На правах рукописи

ГАВРИЛОВ Григорий Николаевич

СИСТЕМА ОБНАРУЖЕНИЯ ВРЕДОНОСНЫХ ПРОГРАММ В ОПЕРАЦИОННОЙ СИСТЕМЕ (ОС) ДЛЯ МОБИЛЬНЫХ УСТРОЙСТВ (НА ПРИМЕРЕ ANDROID) С ПРИМЕНЕНИЕМ ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ

Специальность 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени кандидата технических наук

Уфа – 2017

Работа выполнена в ФГБОУ ВО «Уфимский государственный авиационный технический университет» на кафедре электроники и биомедицинских технологий

Научный руководитель: доктор технических наук, профессор,
Жернаков Сергей Владимирович
ФГБОУ ВО «Уфимский государственный авиационный технический университет» заведующий кафедрой электроники и биомедицинских технологий

Официальные оппоненты: доктор технических наук, профессор,
Маркевич Олег Борисович
ФГАОУ ВО «Южный федеральный университет» главный научный сотрудник кафедры безопасности информационных систем

кандидат технических наук, доцент,
Аникин Игорь Вячеславович
ФГБОУ ВО «Казанский национальный исследовательский технический университет им. А. Н. Туполева – КАИ»
заведующий кафедрой систем информационной безопасности

Ведущая организация: ФГАОУ ВО «Самарский национальный исследовательский университет им. академика С.П. Королева» (г. Самара)

Защита диссертации состоится 29 сентября 2017 г. в 12^{00} часов на заседании диссертационного совета Д 212.288.07 на базе ФГБОУ ВО «Уфимский государственный авиационный технический университет» по адресу: 450008, г. Уфа, ул. К. Маркса, 12.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Уфимский государственный авиационный технический университет» и на сайте *www.ugatu.su*.

Автореферат разослан «___» _____ 20___ года.

Ученый секретарь
диссертационного совета
д-р техн. наук, доцент И. Л. Виноградова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

В настоящее время мобильные устройства являются неотъемлемой частью повседневной жизни. Они используются для выполнения многочисленных операций и хранения личной информации. По статистике Википедии, ОС Android работает на 64 % устройств. Существенные функциональные возможности и наличие личной информации служат причиной, по которой злоумышленники заинтересованы получить неправомерный доступ к мобильному устройству с целью получения выгоды. Данная ОС обладает хорошо организованными защитными механизмами, но имеет ряд уязвимостей, что позволяет вредоносным программам получить несанкционированный доступ к мобильному устройству.

По данным статистики “Лаборатории Касперского”, число вредоносных программ в ОС Android превысило 12 миллионов в 2014 году, что более чем в десять раз выше, чем в 2012 году. Такой рост числа вредоносных программ обусловлен ростом популярности, функциональных возможностей и объема личной информации в процессе использования мобильной ОС. Антивирусные программы работают на основе хорошо изученных признаков вредоносных программ – сигнатур, которые хранятся в базе антивирусных сигнатур как эталоны и с которыми в дальнейшем осуществляется сравнение для последующего обнаружения вредоносных программ. Если в текущий момент времени сигнатура той или иной вредоносной программы отсутствует в антивирусной базе, то антивирусная программа ее не обнаружит, и вредоносная программа может длительное время существовать в мобильной ОС. Процедура получения сигнатуры вредоносной программы требует определенных затрат времени и тщательного её анализа с целью выявления основных признаков вредоносной программы, а также последующего добавления этой сигнатуры в вирусную базу и тиражирования её на все подобные устройства. Сегодня с развитием сети Интернет скорость распространения различной информации и программного обеспечения значительно увеличилась, следовательно, новые или модифицированные вредоносные программы могут распространиться за небольшое количество времени на множество мобильных ОС и нанести огромный ущерб. В связи с этим в диссертационной работе была поставлена задача повысить эффективность обнаружения вредоносных программ в ОС для мобильных устройств Android на основе интеллектуальных технологий путем разработки соответствующей системы обнаружения вредоносных программ.

