области значений одномерной и двумерной случайной величины.

Для двумерной случайной величины $x = (x_v, v = \overline{1, 2})$ оптимальное количество интервалов дискретизации области её значений по каждой компоненте x_v определяется значением

$$N_v = \left\lfloor k_2 \sqrt[4]{n} \right\rfloor + 1, \quad k_2 = \sqrt[4]{\frac{3}{4} \Delta_1 \Delta_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p^2(x_1, x_2) dx_1 dx_2}, \quad v = 1, 2, \quad (1)$$

где Δ_1, Δ_2 – длины интервалов значений случайных величин x_1, x_2 , n – объём исходных статистических данных, а $\alpha > |\alpha|$ – целое число.

Предложенная формула оптимальной дискретизации (1) была использована для ранжирования по эффективности эвристических методов дискретизации $D(m)$ при $m=1, 3, 5, 10$ в задаче оценивания плотности вероятности. Например, для двумерного нормального закона распределения зависимых случайных величин количество оптимальных интервалов дискретизации области их значений зависит от коэффициента корреляции r , объёма исходных данных и метода дискретизации. С ростом абсолютного значения коэффициента корреляции r количество оптимальных интервалов дискретизации увеличивается. При малых значениях $|r| < 0,5$ эффективность оптимальной формулы дискретизации (1) сопоставима с эвристическим методом $D(m=1)$. В условиях $|r| > 0,5$ целесообразно использовать методы дискретизации $D(m=3)$ и $D(m=5)$ (таблица).

Таблица – Значения среднеквадратического отклонения оценки плотности вероятности $\bar{p}(x_1, x_2)$, соответствующие объёму n статистических данных, коэффициенту корреляции $r=0,3$ (а), $r=0,8$ (б) между x_1, x_2 и сравниваемым методам дискретизации

а)

n	Методы дискретизации				
	Формула (1)	$D(m=1)$	$D(m=3)$	$D(m=5)$	$D(m=10)$
100	0,0630	0,0713	0,0630	0,0690	0,0690
300	0,0436	0,0458	0,0436	0,0477	0,0477
500	0,0368	0,0368	0,0380	0,0380	0,0446
700	0,0330	0,0335	0,0357	0,0357	0,0433
900	0,0303	0,0303	0,0308	0,0344	0,0426
1100	0,0283	0,0283	0,0295	0,0335	0,0335
1300	0,0268	0,0270	0,0285	0,0330	0,0330
1500	0,0255	0,0255	0,0278	0,0278	0,0325

б)

n	Методы дискретизации				
	Формула (1)	$D(m=1)$	$D(m=3)$	$D(m=5)$	$D(m=10)$
100	0,1740	0,2100	0,1780	0,1740	0,2100
300	0,1200	0,1478	0,1200	0,1230	0,1440
500	0,1016	0,1432	0,1016	0,1018	0,1120
700	0,0900	0,1300	0,0920	0,0900	0,0940

900	0,0832	0,1250	0,0832	0,0835	0,0900
1100	0,0770	0,1230	0,0797	0,0770	0,0870
1300	0,0730	0,1230	0,0740	0,0730	0,0765
1500	0,0700	0,1100	0,0730	0,0700	0,0740

Отмеченная закономерность не зависит от значений математических ожиданий и среднеквадратических отклонений двумерных случайных величин. Полученные результаты имеют важное значение при проверке гипотез о распределениях случайных величин с использованием критерия Пирсона.

2 Разработка методов интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата

Обобщены и систематизированы работы по созданию технологии автоматизации исследования функциональных характеристик командно-измерительной системы космического аппарата. Построенная технология реализуется программным обеспечением контрольно-проверочной аппаратуры, которое позволяет выполнять измерения и управлять оборудованием командно-измерительной системы, обеспечивает наглядность построения сложных последовательностей испытательных процедур, а также удобство и корректность отображения результатов. Технология поддерживает схему организации испытательных процедур, показанную на рисунке 1.

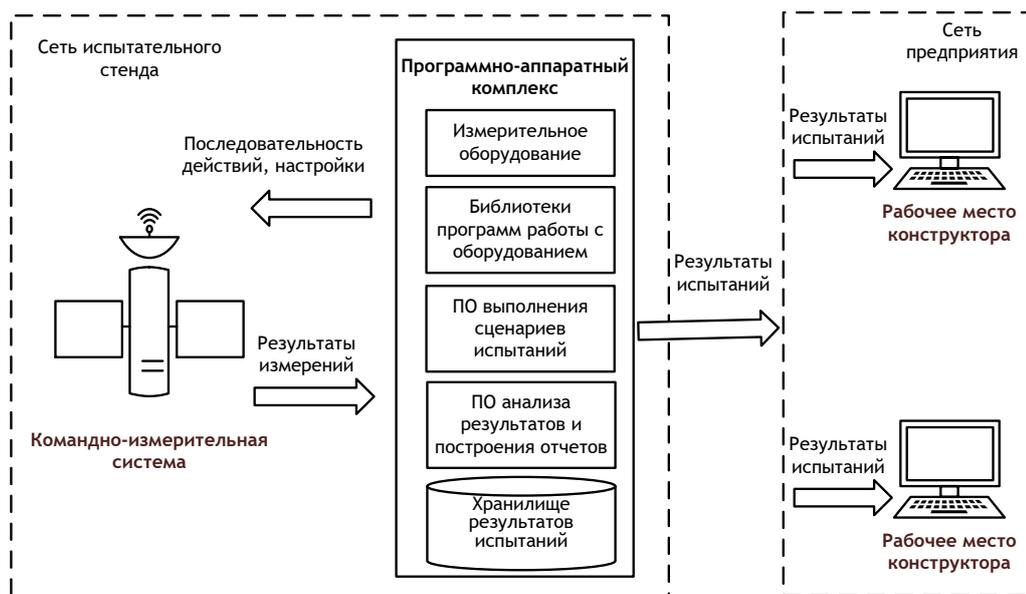


Рисунок 1 – Схема организации выполнения испытаний

В основе технологии испытаний бортовой аппаратуры командно-измерительной системы космического аппарата лежит имитационная модель её функционирования (рисунок 2).

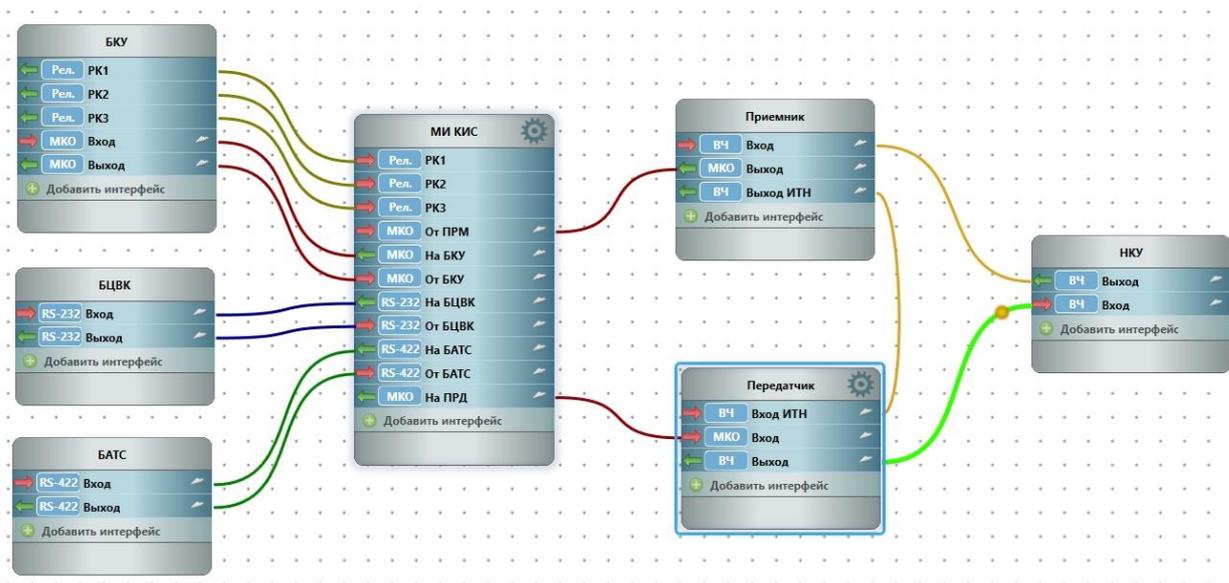


Рисунок 2 – Имитационная модель функционирования бортовой аппаратуры

Применение имитационных моделей позволяет рассматривать структуру и методы взаимодействия подсистем объекта контроля для изучения особенностей их функционирования и построения испытаний.

Для задания различных видов испытательных процедур предложены универсальные принципы их описания. Разработана модель для задания формализованных последовательностей действий по измерению или управлению объектом контроля – сценариев. Множество $A^0 = \{A_1^0, \dots, A_R^0\}$ – действия, которые используются для построения сценариев. $A_r^0 = \langle I_r, M_r \rangle$, $r \in \overline{1, R}$, где I_r – множество настроек действия, M_r – множество параметров мониторинга. Пример элементов из множества действий: генерация импульсов, измерение номинала частоты, управление мультиметром, измерение мощности СВЧ, измерение параметров сигналов релейных команд, измерение спектра и пр.

