Проект РНФ No 21-71-20078

# Аналитическая обработка больших массивов гетерогенных данных о событиях кибербезопасности в интересах оценки состояния, поддержки принятия решений и расследования компьютерных инцидентов в критически важных инфраструктурах

## Описание выполненных в 2022 году работ и полученных научных результатов

1. Разработаны методы, модели, методики и алгоритмы обнаружения в реальном времени атак на основе имитационного и графо-ориентированного моделирования. Разработанный метод обнаружения атак основан на комбинировании аналитического, имитационного и графо-ориентированного моделирования, машинного обучения, искусственных нейронных сетей и технологии больших данных. Метод предполагает возможности распараллеливания процедур обнаружения атак с разбиением по различным сценариям, видам атак, процедурам построения и обработки графов. Модельно-алгоритмическая часть метода специфицирует процедуры, используемые для описания анализируемой системы и ее характеристик, построения графов состояний, определения переходов между состояниями, а также для классификации и прогнозирования будущих состояний.

2. Разработаны методы, модели, методики и алгоритмы обнаружения в реальном времени аномальной активности и нарушений критериев и политик безопасности на основе аналитической обработки больших массивов гетерогенных данных о событиях кибербезопасности. Методы обнаружения предусматривают реализацию четырех этапов: предобработки входных данных, классификации объекта, представленного вектором или матрицей атрибутов, формирования решения о статусе объекта и формирования объяснения решений, сделанных на предыдущем этапе. Разработана модельно-алгоритмическая часть методов, которая включает в себя формальную модель выявления аномальной активности и нарушений критериев и политик безопасности на основе аналитической обработки больших массивов гетерогенных данных, модели данных от разнородных источников, в частности, от технологических процессов киберфизических объектов и политик безопасности, и алгоритмы предобработки входных данных в зависимости от типа их источника. Разработана формальная модель объяснений решений, а также алгоритмы их формирования с учетом требований к времени их генерации.

3. Разработаны методы, модели, методики и алгоритмы оперативной оценки защищенности информационных, телекоммуникационных (ИТР) и других критически важных ресурсов (КВР) на основе аналитической обработки больших массивов гетерогенных данных. В качестве основы разработанных моделей и методов используются аналитические выражения для количественного оценивания рисков защищенности, путей высокого риска защищенности (с учетом имеющихся уязвимостей) и максимально возможного риска с учетом максимально возможных уязвимостей. В состав алгоритмов оперативной оценки защищенности включены алгоритмы построения модели сети, построения графа рисков, выделения путей риска, оценки вероятности и воздействия путей риска, оценки сетевых рисков и выбора пути высокого риска. Разработанная методика оперативной оценки защищенности ресурсов ориентирована на совместное выполнение вышеперечисленных алгоритмов. При этом

рассчитываются следующие показатели: путь высокого риска для конкретного хоста; путь высокого риска для конкретного злоумышленника; путь высокого риска для конкретной точки входа; путь высокого риска для всей совокупности ресурсов.

4. Разработаны методы, модели, методики и алгоритмы оперативного анализа и управления рисками информационной безопасности на основе аналитической обработки больших массивов гетерогенных данных о событиях кибербезопасности в интересах оценки состояния, поддержки принятия решений и расследования инцидентов. Методы отличаются вычислением уровня риска на основе данных об атаках и аномалиях, обнаруженных при совместном анализе сетевого трафика и журналов событий, и применением полученных оценок для прогнозирования развития кибератак с учетом этапа проведения атаки, к которому относится обнаруженное атакующее действие. Разработанная модельно-алгоритмическая часть подхода включает модель анализируемой киберфизической системы, модель атак, модель распространения ущерба в анализируемой системе, а также методики их построения на основе анализа данных журналов событий и сетевого трафика и алгоритмы вычисления уровня риска.

5. Разработаны методы, модели, методики и алгоритмы оперативной визуализации больших массивов гетерогенных данных о событиях кибербезопасности в интересах оценки состояния, поддержки принятия решений и расследования инцидентов. Методика оперативной визуализации больших массивов гетерогенных данных основана на решении задач, описанных на первом этапе проекта, которые возникают при попытке визуализировать большие объемы данных безопасности. При визуализации ставится цель снижения количества объектов, отображаемых на экране с помощью методов агрегации отображаемых измерений и объектов, а также аппроксимации объектов. Решение о применении отдельных алгоритмов и моделей принимается исходя их объема данных, количества измерений и объектов. Решение принимается оператором системы экспериментальным путем. Предложено использовать алгоритмы агрегации измерений (PCA, UMAP, анализа мультиколлинеарности), объектов (K-means, spectral, поиск сообществ на основе модулярности и распространения метки), аппроксимации (KDE, гексогональные решетки) и модели визуализации (столбчатые графики, графики рассеивания, трилинейные координаты, линейные графики, счетчики, параллельные координаты, силовые графы, карты деревьев, радиальные графы, и матрицы).

6. Разработаны методы, модели, методики и алгоритмы принятия решений по защите информационных, телекоммуникационных и других критически важных ресурсов на основе аналитической обработки больших массивов гетерогенных данных о событиях кибербезопасности в интересах оценки состояния, поддержки принятия решений и расследования инцидентов. Методы принятия решений отличаются выделением уровней принятия решений в зависимости от доступного времени, возможности автоматической реализации мер защиты и необходимых для реализации мер защиты прав доступа. Кроме того, обеспечивается проактивное и реактивное принятие решений в зависимости от этапа реализации атаки. Разработанные алгоритмы выбора оптимального набора защитных мер используют результаты обнаружения атак и аномалий, а также анализа рисков. Модельно-алгоритмическая часть методов включает модель защитной меры и иерархическую модель принятия решений, а также методику и алгоритм выбора защитных мер, основанный на решении задач многокритериальной оптимизации.

7. Дополнительно к плану разработана архитектура и функциональная структура системы аналитической обработки больших массивов гетерогенных данных о событиях кибербезопасности. В функциональной структуре системы выделены служебные компоненты, такие как брокер сообщений, брокер потоков, компоненты оперативного и долговременного хранения, компонент менеджмента и управления вычислениями. Исследованы алгоритмы управления потоками для анализа больших массивов гетерогенных данных. Проведено тестирование существующих информационных технологий обработки Больших данных. Проведен сравнительный анализ фреймворков анализа Больших данных, поддерживаемых ими моделей защищенности, алгоритмов и методов управления потоками, а также возможностей разработки системы управления задачами распараллеливания вычислений. Реализован прототип, позволяющий оценить возможность запуска данных технологий на отечественных операционных системах. Предварительно разработаны и протестированы экспериментальные стенды «Умный дом», «Умный транспорт» и «Офис».

Результаты исследований опубликованы в 20 статьях, индексируемых в WoS и Scopus (среди них 5 статей Q1), в одной статье, индексируемой в RSCI, и 21 статье и тезисах докладов, индексируемых в РИНЦ.

При выполнении проекта получены исключительные права на результаты интеллектуальной деятельности (РИД): 1 патент на изобретение, 7 свидетельств о государственной регистрации программ для ЭВМ и 1 свидетельство о государственной регистрации базы данных.

Члены коллектива участвовали в апробации результатов на 13 российских и международных конференциях и семинарах.

URL: http://comsec.spb.ru/ru/projects/
URL: http://comsec.spb.ru/en/projects/