Степень разработанности темы

В настоящее время в данной предметной области ведутся активные разработки, о чем свидетельствуют работы ведущих отечественных и зарубежных исследователей: В. И. Васильева, И. В. Машкиной, С. С. Валеева, В. Д. Котова, О. Б. Макаревича, И. В. Аникина, Fan Yuhui, Xu Ning, Borja Sanz, Igor Santos, Javier Nieves, Carlos Laorden, Inigo Alonso-Gonzalez, Pablo G. Bringas, Daniel Arp, Michael Spreitzenbarth, Malte Hubner, Hugo Gascon, Konrad Rieck, Pehlivan U., Muller K., Mika S. Анализ отечественных и зарубежных публикаций по данной

тематике показывает, что такие исследования активно ведутся, однако в работах отсутствуют практические рекомендации, а также качественные и количественные характеристики разработанных программных проектов для систем комплексной защиты средств мобильной связи ОС Android. Поэтому тема диссертационной работы, посвященная повышению эффективности обнаружения вредоносных программ на основе интеллектуальных методов, является актуальной.

Объект исследования – обнаружение вредоносных программ в ОС для мобильных устройств (на примере Android).

Предмет исследования – методы и алгоритмы обнаружения вредоносных программ в ОС для мобильных устройств (на примере Android) на основе интеллектуальных технологий.

Цель работы – повышение эффективности обнаружения вредоносных программ в ОС для мобильных устройств (на примере Android) путем разработки моделей и алгоритмов на основе интеллектуальных технологий.

Задачи исследования

Для достижения поставленной цели в работе были поставлены и решены следующие задачи:

1. Исследовать работу ОС для мобильных устройств Android с точки зрения защищенности, привести перечень угроз и уязвимостей, также рассмотреть существующие встроенные и сторонние средства защиты информации.
2. Разработать системные модели процесса функционирования системы обнаружения вредоносных программ, исследовать множество прикладных программ и сформировать перечень особенностей их поведения.
3. Разработать алгоритмы обнаружения вредоносных программ в ОС для мобильных устройств Android с применением интеллектуальных технологий.
4. Предложить архитектуру интеллектуальной системы обнаружения вредоносных программ, провести вычислительные эксперименты с целью оценки эффективности предложенных алгоритмов.
5. Разработать программное обеспечение исследовательского прототипа системы обнаружения вредоносных программ в мобильной ОС Android.

Научная новизна

Научная новизна работы заключается в следующем:

1. Предложена и обоснована экспериментальная выборка, на основе комплексного анализа параметров ОС Android, позволяющая описать поведенческий характер потенциальной вредоносной программы и на основе исследования множества прикладных программ классифицировать их, отличающаяся полнотой представления, а также качественными и количественными характеристиками.
2. Разработаны системные модели процесса функционирования программного комплекса обнаружения вредоносных программ для ОС Android на основе SADT-методологии и IDEF-технологий, позволяющие интегрировать систему обнаружения вредоносных программ с компонентами внутренних и внешних механизмов защиты ОС Android, отличающиеся полнотой

представления комплексных процессов защиты информации для ОС Android с учетом факторов неопределенности.

3. Разработан комплекс алгоритмов обнаружения вредоносных программ для ОС Android на основе гибридных интеллектуальных технологий, позволяющий с высокой точностью выполнять обнаружение вредоносных программ, отличающийся от классического сигнатурного метода динамическим анализом поведения вредоносной программы с использованием нейро-нечетких алгоритмов.

4. Предложена архитектура системы обнаружения вредоносных программ, на основе разработанных алгоритмических и программных средств, которая позволила идентифицировать как известные, так и модифицированные и новые вредоносные программы для ОС Android, отличающаяся оперативностью и качеством принимаемых решений по обнаружению вредоносных программ.

Практическая значимость

Практическая значимость разработанной системы обнаружения вредоносных программ заключается в применении комплекса данных алгоритмов с целью увеличения эффективности обнаружения вредоносных программ, а также в применении их в качестве дополнительного средства защиты к уже имеющимся методам обнаружения вредоносных программ в ОС для мобильных устройств Android.

Методы исследования

В процессе исследования использовались теория вероятностей и методы математической статистики, методы иерархической кластеризации, метод к-средних, факторный анализ, кластерный анализ, дискриминантный анализ, нейросетевые методы, машина опорных векторов, аппарат нечеткой логики.

Защищаемые положения

1. Результаты анализа защищенности ОС для мобильных устройств Android, встроенных и сторонних средств защиты информации, структуры прикладных программ и перечень угроз и уязвимостей.

2. Комплекс моделей функционирования системы обнаружения вредоносных программ в ОС для мобильных устройств Android и экспериментальная выборка, описывающая поведение двух типов прикладных программ.

3. Комплекс алгоритмов обнаружения вредоносных программ на основе машины опорных векторов и нечеткой логики.

4. Разработанный в виде программы исследовательский прототип системы обнаружения вредоносных программ в ОС для мобильных устройств Android.