Совокупность действий, направленная на выполнение определенной задачи в сценарии испытаний, называется заданием. Задание в сценарии можно представить в следующем виде: $T_i = \langle A_i, C_i, V_i, k_i, l_i \rangle$, $T_i \in T$, $i \in \overline{1, |T|}$, где A_i – упорядоченная совокупность действий в рамках задания, C_i – описание цикла выполнения задания, V_i – количество этапов в задании, $k_i \in \overline{1, K}$ – начальный этап выполнения задания, $l_i \in \overline{1, K}$,

$l_i \geq k_i$ – конечный этап выполнения задания. Этапы выполнения заданий представляются как: $S = \langle T, K \rangle$, где T – упорядоченная совокупность заданий в рамках сценария, K – количество этапов в сценарии.

Упорядоченная совокупность действий в рамках задания может быть представлена как подмножество действий, для которых определены этапы их выполнения. Действие A_{ij} в рамках задания T_i можно представить как: $A_{ij} = \langle I_{ij}, M_{ij}, L_{ij}, F_{ij}, q_{ij}, p_{ij} \rangle$, $A_{ij} \in A_i$, $j \in \overline{1, |A_i|}$, где I_{ij} – множество настроек действия, M_{ij} – множество параметров мониторинга, L_{ij} – условия запуска выполнения действия, F_{ij} – условия остановки выполнения действия, $q_{ij} \in \overline{1, V_i}$ – начальный этап выполнения действия, $p_{ij} \in \overline{1, V_i}$, $p_{ij} \geq q_{ij}$ – конечный этап выполнения действия.

Пример сценария, состоящего из трех заданий, представлен на рисунке 3.



Рисунок 3 – Пример сценария испытаний

Выполнено описание сценария испытаний и правил его построения в виде контекстно-свободной грамматики $G = (N, T, P, St)$, где N – множество нетерминальных символов, T – множество терминальных символов, P – конечное множество продукций, St – начальный символ.

Множество нетерминалов содержит символы, обозначающие основные конструкции сценария испытаний. Множество терминалов включает элементы исходного множества действий A^0 , цифры, буквы русского и английского алфавитов, знаки операций сравнения и другие символы.

Начальным символом грамматики является <сценарий>. В расширенной форме Бэкуса-Наура определим продукции, которые порождают сценарий как последовательность этапов выполнения заданий: <сценарий> = <наименование>, (<этап_сценария>, {<этап_сценария>}); <этап_сценария> = “Этап сценария:”, (<задание>, {<задание>}).

Задание представляется как последовательность этапов выполнения действий:
<задание> = <наименование>, <продолжительность_этапов_сценария>, <цикл>, (<этап_задания>, {<этап_задания>}); <этап_задания> = “Этап задания:”, (<действие_в_задании>, {<действие_в_задании>}).

Действие в задании представляется как элемент из исходного множества действий и характеристик, определяющих правила его выполнения в задании:
<действие_в_задании> = <действие>, <продолжительность_этапов_задания>, <условия_запуска>, <условия_остановки>.

В модели определены представления значений параметров настроек и интервалов допустимых значений параметров мониторинга. Условия запуска действия задаются как правила вида «Если условие то выполнить (запуск, остановка)». В левой части правил сравниваются значения параметров мониторинга предыдущих действий в задании с определенными значениями.

Представленная формальная модель сценария позволяет описать все возможные варианты испытаний бортовой аппаратуры командно-измерительной системы космического аппарата. Разработан графический метод построения сценариев испытаний, позволяющий формировать испытания, расширять функций управления и измерения параметров объекта контроля. Реализация формальной модели и метода построения сценариев испытаний выполняется в графическом редакторе.

Реализован модуль исполнения сценария испытаний с возможностью контроля параметров в реальном времени или выполнения послесеансной обработки результатов. Проведение испытаний и анализ результатов сопровождаются визуализацией хода выполнения сценария, последовательности и параметров испытательных процедур, контрольных точек, результатов измерений (рисунок 4).

Разработанная технология автоматизации испытаний бортовых систем космического аппарата на основе формальной модели сценария объединяет информационные, измерительные и вычислительные функции. Реализация представленной технологии расширяет возможности проведения экспериментальных исследований и дает возможность принимать научно-обоснованные конструкторские решения.

Применение современных технологий проведения испытаний требует от специалиста – конструктора бортовой аппаратуры высокой квалификации и широких знаний об различных особенностях функционирования объекта испытаний. В этой связи продолжается развитие методов формирования учебно-исследовательских материалов на

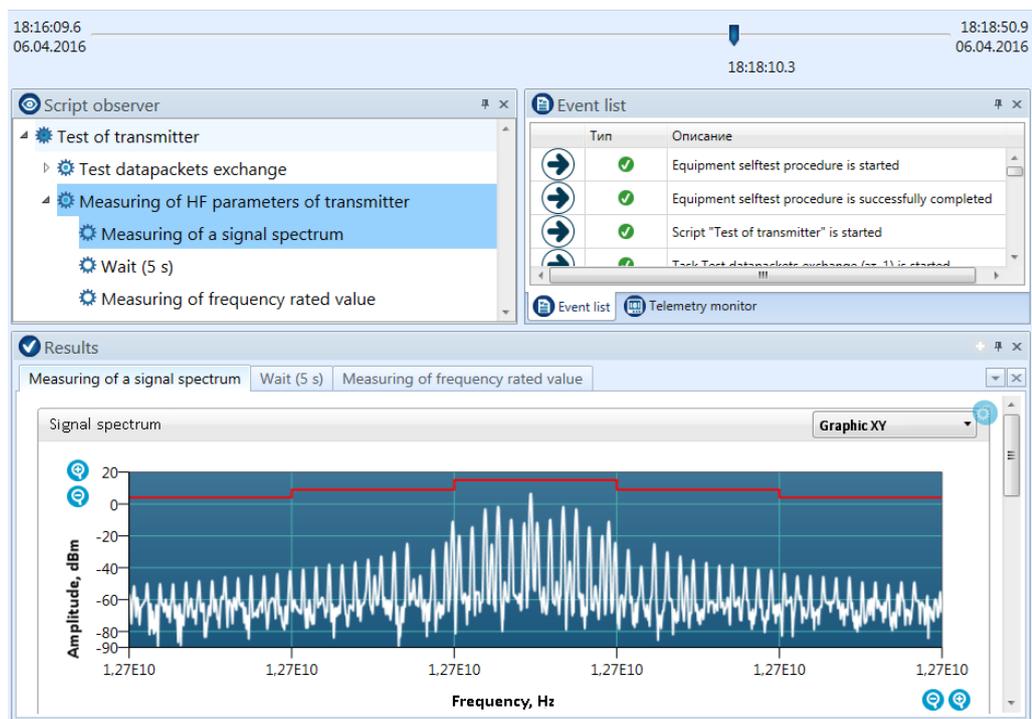


Рисунок 4 – Отображение процесса проведения испытаний

основе имитационных моделей бортовой аппаратуры космического аппарата. Разработаны четыре уровня заданий в учебно-исследовательской системе. Уровни идентификации ранее изученного материала и воспроизведения знаний реализуются в системе контрольно-проверочным модулем, позволяющим создавать тестовые задания. Задания по использованию знаний в стандартных и нестандартных ситуациях формируются при помощи инструментов имитационного моделирования. Инструменты имитационного моделирования, встроенные в учебно-исследовательскую систему, позволяют создавать модель бортовой аппаратуры командно-измерительных устройств из различных элементов. Элементы взаимодействуют друг с другом путём передачи и приема пакетов данных, которые можно создавать в специальной подсистеме «Редактор пакетов данных». Можно задавать логику функционирования, которая определяет поведение, как отдельных элементов, так и системы в целом. Имитационная модель позволяет создавать для обучаемых специалистов задания для развития их навыков работы, как в штатных, так и в нештатных ситуациях.

3 Разработка моделей и методов консолидации и анализа данных, технологий создания интегрированных систем территориального и корпоративного управления

Решение задачи консолидации данных гетерогенных источников сопряжено с созданием теоретического базиса для объединения разнородных и разрозненных данных.

В качестве консолидирующей основы разработана оригинальная реализация многомерной модели данных, которая, с одной стороны, разделяет аналитические концепты на аспекты анализа (измерения) и числовые характеристики анализируемого процесса (показатели), обеспечивая возможность оперативной агрегации и детализации данных OLAP-куба, а с другой стороны, сохраняет аналитические связи между ними, позволяя осуществлять адаптивное манипулирование концептами и строить произвольные аналитические запросы. Исследование существующих подходов к построению многомерной модели данных показало, что, как правило, эта задача решается в ходе одного из этапов проектирования хранилища данных и направлена сугубо на разработку схемы эффективного хранения и доступа к данным. Оригинальность предложенного подхода заключается в построении связной многомерной модели данных в виде семантической надстройки над физической схемой хранения данных, независимо от формата. Более высокий уровень абстракции позволяет сохранить атомарность аналитических концептов, их естественное представление и связи между ними. Единая методология многомерного моделирования разнородных данных обеспечивает возможность объединения частных моделей в единую консолидированную модель. Разработаны алгоритмы генерации связной концептуальной модели многомерных данных для нормализованных реляционных баз данных на основе графа функциональных зависимостей, а также для XML-файлов – на основе структурных особенностей элементов документа.

