Сведения о ходе выполнения проекта по соглашению
№ 14.604.21.0147 от «14» ноября 2014 г.

Тема: Разработка методов агрегации, нормализации, анализа и визуализации больших массивов гетерогенных структурированных, полуструктурированных и неструктурированных данных для мониторинга и управления безопасностью распределенной сети электронных потребительских устройств (internet of things).

Целью проекта является создание комплекса научных/научно-технических решений в области разработки методов и алгоритмов, обеспечивающих повышение эффективности научных исследований посредством агрегации, нормализации, анализа и визуализации больших массивов гетерогенных структурированных, полуструктурированных и неструктурированных данных в распределенной сети электронных потребительских устройств (internet of things).

На **первом этапе** ПНИ проводился аналитический обзор современной научно-технической, нормативной, методической литературы, затрагивающей исследуемую научно-техническую проблему. Исследовались особенности построения современных сетей «Интернет вещей» и особенности обеспечения безопасности в этих сетях. Анализировались достигнутые научно-технические результаты в области агрегации, нормирования, корреляции и визуализации событий безопасности в сетях «Интернет вещей», базирующиеся на концепции больших данных. На основе полученных результатов анализа были сделаны выбор и обоснование направления исследования, а также разработаны основные принципы и методические подходы в области создания средств агрегации, нормализации, анализа и визуализации больших массивов данных для мониторинга и управления безопасностью сети «Интернет вещей». Учитывая полученные требования по хранению данных, были разработаны методические подходы к созданию хранилища данных, обеспечивающего гибридное онтологическое представление больших массивов гетерогенных данных для мониторинга и управления сетью «Интернет вещей», которые касаются архитектуры и реализации хранилища данных о событиях безопасности.

Новыми результатами, полученными в ходе исследований на первом этапе, являются основные принципы и методические подходы в области создания средств агрегации, нормализации, анализа и визуализации больших массивов структурированных, полуструктурированных и неструктурированных гетерогенных данных для мониторинга и управления безопасностью распределенной сети электронных потребительских устройств (internet of things), а также методические подходы к созданию хранилища данных, обеспечивающего гибридное онтологическое представление этих данных.

На **втором этапе** ПНИ разрабатывались математические методы и алгоритмы агрегации больших массивов данных, в том числе с применением приемов и методов параллельных вычислений, математические методы и алгоритмы нормализации и анализа больших массивов гетерогенных данных, поступающих синхронно и асинхронно от распределенной сети источников, а также методические подходы к созданию протоколов для синхронной и асинхронной передачи данных от потребительских устройств к центрам обработки данных. В состав разработанных методов и алгоритмов агрегации данных вошли методы и алгоритмы формирования последовательностей операторов потоковой обработки данных для расчета мер центральной тенденции и экстремумов, методы и алгоритмы распараллеливания операторов потоковой обработки данных, методы и алгоритмы распределения операторов потоковой обработки по узлам транспортной сети, а также методы и алгоритмы агрегации больших массивов данных для мониторинга и управления безопасностью сетей «Интернет вещей». Разработанные методы и алгоритмы нормализации и анализа данных основаны на правило-ориентированном, шаблонно-ориентированном и нейросетевом подходах. Разработанные методические подходы к созданию протоколов для синхронной и асинхронной передачи данных включают подходы к управлению блоками данных и подходы к управлению потоками данных.

Новыми результатами, полученными в ходе исследований, являются методы и алгоритмы распараллеливания операций агрегации и анализа больших массивов гетерогенных

данных о безопасности сети «Интернет» за счет рационального распределения потоковых операторов по узлам сети, усовершенствованный генетический алгоритм с мультихромосомным кодированием решений оптимизационной задачи, нейросетевой метод анализа состояния элементов сети «Интернет вещей» на основе комбинированной искусственной нейронной сети и методические подходы к созданию протоколов для синхронной и асинхронной передачи данных в сети «Интернет вещей».

На **третьем этапе** ПНИ разрабатывались математические методы и алгоритмы визуализации больших массивов гетерогенных структурированных, полу-структурированных и неструктурированных данных и технические принципы и методические подходы к организации и развертыванию решений по агрегации, нормализации, анализу и визуализации больших массивов гетерогенных структурированных, полуструктурированных и неструктурированных данных в различных средах. В состав разработанных методов и алгоритмов визуализации больших массивов гетерогенных вошли методы и алгоритмы визуализации с использованием GIS-систем, с помощью диаграмм и графиков, инфографические методы визуализации, табличные методы визуализации и иерархические методы визуализации, реализующие технологии кластеризации, автоматической классификации и «ленивой» загрузки массивов данных, а также детализации кластеризованных данных (Drill-down), перехода между связанными данными и микширования данных (mesh-up). Разработанные технические принципы и методические подходы определяют концептуальные решения по реализации процессов сбора и обработки информации в распределенной сети электронных потребительских устройств (internet of things).

Новыми результатами, полученными в ходе исследований, являются математические методы и алгоритмы визуализации больших массивов гетерогенных данных, реализованные с применением принципов потоковой обработки, а также методы и алгоритмы классификации, применимые для обработки полуструктурированных и неструктурированных данных.

На **четвертом этапе** ПНИ проводились экспериментальные исследования поставленных перед ПНИ задач. Разрабатывался экспериментальный образец программного обеспечения (ЭО ПО) для агрегации, нормализации, анализа и визуализации данных для мониторинга и управления безопасностью распределенной сети электронных потребительских устройств (internet of things). В состав ЭО ПО вошли модуль сбора данных, модуль агрегации данных, модуль нормализации и анализа данных, модуль визуализации и трансляции данных и модуль работы с хранилищем данных. Разработаны Программа и методика экспериментальных исследований разработанного ЭО ПО, позволяющие подтвердить соответствие характеристик ЭО ПО предъявляемым требованиям в условиях, максимально приближенных к условиям реальной эксплуатации, а также оценить возможность реализации ЭО ПО. Разработана программная документация на ЭО ПО, включающая описания и тексты программ всех модулей ЭО ПО, описание применения ЭО ПО и описание логической и физической структуры базы данных. Проводились экспериментальные исследования ЭО ПО с применением тестовых гетерогенных данных, учитывающие в качестве входных данных как данные, сгенерированные программно-инструментальным стендом, так и данные, отражающие реальный трафик в компьютерной сети, загруженные из внешней базы данных.

Новыми результатами, полученными в ходе исследований, является ЭО ПО, реализующий технологию параллельной потоковой предварительной обработки данных о событиях безопасности в сетях «Интернет вещей» и методика проверки работоспособности ЭО ПО, учитывающая предложенный показатель суммарных потерь в активах при ограничениях на вероятности реализации угроз и на время реализации мероприятий защиты сети.