Достоверность результатов

Достоверность научных положений и выводов и обоснованность полученных в диссертационной работе результатов подтверждается корректной постановкой задач, строгостью применяемого математического аппарата, результатами математического моделирования и апробации комплекса алгоритмов и программы, реализующих предложенные подходы обнаружения вредоносных программ.

Личный вклад

Все исследования, изложенные в диссертационной работе, проведены автором в процессе научной деятельности. Результаты, выносимые на защиту, получены автором лично, заимствованный материал обозначен в работе ссылками.

Апробация результатов

По теме диссертации опубликовано 17 научных статей и тезисов докладов, из них 6 статей в изданиях, рекомендованных ВАК. Имеется свидетельство о государственной регистрации программы для ЭВМ.

Основные положения, представленные в диссертационной работе, докладывались и обсуждались на следующих конференциях:

- Международная научно-практическая конференция “Интеграционные процессы науки XXI века”, г. Стерлитамак, 2015 г.,
- Международная научно-практическая конференция “Научные перспективы XXI века”, г. Нефтекамск, 2015 г.,
- XIII Международная научно-практическая конференция “Научные перспективы XXI века. Достижения и перспективы нового столетия”, г. Новосибирск, 2015 г.,
- XV Международная научно-практическая конференция “Научное обозрение физико-математических и технических наук в XXI веке”, г. Москва, 2015 г.,
- XVII Международная научно-практическая конференция “Актуальные вопросы развития инновационной деятельности в новом тысячелетии”, г. Новосибирск, 2015 г.,
- Международная молодежная научная конференция “XXII Туполевские чтения (школа молодых ученых)”, г. Казань, 2015 г.,
- XVI Международная научно-техническая конференция “Проблемы техники и технологии телекоммуникаций”, г. Уфа, 2015 г.,
- IX Всероссийская молодежная научная конференция “Мавлютовские чтения”, г. Уфа, 2015 г.,
- XV Международная научная конференция “Перспективы направления развития современной науки”, г. Москва, 2016 г.,
- VIII Международная научно-практическая конференция “Актуальные проблемы науки XXI века”, г. Москва, 2016 г.,
- XII Международная научно-техническая конференция «Актуальные проблемы электронного приборостроения», г. Новосибирск, 2016 г.

Результаты диссертационного исследования внедрены в производственный процесс ЗАО «Республиканский центр защиты информации», ОАО «Инфотекс Интернет Траст» в и учебный процесс кафедры «Вычислительная техника и защиты информации» ФГБОУ ВО «Уфимский государственный авиационный технический университет».

Объем и структура работы

Диссертационная работа включает введение, четыре главы основного материала, заключение, приложения А, Б, В, Г и список литературы. Работа

изложена на 195 страницах машинописного текста, список литературы включает 115 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении к диссертации дается краткая характеристика работы, сформулированы ее цели и задачи, обоснована актуальность исследований в данной предметной области, показана научная новизна и практическая ценность полученных результатов.

В первой главе выполнен обзор современного состояния защищенности в мобильных средствах связи. Выявлены существующие в настоящее время угрозы и уязвимости, такие как вредоносные программы, система разрешений, человеческий фактор, модифицированная мобильная ОС. Рассмотрены встроенные средства защиты, такие как песочница, система разрешений, межпроцессное взаимодействие, подпись кода и платформы ключей, а также структура программ и архитектура мобильной ОС Android. В результате выполненной работы сделан вывод о том, что мобильная ОС Android содержит в своем арсенале надежные средства защиты, однако имеется перечень уязвимостей, в частности вредоносные программы, против которых они малоэффективны. В связи с этим мотивацией дальнейшего исследования в данной области является повышение эффективности обнаружения вредоносных программ в мобильной ОС Android. Ставятся цели и задачи исследования.

Во второй главе разработана функциональная модель IDEF0, которая выделяет и обосновывает основные функции системы обнаружения вредоносных программ. Система обнаружения вредоносных программ в ОС для мобильных устройств представляет собой решение, реализованное в виде программы (на уровне ОС), функционирующей по определенным правилам, инструкциям и требованиям. Основными компонентами программы являются файлы androidmanifest.xml, classes.dex, а также информация о системных вызовах. Управляющими стрелками на IDEF-диаграммах являются правила, условия и требования, согласно которым регулируется деятельность и порядок работы исследовательского прототипа системы обнаружения вредоносных программ в ОС для мобильных устройств. На выходе блока поддержки и принятия решений находятся результаты классификации, выполненные машиной опорных векторов (ok, virus) и на основе аппарата нечеткой логики (в процентах и в присвоении категории программе (безопасная, подозрительная и опасная)). Основными составными частями процесса моделирования являются: аппаратно-программные средства, метод, методология, алгоритмическое и программное обеспечение для мобильной ОС Android. Функциональная модель, отображающая принцип работы системы обнаружения вредоносных программ, представлена на рисунке 1.