Разработаны методические и программные средства оперативного межсистемного обмена данными на основе интерпретируемых XML-шаблонов. Шаблон задает структуру и информационное наполнение, устанавливает соответствие между форматами обмена – узлами XML-документов и таблицами реляционной базы данных. Разработаны алгоритмы анализа XML-шаблонов и построения дерева структурных элементов, обеспечивающего визуальную настройку условий и параметров обмена. Дерево структурных элементов представляет собой информационную модель соответствия, где каждый объект содержит описание таблицы реляционной базы данных в формате {свойство: значение}. Для обеспечения двустороннего обмена, с учетом направления передачи, разработаны алгоритм наполнения базы данных и алгоритм генерации XML-документов. На основе анализа информационной модели соответствия первый алгоритм выполняет извлечение данных из XML-документов и размещает их в соответствующих таблицах реляционной базы данных, второй алгоритм заполняет шаблон данными из таблиц базы данных и формирует XML-документ. Выполнена программная реализация разработанных алгоритмов. Предложенный подход к организации межсистемного обмена данными позволяет вносить изменения в структуру обрабатываемых документов без

перепрограммирования и реализовать требуемый функционал по интеграции гетерогенных источников данных в отдельном программном модуле. Применение XML обеспечивает универсальный формат обмена и хорошую структурированность данных.

Выполнено развитие ранее предложенного подхода к унификации справочной информации, позволяющего систематизировать структуру корпоративных справочников и классификаторов. В состав типового справочника введены информационные поля, позволяющие однозначно идентифицировать записи справочника, как при локальном использовании, так и при широком распространении среди участников информационного обмена. Записи справочников обладают установленным периодом действия, все изменения записей датируются и фиксируются в системных журналах. При необходимости справочник может быть дополнен необходимой атрибутивной информацией, хранящейся в таблицах свойств. В системе может создаваться произвольное количество древовидных структур – оглавлений, которые позволяют иерархически упорядочить данные из одного, или нескольких справочников одновременно. Для установки связей между справочниками используются поля-ссылки, которые могут вводиться в состав справочника. Связи типа «многие-ко-многим» устанавливаются с помощью специальных таблиц связей. Принцип однородности структуры различных справочников и классификаторов, используемых на отраслевом уровне, позволяет применять единые механизмы обработки в работе с разнотипными реестрами и справочниками. Полученные теоретические результаты межсистемного взаимодействия и интеграции данных апробированы в задачах организации муниципальных закупок.

Разработана технология консолидации данных комплексного мониторинга природно-техногенной безопасности в централизованное хранилище данных (рисунок 5). Технология позволила решить проблему межведомственной интеграции данных, в том числе данных инструментального мониторинга, и обеспечить их совместную аналитическую on-line обработку. Разработаны принципы систематизации мониторинговых данных, позволяющие формализовать представление первичной информации и результатов её аналитической обработки (рисунок 6).

Технология консолидации данных апробирована в Территориальном центре мониторинга и прогнозирования ЧС Красноярского края. Сбор данных осуществляется в автоматическом режиме из ведомственных сетей, с приборов контроля и распределённой системы сбора данных. Аналитическая обработка мониторинговых данных выполняется на основе методов OLAP и динамического картографирования.

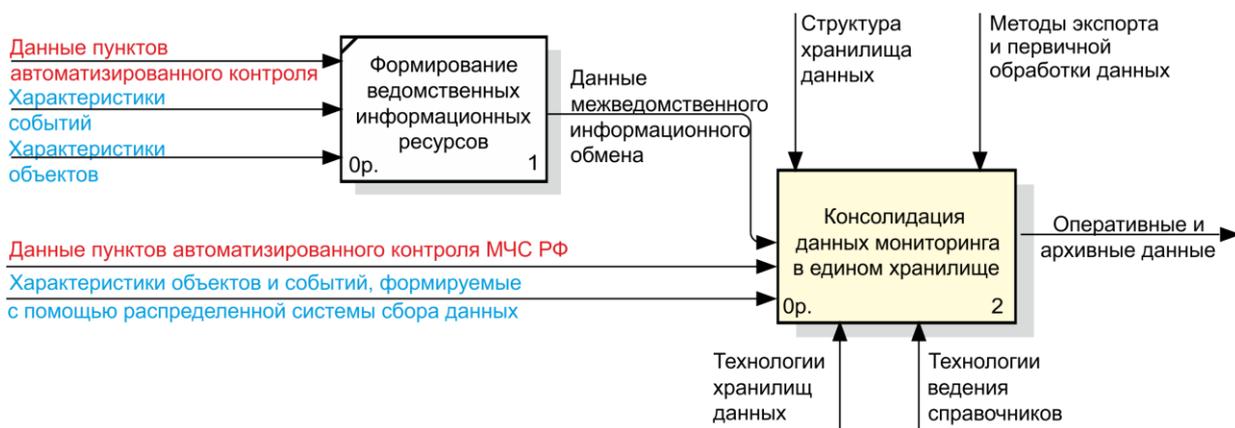


Рисунок 5 – Процесс консолидации данных оперативного мониторинга

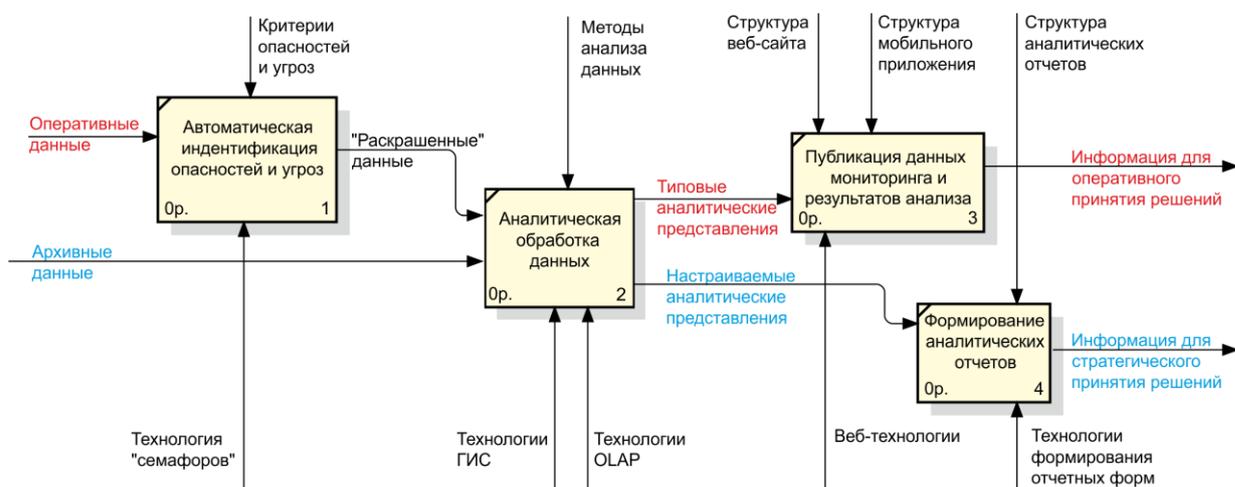


Рисунок 6 – Технологии обработки данных мониторинга

В рамках исследования структурных характеристик территорий, выполнен комплексный анализ показателей природной и техногенной безопасности Красноярского края, направленный на исследование географических особенностей и закономерностей возникновения ЧС, на основе применения методов интеллектуальной обработки данных – метода главных компонент и кластерного анализа – к данным паспортов безопасности населенных пунктов. Выполнена оценка главных компонент и их интерпретация с учётом вклада анализируемых характеристик. Выполнен анализ распределения данных по главным компонентам на разных уровнях детализации территории: группы районов, муниципальные образования и населённые пункты. В многомерном пространстве данных построены двух- и трёхкластерные структуры. Результаты анализа показали, что наибольшую угрозу территориям Красноярского края представляют техногенные и лесные пожары. Высокий риск техногенных пожаров наблюдается в крупных населённых пунктах, в которых численность населения, количество социально-значимых и

промышленных объектов выше средних значений. Высокий риск природных пожаров наблюдается в крупных поселениях, непосредственно прилегающих к лесным массивам, а также в городах, в границах которых расположены лесные массивы. Результаты анализа позволили определить зоны повышенного риска, ранжировать населённые пункты по степени опасности возникновения ЧС природного и техногенного характера, что позволяет более эффективно планировать и проводить мероприятия по предупреждению и смягчению последствий ЧС на территории Красноярского края.

На основе интеграции вычислительных методов и с применением экспертных знаний разработаны методы стандартизации и классификации текстуальных записей. На примере записей, полученных по данным переписи Великобритании за 1851-1911 годы, разработаны два метода: метод классификации записей о профессиональной деятельности и метод стандартизации записей о месте рождения. Для классификации записей о профессиональной деятельности на основе теории принятия решений и байесовского подхода разработаны два алгоритма: первый алгоритм позволяет определить классификационный код для строки описания профессии путем выявления наиболее значимого слова описания и определения его классификационного кода; второй алгоритм позволяет определить классификационный код для строки описания профессии путем комбинации классификационных кодов всех слов описания. Для стандартизации записей о месте рождения разработан алгоритм, реализующий два основных этапа: идентификации приходов путём проверки строки описания и идентификации приходов путём проверки отдельных слов из строки описания. Выполнено формальное описание принципов идентификации в виде диаграммы событий управляющей модели ARIS. С целью реализации алгоритмов разработаны специализированные справочники и классификаторы. Разработанные алгоритмы применены для автоматической обработки миллионов записей переписи населения и могут успешно применяться для решения аналогичных задач формализации текстуальных записей.

Для оценки и исследования состояния пожарной безопасности здания разработан метод консолидации и анализа данных результатов моделирования эвакуации людей из здания и распространения опасных факторов пожара, позволяющий выявлять опасные участки эвакуационных путей, оценивать степень влияния опасных факторов пожара на людей, получать агрегированные показатели, необходимые для расчета пожарного риска и поддержки управленческих решений по повышению уровня пожарной безопасности объекта. Предложенный метод позволяет решить задачи консолидации данных и их анализа. В процессе консолидации происходит приведение двух массивов данных к виду, необходимому для быстрого выполнения аналитических запросов. Суть метода

заключается в совместном анализе двух одновременно происходящих процессов (распространения опасных факторов пожара и эвакуации людей) с целью выявления угроз воздействия опасных факторов пожара на участников эвакуации. В результате применения метода строятся сравнительные таблицы стадий развития пожара и эвакуации, которые могут быть использованы специалистами по пожарной безопасности для принятия решений и анализа условий текущей эксплуатации объекта. Метод консолидации и анализа результатов моделирования эвакуации и распространения опасных факторов пожара позволяет выполнять расчет времени эвакуации и времени блокирования эвакуационных путей – эти параметры участвуют в процессе оценки состояния пожарной безопасности объекта в условиях текущей эксплуатации и формировании рекомендаций по снижению пожарного риска.

Выполнено развитие средств интерактивного представления результатов оперативной аналитической обработки данных, основанное на применении принципов инфографического моделирования, которые позволяют повысить когнитивную эффективность оперативного анализа и максимально полно использовать аналитический потенциал OLAP-средств. Выполнен обзор существующих подходов и инструментов построения инфограмм, определены основные виды инфограмм. Предложен подход к построению аналитических инфографических моделей для отображения многомерных результатов оперативного анализа с учетом состава и характера данных. Для построения инфограмм необходимо определить размерность отображаемых данных в терминах технологии OLAP, проанализировать способы отображения многомерных данных, выявить основные когнитивные элементы инфограмм. Систематизация инфограмм позволила создать фундамент для разработки аналитических инфографических моделей. На рисунке 7 представлен пример анализа статистики о проведенных научных мероприятиях.

Инфограмма состоит из двух диаграмм, отображающих два разных показателя одного многомерного куба данных. Измерениями построенной OLAP-модели являются: год, наименование мероприятия, организация; показателями – количество статей, их авторов и сделанных докладов. Выполнен срез куба, где фиксирован год – 2015 и мероприятие – XIV Всероссийская конференция «Проблемы информатизации региона». Отображены показатели: количество авторов статей, количество прочитанных докладов сотрудниками различных организаций. Актуальные задачи дальнейшего исследования – создание методических основ многомерного когнитивного моделирования и разработка аналитических инфографических моделей. Необходимо решить проблему взаимодействия с OLAP-системами, увеличить скорость построения инфограмм при выполнении операций

XIV Всероссийская конференция «Проблемы информатизации региона» (ПИР-2015)

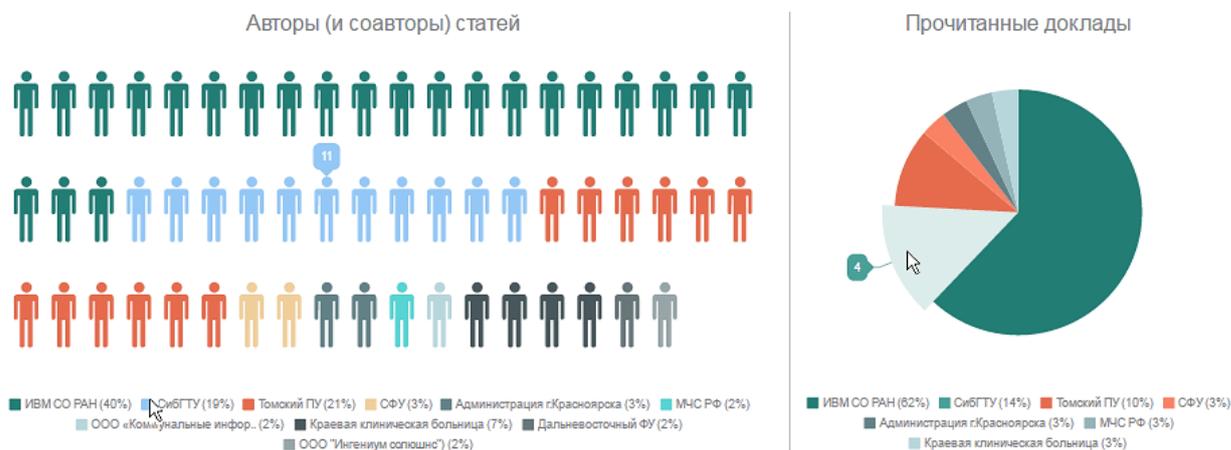


Рисунок 7 – Динамическая инфограмма для отображения статистики конференции

над гиперкубом. Особого внимания требует задача восприятия информации – выбора разных способов графического представления при визуализации результатов многомерного анализа данных.

Продолжены работы над архивом (репозиторием) открытого доступа к научным публикациям ИВМ СО РАН. Архивы открытого доступа (Open Access), создаваемые на базе Интернет-протоколов OAI-PMH, позволяют существенно повысить показатели цитирования научных публикаций. Для формирования репозитория использована программная платформа DSpace. Выполнены работы по отладке схем преобразования данных между архивом открытого доступа и системой автоматизации библиотек ИРБИС, в которой ранее формировалась база данных трудов сотрудников института. Определен набор полей метаданных для ввода публикаций. Выполнены работы по дифференциации доступа к публикациям на основе сетевых протоколов LDAP. Базы LDAP развернуты на UNIX-сервере института. Выполнены работы по внедрению методов автоматической поддержки издательского эмбарго, свободного лицензирования публикаций и дифференциации доступа к данным. Выполнены работы по интеграции архива открытого доступа с поисковыми системами Интернет, по самодепонированию публикаций и свободному лицензированию. Архив открытого доступа запущен в тестовом режиме.

Завершены работы по формированию базы трудов сотрудников института, как полнофункционального средства учета публикационной активности. База данных трудов сотрудников аккумулирует многолетние результаты научной деятельности и предоставляет хорошее средство учета научных публикаций. В соответствии с новыми требованиями, расширен набор полей базы, которые пополнились данными об импакт-факторе источника, о цитировании публикации, списком пристатейной библиографии и

т.д. Расширены взаимосвязи с локальными и онлайн-ресурсами, что позволяет обогатить описания идентификаторами публикаций в индексах научного цитирования (Web of Science, Scopus, РИНЦ) и обеспечить динамическое обновление показателей цитирования. Обеспечена связь электронных полных текстов публикаций цифрового репозитория с записями в базе данных с помощью URL-ссылок, что позволяет контролировать точность и полноту описаний. Выполнена интеграция текущей база трудов сотрудников института и электронного каталога библиотеки в информационную систему СО РАН (программная платформа сформирована ранее, как результат ряда научных проектов). Данные базы трудов сотрудников используются в библиометрических и наукометрических исследованиях.

В сфере разработки технологий создания интегрированных систем предложена оригинальная реализация технологии модельно-ориентированной разработки программного обеспечения (Model Driven Development, MDD), учитывающая требования динамического перестроения тематического наполнения системы. Построение модельно-ориентированных систем сбора данных включает в себя этапы проектирования моделей абстрактного уровня и моделей прикладного уровня. Проектирование моделей абстрактного уровня включает процесс построения мета-метамодели (M3), модели языка моделирования и процесс построения метамодели (M2), метамодели различных специфичных областей приложения. Проектирование прикладных моделей состоит из процесса построения моделей прикладного уровня (M1) и этапа формирования экземпляров концептов (M0) определённых на уровне M1. В отличие от классического подхода в авторской реализации, во-первых, предлагается объединить процессы формирования управляющей и прикладной моделей (рисунок 8), включение управляющей модели в состав системы позволяет оперативно реагировать на изменяющиеся требования к тематическому наполнению. Во-вторых, предлагается добавить процесс формирования модельно-ориентированной системы, который позволит получить в результате выполнения всех процессов не программный код, а рабочую систему.

Предложенная реализация модельно-ориентированного подхода к разработке программного обеспечения позволяет получать модельно-ориентированную систему с функцией управления метамоделью и возможностью изменения своего функционала силами обычных пользователей. Разработан алгоритм генерации программных компонентов (рисунок 9). Применение данного алгоритма позволяет достичь высокого уровня адаптируемости информационной системы в соответствии с меняющимися потребностями пользователей.