В блоке А1 “извлечь признаки” выполняется анализ поступающих на вход компонентов программы (файлов androidmanifest.xml, classes.dex и системных вызовов) и формируется выходной вектор данной программы. Данный вектор подается на блоки А3 “машина опорных векторов” и А4 “нечеткая логика”, в которых выполняется его классификация, а также в условиях сильной размытости кластеров осуществляется уточнение – дополнительная классификация с учетом

условий и результата работы машины опорных векторов. Функциональная модель IDEF0 позволила отобразить структуру и функции системы обнаружения вредоносных программ, а также объекты и потоки информации, связывающие эти функции. Чтобы визуально оценить и рассмотреть процесс работы системы обнаружения вредоносных программ, необходимо построить информационную модель, отображающую структуру и содержание информационных потоков, необходимых для реализации основных функций данной системы.

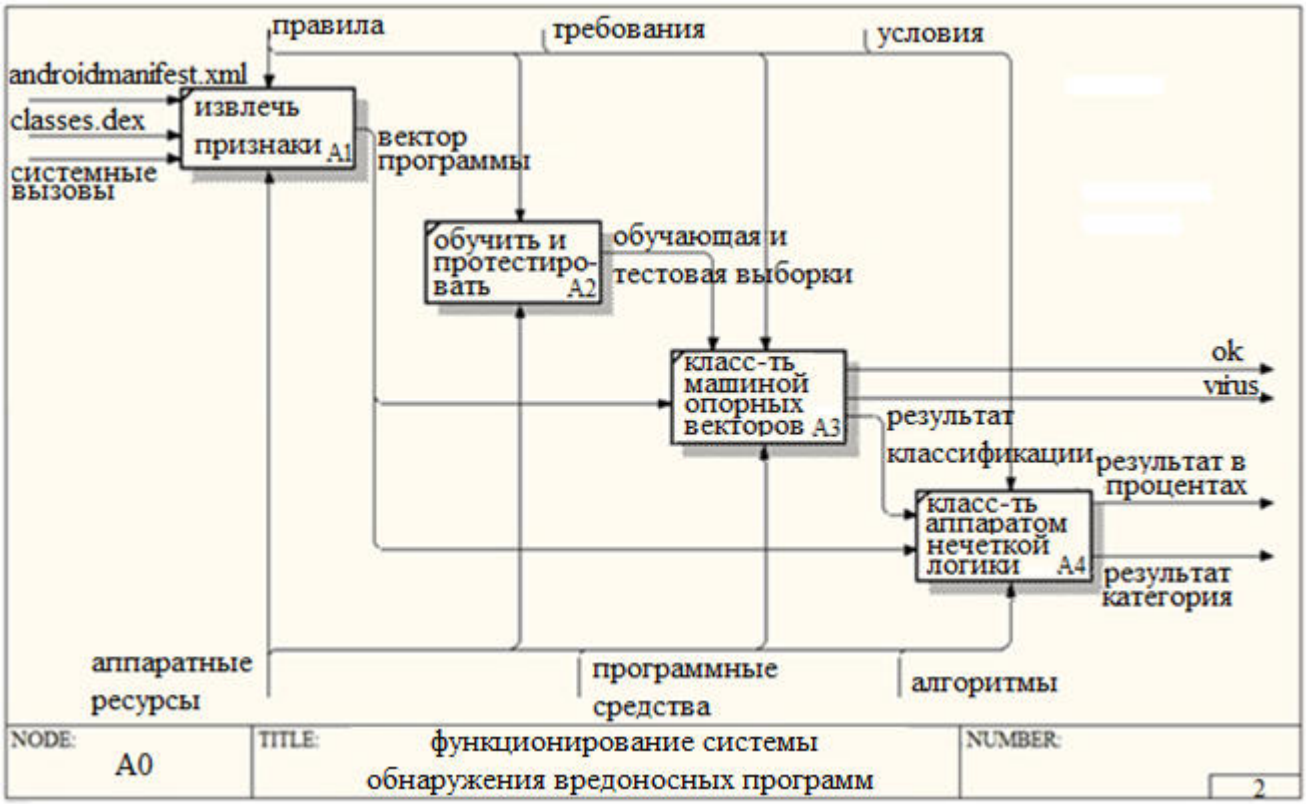

Рисунок 1 – Функциональная модель A1 работы системы обнаружения вредоносных программ в ОС Android