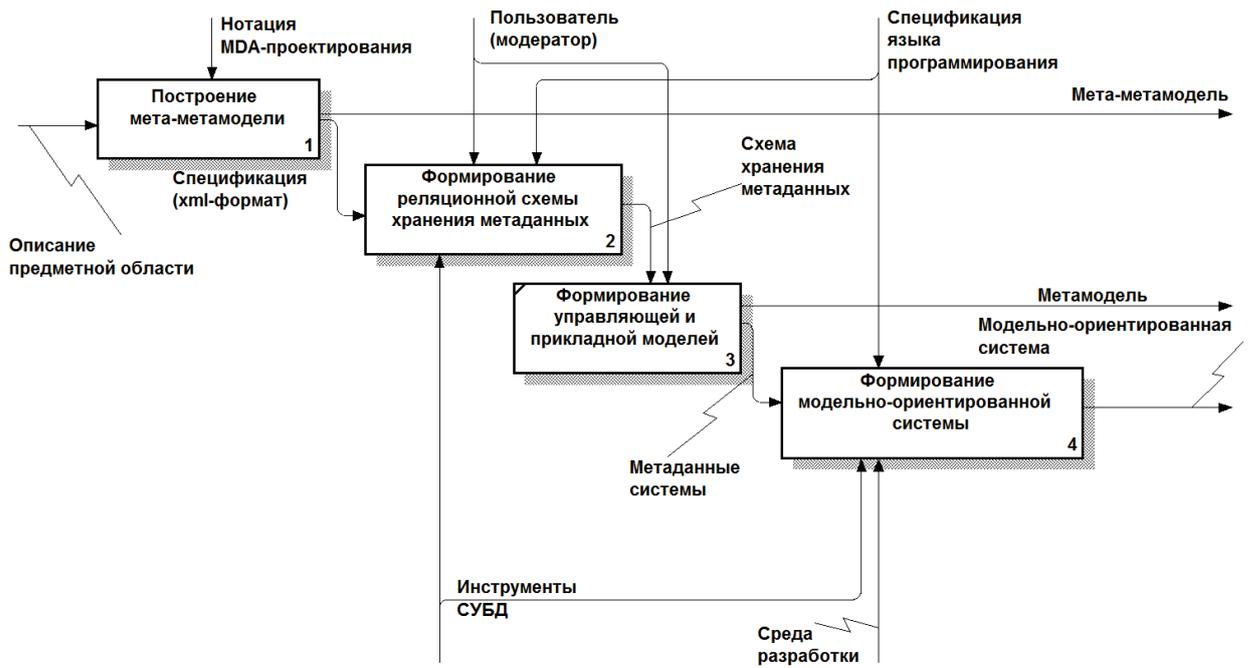


Рисунок 8 – Диаграмма оригинальной схемы проектирования системы

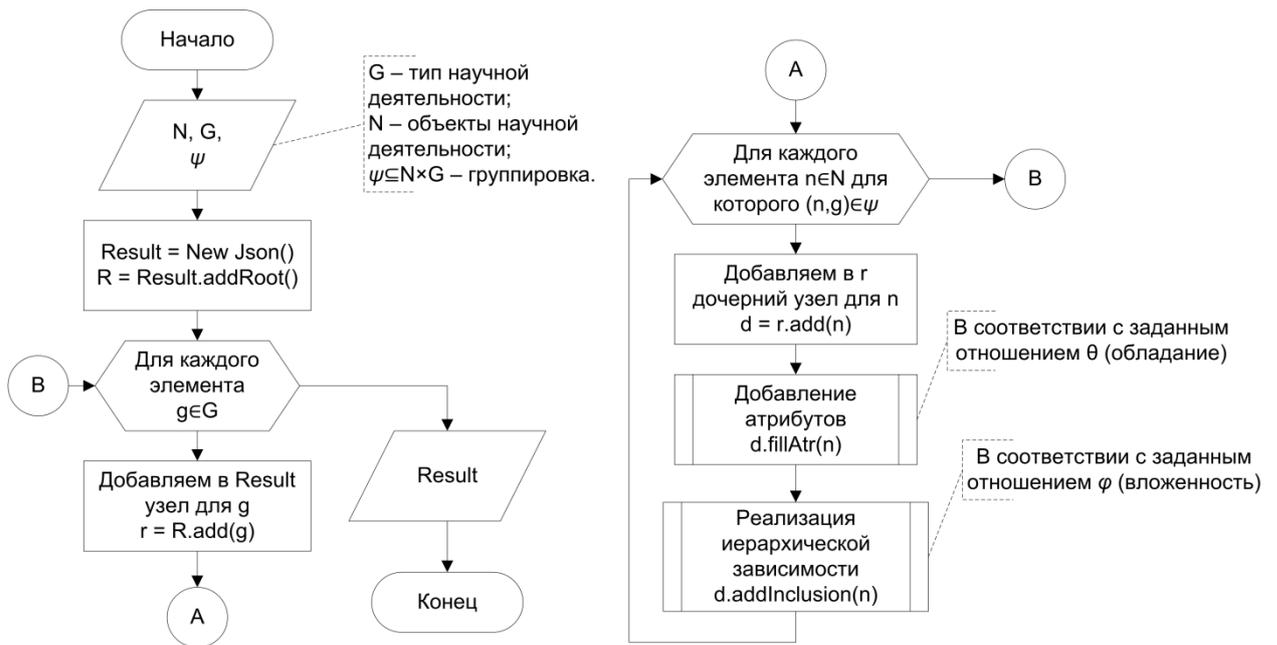


Рисунок 9 – Алгоритм генерации программных компонентов

Предложенный алгоритм апробирован при создании автоматизированной системы учёта результатов научной деятельности ИВМ СО РАН и веб-системы распределенного сбора данных, являющейся частью системы консолидации и анализа данных арктического и территориального центров мониторинга и прогнозирования ЧС.

4 Исследование моделей управления безопасностью в информационно-телекоммуникационных системах

Проведены исследования в области безопасности интернет-сервисов и телекоммуникационных сетей. Проведен анализ моделей управления безопасностью в информационно-телекоммуникационных системах. Рассмотрены классические модели управления безопасностью: дискреционные, мандатные, ролевые и изолированной программной среды. Установлено, что данные модели не обеспечивают необходимой гибкости конфигурирования и разграничения доступа. Также проанализирована расширенная ролевая модель, основанная на иерархии путей и представляющая развитие базовой ролевой модели RBAC. Исследования показали, что элементы данной модели хорошо отражают структуру информационно-телекоммуникационных систем при условии централизованного управления. В частности, модель может применяться в информационно-телекоммуникационных сетях, построенных в соответствии с классической иерархией (ядро сети, уровень распределения, уровень доступа).

Выявлены основные источники информации для идентификации агентов угроз. На основе накопленной информации проведено исследование взаимосвязи угроз по отдельным интернет-сервисам. Выявлены сильные корреляции между отдельными группами сервисов (рисунок 10), такими как: HTTP-HTTPS – доступ к веб-серверам по обычному и защищенному соединениям (0,957), SSH-HTTPa – сервисы удаленного входа и прокси-серверы (0,607); MySQL-SSH – доступ к сетевой базе данных и сервисы удаленного входа (0,618).

	HTTP	HTTPS	Telnet	HTTPa	SIP	RDP	SSH	mSQL	Radmin	MySQL
HTTP	1	0,957	-0,2	0,016	-0,104	0,202	-0,072	-0,114	-0,06	-0,017
HTTPS	0,957	1	-0,189	-0,036	-0,086	0,208	-0,03	-0,115	-0,06	0,054
Telnet	-0,2	-0,189	1	-0,17	0,179	-0,191	0,067	0,367	0,027	-0,163
HTTPa	0,016	-0,036	-0,17	1	0,134	-0,031	0,607	-0,351	0,295	0,468
SIP	-0,104	-0,086	0,179	0,134	1	0,464	0,27	0,386	0,4	0,174
RDP	0,202	0,208	-0,191	-0,031	0,464	1	0,279	-0,036	0,241	0,153
SSH	-0,072	-0,03	0,067	0,607	0,27	0,279	1	0,286	0,191	0,618
mSQL	-0,114	-0,115	0,367	-0,351	0,386	-0,036	0,286	1	0,014	-0,204
Radmin	-0,06	-0,06	0,027	0,295	0,4	0,241	0,191	0,014	1	-0,06
MySQL	-0,017	0,054	-0,163	0,468	0,174	0,153	0,618	-0,204	-0,06	1

Рисунок 10 – Коэффициент линейной корреляции сервисов

Кроме вышперечисленных связей, на другой выборке был получен коэффициент корреляции 0,497 на паре протоколов SSH-SMTP – попытки входа на удаленный сервер и

доступа к почтовому серверу. Таким образом, существует сильная связь между сервисами HTTP-HTTPS, и при обращении к одному из них велика вероятность попытки доступа по второму. В свою очередь, доступ по этим протоколам не коррелирует с остальными протоколами, и можно сделать вывод о существенно различных источниках угроз. Среди остальных сервисов наиболее показательным, по попыткам доступа, является протокол SSH (Secure Shell), доступ по которому может служить индикатором попыток обнаружения слабо защищенных сервисов из второй группы (SSH, MYSQL, HTTPa, SMTP). Проведенное исследование позволило выявить взаимосвязи между отдельными видами угроз и усовершенствовать применяющиеся в организации модели управления безопасностью и методы превентивной защиты.