Для этого построим информационную модель IDEF1X, раскрывающую информационное содержание работы системы обнаружения вредоносных программ. Она представляет собой логическую структуру, содержащую информацию об объекте исследования, которая позволяет оценить процесс работы системы обнаружения вредоносных программ в мобильной ОС. На рисунке 2 представлена построенная информационная модель работы системы обнаружения вредоносных программ в мобильной ОС Android. Блок “androidmanifes.xml” содержит в себе перечень разрешений, запрашиваемых прикладной программой, блок “classes.dex” содержит исполняемый код прикладной программы, блок “системные вызовы” содержит информацию о статистике времени обращения к системным вызовам. Далее осуществляется комплексный анализ информации из описанных выше блоков для дальнейшего формирования выходного вектора данных. Данное действие выполняется блоком с названием “извлечение признаков”.

Блок “база знаний” включает в себя экспериментальные данные, представленные в виде выборки, на которой осуществляется процесс обучения и тестирования машины опорных векторов. На выход обученной и протестированной машины опорных векторов подаются данные описанных выше параметров и производится классификация. В случае сильной размытости кластеров используется аппарат нечеткой логики, позволяющий уточнить конечный результат. Нечеткая логика показывает результат в процентах и категории опасности рассматриваемой программы.

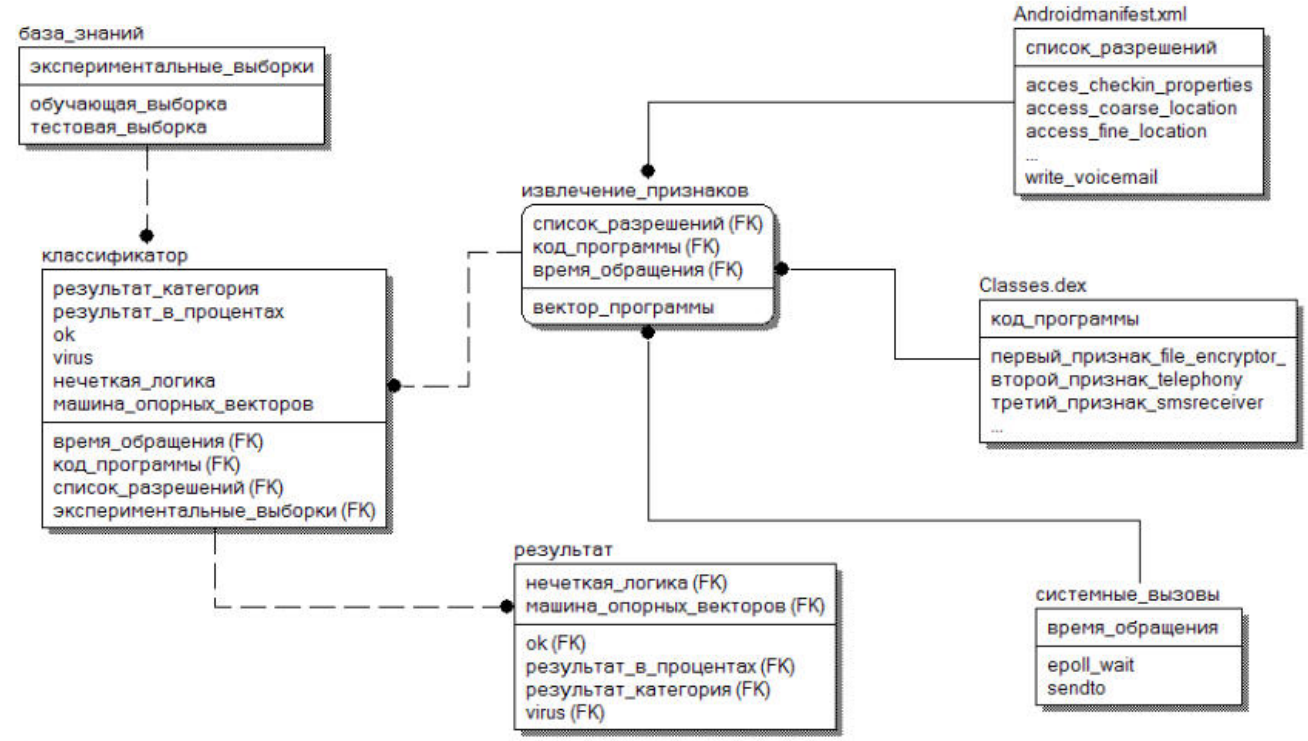

Рисунок 2 – Информационная модель процесса функционирования системы обнаружения вредоносных программ в ОС для мобильных устройств

В процессе функционирования мобильное устройство может взаимодействовать с облачными сервисами (облачные вычисления), средой виртуализации и прочими современными сервисами, так как в нем реализован широкий спектр услуг и множество современных функциональных возможностей. Облачный сервис выступает в качестве источника распространения прикладных программ, хранения информации и т.д.

В рамках данной работы разрабатываемая методика нацелена на обнаружение вредоносных программ в ОС Android. Она может быть реализована как на других уровнях относительно виртуальной машины и облачного сервиса, так и в других альтернативных ОС для мобильных устройств.

Таким образом, построенная функциональная модель IDEF0 позволила отобразить основную структуру и функции разрабатываемой системы обнаружения вредоносных программ, а также объекты и потоки информации, связывающие эти функции. Информационная модель IDEF1X позволила

отобразить структуру и содержание информационных потоков, необходимых для поддержки функций системы обнаружения вредоносных программ.

Выполнен анализ исходного кода множества вредоносных программ, перечня разрешений. В процессе анализа исходного кода были выявлены свойства, присущие поведению вредоносных программ, что позволило выделить экспериментальные данные, а также рассмотреть список существующих разрешений. В приложении А описаны все 152 разрешения. Путем анализа критичности и значимости каждому разрешению присвоено значение 0 – приемлемо (безопасно), 1 – опасно, то есть, при предоставлении доступа на данное разрешение, могут быть задействованы программно-аппаратные ресурсы, которые вследствие управления ими вредоносной программой могут нанести ущерб.

Разработана скрытая марковская модель процесса работы мобильной ОС с учетом воздействия вредоносных программ. Она позволила перейти от качественных оценок динамики процесса работы мобильной ОС к количественным характеристикам. По результатам работы модели сделан вывод, что система находится в рабочем состоянии 68,7 % времени, остальное время – под воздействием вредоносных программ. Следовательно, достаточно большой процент времени система находится в нерабочем состоянии, что свидетельствует о том, что 31,3 % ресурсов мобильных ОС расходует неэффективно.

Анализ результатов первой главы, а также системного моделирования во второй главе позволил разработать нечеткую когнитивную карту. Риск влияния угроз на её основе (с учетом общего риска) с делением на целевые факторы до внедрения системы обнаружения вредоносных программ составил 93 %, что свидетельствует о необходимости применения дополнительных средств защиты. Риск влияния угроз (с учетом анализа общего риска) после внедрения системы обнаружения вредоносных программ составил 47 %. Эффективность внедрения дополнительных средств защиты 49 %, что позволило значительно повысить уровень защищенности ОС для мобильных устройств в целом.

В третьей главе с целью выявления основных признаков, необходимых для описания поведения вредоносных программ, выполнен анализ перечня разрешений системы разрешений, системных вызовов множества вредоносных программ ОС для мобильных устройств. На основе полученных данных разработана экспериментальная выборка, описывающая поведение как вредоносных, так и безопасных программ. В таблице 1 представлен фрагмент экспериментальной выборки, заданной в бинарной форме (0 – отсутствует, 1 – присутствует) и включающей в себя 100 векторов поведения программ. Столбцы с 1 по 10 содержат в себе информацию о выявленных в процессе формализации признаках программ, с 10 по 162 – список всех разрешений. Столбцы 163 и 164 содержат в себе значения, полученные путем анализа используемых системных процессов в ОС для мобильных устройств.

Задача обнаружения сводится к задаче классификации предложенной экспериментальной выборки, описывающей признаки того или иного типа программ (вредоносная или безопасная). Для выбора наиболее подходящего

метода классификации были проведены эксперименты с применением следующих методов классификации:

1. Классических:

– иерархической кластеризации, метод к-средних.

2. Нейросетевых:

– радиально-базисная функция (РБФ), линейная нейронная сеть, персептрон, сеть Ворда, модульная нейронная сеть, прямого распространения, прямого распространения с временным окном равным 12 шагам, сеть Элмана, рекуррентная нейронная сеть, сеть Кохонена.

3. Машина опорных векторов на основе нейронной сети с радиально-базисной функцией активных нейронов скрытого слоя, которая показала лучший результат при решении задачи классификации.