Проведен мониторинг кибербезопасности корпоративной сети Федерального исследовательского центра «Красноярский научный центр СО РАН». Выявлены наиболее активные источники угроз и адаптированы средства защиты. Осуществлялась поддержка системы противодействия массовым рассылкам и антивирусной системы корпоративной почты. Выявлено существенное увеличение (более чем в 2 раза) количества массовых рассылок в четвертом квартале 2016 года (рисунок 11).

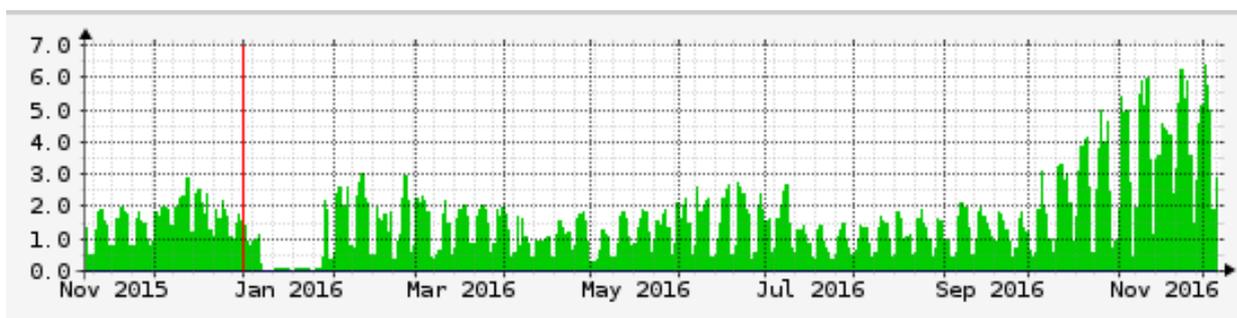


Рисунок 11 – Скорость поступления писем, идентифицированных как спам.

Для усиления безопасности созданы виртуальные машины в кластере Hyper-V для изолированной работы критически важных веб-приложений, функционирующих в составе интегрированных информационно-управляющих систем. Выполнены работы по оптимизации автоматизированной системы выявления и блокирования потенциально опасных хостов, осуществляющих попытки несанкционированного доступа к веб-ресурсам.

На основе инструментов Google maps с помощью библиотеки Google maps clusters создан действующий прототип системы визуализации источников интернет-атак картодиаграммами (рисунок 12). Введённые оси времени позволяют детализировать

интересующий интервал, а картографические средства отобразить необходимый элемент пространства.

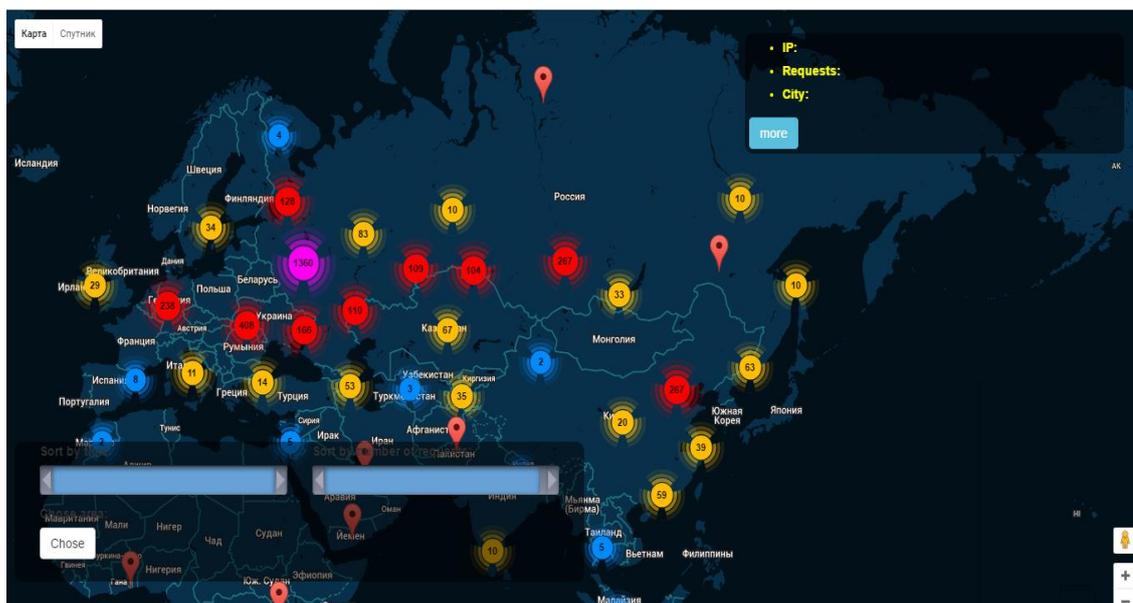


Рисунок 12 – Визуализация источников интернет-угроз

Выполненные работы позволили выявить основные методы для эффективной визуализации и анализа источников интернет угроз. Развитие данной системы позволит повысить эффективность работы специалистов по защите информации.

5 Разработка методов доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей на основе регрессионной оценки плотности вероятности

Разработана новая методика и алгоритмические средства доверительного оценивания плотности вероятности случайных величин с использованием её регрессионной оценки. Идея предлагаемого подхода состоит в выполнении следующих действий:

- декомпозиция исходных статистических данных на множество случайных величин и формирование массива соответствующих им вероятностных характеристик;
- оценивание верхней и нижней границ вероятностей принадлежности случайной величины введённым интервалам дискретизации при заданном коэффициенте доверия;
- по полученной информации, используя непараметрическую оценку кривой регрессии, синтез верхней и нижней доверительных границ для плотности вероятности.

Размеры области, определяемые доверительными границами, зависят от количества интервалов дискретизации случайных величин, их вероятностных характеристик и объёма исходных статистических данных. Установлено, что при большом объёме статистических данных применение формулы Хайнкольда и Гаеде для дискретизации интервала изменения случайной величины является предпочтительным. При относительно малых объёмах результаты использования формул Хайнкольда-Гаеде и Брукса-Каррузера сопоставимы.

Полученные результаты обобщены при доверительном оценивании байесовского уравнения разделяющей поверхности в дуальтернативной задаче распознавания образов и оптимальной решающей функции в среднеквадратическом смысле при восстановлении стохастической зависимости. Предлагаемый подход, например, при доверительном оценивании решающей функции в задаче распознавания образов предполагает разбиение области значений случайной величины на непересекающиеся интервалы в каждом классе и последующее доверительное оценивание вероятностей им принадлежности по исходной статистической информации. На этой основе осуществляется синтез доверительных границ плотностей вероятности распределения случайных величин в классах и уравнения разделяющей поверхности. Установлена зависимость размеров доверительных границ от количества интервалов дискретизации, объёма обучающей выборки и коэффициента доверия.

Исследованы асимптотические свойства регрессионной оценки плотности вероятности ядерного типа, что позволило впервые сформировать критерий выбора оптимального количества двумерных интервалов дискретизации области значений случайных величин. С этих позиций определена зависимость оптимального количества интервалов дискретизации области значений двумерной случайной величины от объёма статистических данных и характеристик восстанавливаемой плотности вероятности. Полученные результаты открывают возможность вывести формулу оптимальной дискретизации области значений многомерной случайной величины, которая необходима для задач проверки гипотез о распределениях случайных величин с использованием критерия Пирсона.

Разработана непараметрическая оценка плотности вероятности с неявно заданной формой ядерной функции. Процедура оптимизации исследуемой статистики реализуется путём последовательного выбора коэффициента размытости ядерной функции из условия максимума функции правдоподобия и параметра сглаживающего оператора с использованием формул дискретизации интервала значений случайной величины. По сравнению с непараметрической оценкой плотности вероятности Розенблатта-Парзена,

исследуемая её модификация характеризуется большим значением смещения, но меньшей дисперсией и среднеквадратическим отклонением. Отмеченные свойства сохраняются при значительно отличающихся условиях вычислительных экспериментов.

6 Последовательный анализ баз генетических данных с целью выявления в них структурированности и порядка

Рассмотрена модель пространственно распределенного сообщества, состоящего из двух видов, взаимодействующих по типу «хищник-жертва»; каждый из видов занимает две станции. Перемещение особей между станциями (миграция) не является случайным, и миграция предусматривает максимизацию коэффициента размножения каждого вида. Модель дискретна по пространству и по времени. Реализованы три версии базовой модели, где каждый вид реализует рефлексивную стратегию поведения для определения оптимального миграционного потока. Установлено, что рефлексивное поведение дает преимущество для рассматриваемых видов для некоторых конкретных наборов параметров. Однако регулярное сканирование области параметров показывает, что нерефлексивное поведение дает преимущество в большинстве комбинаций параметров.

Получены предварительные результаты исследования структуры геномов хлоропластов. Под структурой понимается кластеризация точек, соответствующих отдельным фрагментам генома хлоропласта (длиной порядка 200 нуклеотидов) в пространстве частот триплетов; частоты триплетов подсчитывались с пересечением, так, что каждый нуклеотид давал старт триплету. В ходе эксперимента проанализировано 188 геномов хлоропластов растений самых разных таксономических уровней. Геном хлоропластов преобразовывался в частотные словари триплетов. Затем для каждого генома в 63-мерном пространстве этих частот с помощью программы VidaExpert построены проекции данных, соответствующих выделяемым фрагментам генома, в пространстве первых трёх главных компонент (рисунок 13).

Обнаружено, что подавляющее большинство геномов в пространстве первых трёх главных компонент имеет очень похожую пространственную структуру. Кроме того, для геномов вычислен CG-контент, который является ведущим параметром классификации семикластерных структур геномов бактерий; ожидалось, что хлоропласты, ведущие своё происхождение от бактерий, будут также подчиняться этому правилу. Установлено, что для хлоропластов данный параметр не является ведущим в классификации структур, наблюдаемых при анализе главных компонент распределения фрагментов геномов в пространстве частот триплетов.

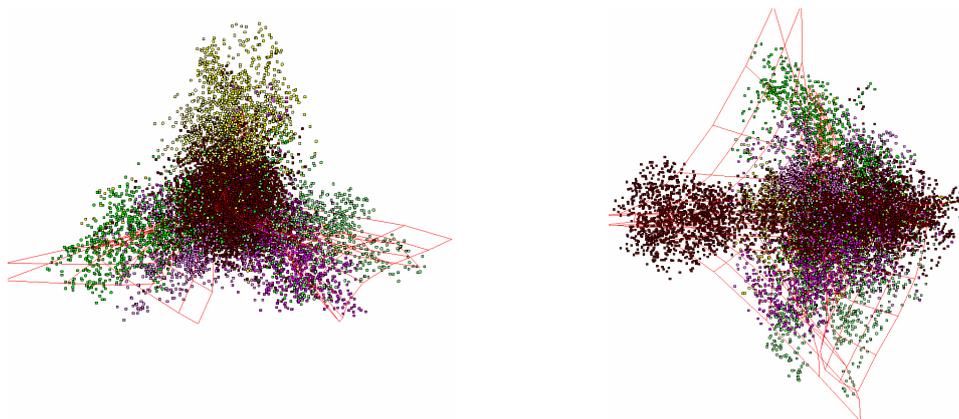


Рисунок 13 – Типичный вид структуры данных геномов наземных растений в проекциях пространства первых трех главных компонент. Представлена структура генома *Nimphaea alba*, слева – «анфас», справа – «в профиль»

ЗАКЛЮЧЕНИЕ

Задачи, запланированные на отчетный период, выполнены полностью.

Получены следующие основные результаты работы:

- Методы интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата.
- Модели консолидации данных гетерогенных источников в соответствии с форматами хранения.
- Модели и алгоритмы анализа структурных характеристик территорий.
- Алгоритмы построения управляющей модели для разработки модельно-ориентированного программного обеспечения.
- Технологии создания интегрированных систем территориального и корпоративного управления.
- Модели управления безопасностью в информационно-телекоммуникационных системах.
- Новые методы доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей.
- Анализ корреляции между кластерами, выделяемыми по статистическим и комбинаторным свойствам нуклеотидных последовательностей и соответствующими биологическими объектами.

Область применения – построение интегрированных информационных систем для комплексной поддержки принятия решений в разных прикладных областях.

ПРИЛОЖЕНИЕ А

Основные публикации

1. Зеньков И.В., Им С.Т., Лапко А.В., Лапко В.А., Музалевский К.В., Охоткина Е.А., Ружечка З.З., Харук В.И., Юронен Ю.П. Развитие и применение информационных технологий исследования природных ресурсов территорий Сибири на основе данных дистанционного зондирования. – Красноярск: СибГАУ, 2016. – 238 с. (в печати, план издания СибГАУ 2016 год, 19 уч. изд. л., тираж 500 экз.)
2. Lapko A.V., Lapko V.A. Optimal selection of the number of sampling intervals in domain of variation of a one-dimensional random variable in estimation of the probability density // *Measurement Techniques*. – 2013. – Vol. 56, No. 7. – P. 24-27. (DOI: 10.1007/s11018-013-0279-x)
3. Lapko A.V., Lapko V.A. Selection of the Optimal Number of Intervals Sampling the Region of Values of a Two-Dimensional Random Variable // *Measurement Techniques*. – 2016. – Vol. 59. No.2. – P. 122-126. (DOI: 10.1007/s11018-016-0928-y)
4. Lapko A.V., Lapko V.A., Yuronen E.A. Testing of hypothesis of random variables independence on the basis of nonparametric algorithm of pattern recognition // *International Siberian Conference on Control and Communications (SIBCON 2016)*, National Research University "Higher School of Economics" Moscow, 12-14 May 2016, Code 122173. – P.1-4 (DOI: 10.1109/SIBCON.2016.7491707)
5. Nozhenkova L.F., Isaeva O.S., Vogorovski R.V. Automation of Spacecraft Onboard Equipment Testing // *International Conference on Advanced Material Science and Environmental Engineering*, 2016, pp. 215-217 (DOI: 10.2991/amsee-16.2016.57).
6. Sadovsky M., Senashova M. Model of Prey-Predator Dynamics with Reflexive Spatial Behaviour of Species Based on Optimal Migration // *Bull Math Biol*. – 2016. – Vol. 78. – P.736–753 (<http://link.springer.com/article/10.1007/s11538-016-0159-z>, DOI: 10.1007/s11538-016-0159-z).
7. Korobko A.A., Nozhenkova L.F. Application of Model Driven Architecture for Development of Data Consolidation Web-system // *Proc. 39th International Convention. (MIPRO 2016)*. – 2016. – P. 1729-1734.
8. Lapko A.V., Lapko V.A., Yuronen E.A. Testing of hypothesis of random variables independence on the basis of nonparametric method of confidence estimation of probability density // *IOP Conference Series: Materials Science and Engineering* – (в печати).
9. Lapko A.V., Lapko V.A., Yuronen E.A. Testing of hypothesis of two-dimensional random variables independence on the basis of algorithm of pattern recognition // *IOP Conference Series: Materials Science and Engineering* (в печати).

10. Lapko V.A., Lapko A.V., Yuronen Yu.P. Synthesis of hybrid systems of pattern recognition on the basis of procedure of consecutive correction of decision functions // IOP Conference Series: Materials Science and Engineering (в печати).
11. Sadovsky M.G., Putinseva Yu.A., Birukov V.V., Krutovsky K.V. De novo assembly and cluster analysis of Siberian larch transcriptome and genome // LNBI, 2016. – Vol. 9656. – P.455-464.
12. Sadovsky M.G., Salmina A.B., Morgun V.A., Pozhilenkova E.A., Khilazheva E.V., Kuvacheva N.V. Enriched environment affects positively a progression of neurodegeneration: elastic maps-based analysis // LNBI. – 2016. – Vol.9656. – P. 505-514.
13. Tsarev S.P., Sadovsky M.G. New error tolerant method for search of long repeats in DNA sequences // LNBI, 2016. – Vol.9702. – P.171-182.
14. Лапко А.В., Лапко В.А. Выбор оптимального количества интервалов дискретизации области значений двумерной случайной величины // Измерительная техника. – 2016. – №2. – С. 14-17.
15. Лапко А.В., Лапко В.А. Непараметрическая оценка плотности вероятности парзеновского типа с неявно заданной формой ядерной функции // Измерительная техника. – 2016. – №6. – С. 14-17.
16. Белорусов А.И., Ноженков А.И. Формирование XML-документов на основе интерпретируемых шаблонов // Информатизация и связь. – 2016. – №3. – С.105-110.
17. Жучков Д.В. Информационная модель унифицированного справочника // Информатизация и связь. – 2016. – №3. – С. 133-137.
18. Лапко А.В., Лапко В.А. Анализ эффективности методов дискретизации области значений двумерной случайной величины при синтезе непараметрической оценки плотности вероятности // Информатика и системы управления. – 2016. – №3 (49). – С. 78-85 (DOI: 10.22250/isu.2016.49.78-85) (http://media/2016/N49_09.pdf)
19. Лапко А.В., Лапко В.А. Гибридные системы распознавания образов в условиях неоднородных данных // Информатика и системы управления. – 2016. – №1(47). – С. 73-81 (http://media/2016/N47_09.pdf).
20. Кононов Д.Д., Исаев С.В. Расширенная ролевая модель безопасности, основанная на иерархии путей // Вопросы защиты информации. 2016. – №4 (в печати).
21. Лапко А.В., Лапко В.А. Статистические модели пространственно распределённых временных процессов // Информатика и системы управления. – 2016. – №4 (50) (в печати).
22. Вогоровский Р.В., Ноженкова Л.Ф. Формализация процессов испытаний бортовой аппаратуры космического аппарата // Образовательные ресурсы и технологии. – 2016. – № 2(14). – С. 131-138.