Таблица 1 – Фрагмент экспериментальных данных

	1	2	3	4	5	6	7	8	9	10	…	162	163	164	165
1	0	0	0	0	0	0	0	0	0	0	…	1	40	70	ok
2	0	0	0	0	0	0	0	0	0	0	…	0	0	11	ok
3	1	0	0	0	0	0	0	0	0	0	…	0	95	10	virus
4	1	1	1	1	1	1	1	1	1	1	…	1	41	0	virus
5	1	1	0	0	0	0	0	0	1	1	…	1	94	9	virus
6	1	1	0	1	0	0	0	0	1	0	…	0	93	8	virus
7	1	1	1	1	1	1	1	1	1	0	…	0	92	7	virus
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
95	1	1	0	0	0	0	1	1	0	0	…	0	42	12	virus
96	1	0	0	0	0	1	1	0	1	0	…	0	41	11	virus
97	0	0	0	0	0	1	1	1	0	0	…	0	41	10	ok
98	0	0	1	0	0	0	0	0	0	0	…	1	39	9	ok
99	0	0	0	0	0	0	0	0	0	0	…	1	38	8	ok
100	0	0	1	0	0	0	0	0	0	0	…	0	36	7	ok

Пусть имеется исходное множество программ X, о котором мы знаем которые их них относятся к вредоносным, какие к безопасным, то есть множество разбито на два класса: $X = X_U \cup X_O$. Задана экспериментальная выборка X^M.

Требуется разбить выборку на непересекающиеся подмножества. Для этого для каждой программы $X_i^M \in X^M$ вычисляются две нормы:

$$P_U(X_i^M; X_U) = \sqrt{\sum_{l \in X_U} \sum_{j=1}^{n} (P_j(X_i^M) - P_j(X_l^U))^2} .$$

$$P_O(X_i^M; X_O) = \sqrt{\sum_{k \in X_O} \sum_{j=1}^{n} (P_j(X_i^M) - P_j(X_k^O))^2} .$$

Если $P_U(X_i^M; X_U) \prec \Delta_U$ тогда $X_i^M \in X_U$.

Если $P_O(X_i^M; X_O) \prec \Delta_O$ тогда $X_i^M \in X_O$.

Архитектура нейронной сети выбирается в соответствии с типом решаемой задачи. Для классификации экспериментальной выборки были выбраны типы нейронных сетей, перечисленные выше.

Для обучения выбранных нейронных сетей применялся метод обратного распространения ошибки. Данный алгоритм обеспечивает настройки весов с учетом многослойной структуры сети. Ошибка обучения на выходе нейронной сети распространяется в обратном направлении к скрытым слоям. Для нейронов выходного слоя величина ошибки вычисляется просто как разность между ожидаемым и реальным выходным значением.

Выбор функции активации осуществляется в зависимости от задачи, удобства программно-аппаратной реализации, а также алгоритма обучения. Аргумент функции активации каждого скрытого узла сети радиальной базисной функции представляет собой евклидову норму между входным вектором и центром радиальной функции. Аргумент функции активации каждого скрытого узла сети многослойного персептрона является скалярным произведением входного вектора и вектора синаптических весов данного нейрона. Аргумент функции активации для линейной нейронной сети представляет собой линейную дискриминантную функцию.

Для машины опорных векторов выбрана в качестве ядра радиально-базисная функция:

$$\exp(-\gamma \mid x_i - x_j \mid^2).$$

Нейронные сети и машина опорных векторов обучены на первых 68 и тестировались на остальных 32 значениях векторов экспериментальной выборки, приведенной в таблице 1.

Требуется определить функцию $a: X \to Y$, которая любому объекту $x \in X$ ставит в соответствие номер кластера $y \in Y$. Множество Y в некоторых случаях известно заранее, однако чаще ставится задача определить оптимальное число кластеров с точки зрения того или иного критерия качества кластеризации.

В качестве критерия точности и качества работы классификаторов будем использовать следующую формулу:

$$\text{ОК} = \frac{\text{ЧО} \times 100}{\text{ЧН}},$$

где ОК – общий процент как вредоносных, так и безопасных программ ошибки классификации;
ЧО – число ошибок классификации;
ЧН – суммарное число наблюдений.