23. Евсюков А.А., Коробко А.В. Применение методов инфографического моделирования для представления многомерных данных // Образовательные ресурсы и технологии. – 2016. – №2(14) – С. 16-23.
24. Евсюков, А.А. Инфографическое моделирование многомерных данных // Современные информационные технологии и ИТ-образование. – 2016. – Т. 12. №2. – С. 127-133.
25. Исаев С.В. Анализ киберугроз и их источников для корпоративной сети Красноярского научного центра СО РАН // Информационные и математические технологии в науке и управлении. Иркутск: ИСЭМ СО РАН. – 2016. – № 4. – С. 76-85.
26. Колдырев А.Ю., Грузенко Е.А. Формирование заданий в учебно-исследовательской системе конструктора бортовой аппаратуры // Образовательные ресурсы и технологии. – 2016. – №2(14). – С. 30-35.
27. Коробко А.А. Алгоритм генерации программных компонентов модельно-ориентированной системы // Образовательные ресурсы и технологии – 2016. – №2 (14). – С. 173-180.
28. Ничепорчук В.В. Информационное обеспечение управления природно-техногенной безопасностью // Сибирский пожарно-спасательный вестник, 2016. – №1. – С.49-54.
29. Ничепорчук В.В. Системный анализ информационных потоков органов управления РСЧС регионального уровня // Сибирский пожарно-спасательный вестник, 2016. – №2. – С.41-46.
30. Ничепорчук В.В. Формирование базы показателей природно-техногенной безопасности территорий Красноярского края // Образовательные ресурсы и технологии, 2016. – №2(14). – С. 226-234.
31. Ничепорчук В.В., Ноженков А.И. Интеграция технологий в системе комплексного мониторинга чрезвычайных ситуаций // Образовательные ресурсы и технологии, 2016. – №2(14). – С. 281-287.
32. Ноженкова Л.Ф., Исаева О.С., Вогоровский Р.В., Грузенко Е.А. Автоматизация испытаний параметров и логики функционирования командно-измерительной системы // Исследования наукограда. – 2016 (в печати).
33. Сенашова М.Ю., Садовский М.Г. Семикластерная структура геномов хлоропластов отражает филогению их носителей // Международный журнал прикладных и фундаментальных исследований. – 2016. – №12(7) (в печати).
34. Шурер К., Пенькова Т. Методы стандартизации и классификации записей о месте рождения и профессиональной деятельности в данных переписи Великобритании 1851-1911 гг. // Историческая информатика. Информационные технологии и

- математические методы в исторических исследованиях и образовании. – 2016 (в печати).
35. Исаев С.В. Идентификация и визуализация источников интернет-угроз/ Материалы XX Юбилейной междунар. науч.-практ. конф., посвящ. памяти генерального конструктора ракетно-космических систем академика М. Ф. Решетнева (09-12 нояб. 2016, г. Красноярск): в 2 ч. / под общ. ред. Ю. Ю. Логинова; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2016. – Ч. – С. 260-262.
 36. Исаев С.В., Кулясов Н.В. Исследование взаимосвязи атак на различные интернет-сервисы // Материалы XX Юбилейной междунар. науч.-практ. конф., посвящ. памяти генерального конструктора ракетно-космических систем академика М. Ф. Решетнева (09-12 нояб. 2016, г. Красноярск) : в 2 ч. / под общ. ред. Ю.Ю. Логинова; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2016. – Ч. 2. – С. 258-260.
 37. Лапко А.В., Лапко В.А., Юронен Е.А. Проверка гипотезы о независимости двумерных случайных величин на основе алгоритма распознавания образов // Материалы XX Юбилейной междунар. науч.-практ. конф., посвящ. памяти генерального конструктора ракетно-космических систем академика М. Ф. Решетнева (09-12 нояб. 2016, г. Красноярск) : в 2 ч. / под общ. ред. Ю.Ю. Логинова; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2016. – Ч. 2. – С. 50-52. (<https://disk.sibsau.ru/index.php/s/UGP7UcMvdmYV7o>)
 38. Лапко А.В., Лапко В.А., Юронен Е.А. Проверка гипотезы о распределениях случайных величин на основе непараметрического метода доверительного оценивания плотности вероятности // Материалы XX Юбилейной междунар. науч.-практ. конф., посвящ. памяти генерального конструктора ракетно-космических систем академика М. Ф. Решетнева (09-12 нояб. 2016, г. Красноярск) : в 2 ч. / под общ. ред. Ю.Ю. Логинова; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2016. – Ч. 2. – С. 52-54. (<https://disk.sibsau.ru/index.php/s/UGP7UcMvdmYV7o>)
 39. Ничепорчук В.В. Системная модель оперативного комплексного мониторинга / Мониторинг, моделирование и прогнозирование опасных природных явлений и чрезвычайных ситуаций: Сборник статей по материалам VI Всероссийской научно-практической конференции, Железногорск, 2016. – С. 24-28.
 40. Ничепорчук В.В., О.В. Тасейко Концепция создания региональной системы мониторинга рисков // Материалы XX Юбилейной междунар. науч.-практ. конф., посвящ. памяти генерального конструктора ракетно-космических систем академика М. Ф. Решетнева (09-12 нояб. 2016, г. Красноярск) : в 2 ч. / под общ. ред. Ю.Ю. Логинова; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2016. – Ч. 2. – С.294-296.

41. Ковязина Е.В. Архив открытого доступа как средство продвижения научных публикаций // Информатизация образования и методика электронного обучения: материалы I Международной научной конференции в рамках IV Международного форума «Человек, семья и общество: история и перспективы развития» (Красноярск, 27-30 сентября 2016 г.) – Красноярск : Сиб. федер. ун-т, 2016. – С. 134-139. (<http://lib3.sfu-kras.ru/ft/LIB2/ELIB/b74/free/i-599417.pdf>)
42. Ковязина Е.В. Корпоративные репозитории научных публикаций и проблемы обмена данными // Труды ГПНТБ СО РАН. – 2016. – Вып. 10. – С. 288-292. (<http://elibrary.ru/item.asp?id=26743821>)
43. Ковязина Е.В. Открытый архив в научной библиотеке [Электронный ресурс] // Материалы Второго Международного профессионального форума «Книга. Культура. Образование. Инновации» (“Крым-2016”), 4-12 июня 2016 г., Судак, Республика Крым, Россия. – Электрон. дан. – Москва : ГПНТБ России : Ассоциация ЭБНИТ, 2016. – 1 электрон. опт. диск (CD-ROM). – Систем. требования: IBM PC, Windows 2000 или выше. – Загл. с этикетки диска. – ISBN 978-5-85638-194-7. (<http://www.gpntb.ru/win/inter-events/crimea2016/disk/036.pdf> - 4 стр.)
44. Ничепорчук В.В. Оценивание территориальных рисков с использованием данных мониторинга // Безопасность критичных инфраструктур и территорий: Материалы VII Всероссийской конференции и XVII школы молодых учёных. – Екатеринбург: УрО РАН, 2016. – С. 151-154.
45. Ничепорчук В.В. Принципы создания информационно-аналитических систем поддержки управления природно-техногенной безопасностью // Сборник материалов XXVI Международной научно-практической конференции «Предупреждение. Спасение. Помощь». Секция № 10 «Моделирование сложных процессов и систем». – Химки, 2016. – С. 12-17.
46. Ничепорчук В.В. Реализация системного интегратора центра мониторинга и прогнозирования ЧС // Материалы двадцать пятой международной научно-технической конференции «Системы безопасности – 2016», М.: Академия ГПС МЧС России, 2016. – С. 152-155.

ПРИЛОЖЕНИЕ Б

Количество публикации по тематике проекта за весь период выполнения проекта по годам

индикатор	Ед. измерения	Всего за 2013-2016 гг.	2013	2014	2015	2016
Количество публикаций в ведущих российских и международных журналах по результатам исследований, полученных в процессе реализации проекта	единиц	86	15	12	32	27
Количество публикаций в мировых научных журналах, индексируемых в базе данных «Сеть науки» (WEB of Science)	единиц	17	2	5	5	5
Число тезисов в конференциях	единиц	103	46	14	30	12
Число охраняемых объектов интеллектуальной собственности зарегистрированных в России	единиц	7	2	5	0	0

ПРИЛОЖЕНИЕ В

Выписка из плана научно-исследовательской работы на 2016 год

Содержание работы	Планируемый результат выполнения работы
1. Разработка методов интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата	1. Методы интеллектуальной поддержки моделирования и анализа функционирования бортовой аппаратуры командно-измерительной системы космического аппарата.
2. Разработка моделей и методов консолидации и анализа данных и технологий создания интегрированных систем территориального и корпоративного управления	2. Модели консолидации данных гетерогенных источников в соответствии с форматами хранения. Модели и алгоритмы анализа структурных характеристик территорий. Алгоритмы построения управляющей модели для разработки модельно-ориентированного программного обеспечения. Технологии создания интегрированных систем территориального и корпоративного управления.
3. Исследование моделей управления безопасностью в информационно-телекоммуникационных системах.	3. Модели управления безопасностью в информационно-телекоммуникационных системах.
4. Разработка методов доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей на основе регрессионной оценки плотности вероятности.	4. Новые методы доверительного оценивания законов распределения случайных величин, решающих функций в задачах распознавания образов и восстановления стохастических зависимостей.
5. Последовательный анализ созданных ранее баз генетических данных с целью выявления в них структурированности и порядка.	5. Выявление корреляции между кластерами, выделяемыми по статистическим и комбинаторным свойствам нуклеотидных последовательностей и соответствующими биологическими объектами.