С целью выбора лучшего метода классификации для решения поставленной задачи был проведен эксперимент, в результате которого выбран наиболее оптимальный метод, который реализует поставленную задачу с меньшим количеством ошибок первого и второго рода. По результатам проделанного эксперимента установлено, что точность работы классических методов

классификации (иерархической кластеризации и к-средних) невысокая: ошибки I рода – 26,08 %, ошибки II рода – 22,7 %. Классификация с применением нейросетевых методов показала лучшие результаты по сравнению с классическими методами в условиях с присутствием аддитивной помехи (M = 0, σ = ± 0,01 %) преимущественно к большому уровню помех. Результаты эксперимента представлены в таблице 2.

Таблица 2 – Общая таблица результатов классификации

% ошибочных	% правильных	Ошибочно	Правильно	Всего	Классы	Тип нейронной сети
11	88.9	2	16	18	ok	РБФ
20	80	3	12	15	virus	
11.1	88.8	2	16	18	ok	Линейная
26.6	73.3	4	11	15	virus	
66.6	33.3	12	6	18	ok	Персептрон
0	100	0	15	15	virus	
100	0	48	0	48	ok	Ворда
0	100	0	52	52	virus	
61.1	38.9	11	7	18	ok	Модульная
0	100	0	15	15	virus	
33.3	66.7	6	12	18	ok	Прямого распространения
7.1	92.9	1	14	15	virus	
66.6	33.4	12	6	18	ok	Прямого распространения (12)
20	80	3	12	15	virus	
50	50	9	9	18	ok	Элмана
7.1	92.9	1	14	15	virus	
22.2	77.8	4	14	18	ok	Рекуррентная
66.6	33.4	10	5	15	virus	
26.6	73.3	4	11	15	ok	Кохонена
61.1	38.8	11	7	18	virus	
27.7	72.3	5	13	18	ok	Машина опорных векторов
0	100	0	15	15	virus	

Анализ таблицы показывает, что машина опорных векторов лучше осуществляет классификацию по сравнению с нейронными сетями и классическими методами. Ошибочно выполнена классификация 5 типов программ, процент правильно классифицированных составил 72,3 %. Количество ошибок I рода – 0 %, ошибок II рода – 27,7 %. В классе virus классификация была выполнена безошибочно.

Следовательно, машина опорных векторов выполняет классификацию предложенной выборки с меньшим количеством ошибок. Данный метод был взят за основу для разработки общей методики.

В основе данного метода реализована машина опорных векторов. В качестве корректирующего инструмента процесса классификации используется

нечеткая логика, позволяющая выполнять правильную классификацию, а также дополнять результат работы машины опорных векторов с учетом сильных помех. Метод обнаружения вредоносных программ, представленный на рисунке 3, объединяет в себе два метода.

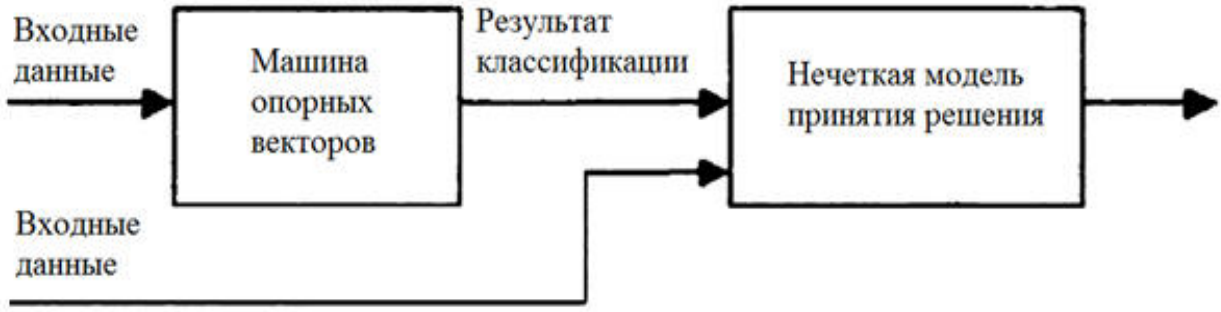

Рисунок 3 – Схема метода обнаружения вредоносных программ

В качестве входных данных используется разработанная и описанная выше экспериментальная выборка. Далее формализуются дополнительные признаки и задаются в виде функций принадлежности для базы нечетких правил. В состав функций принадлежности также входит результат предсказания машиной опорных векторов. На основании всех признаков и работы машины опорных векторов получаем результат, выраженный в процентах.

Алгоритм функционирования метода обнаружения вредоносных программ включает в себя следующие шаги:

1. Извлечение и формирование вектора признаков из программы путем анализа файлов androidmanifest.xml, clases.dex и системных вызовов.

2. Подается вектор признаков программы (вредоносная или безопасная программа) на вход SVM-классификатора.

В результате обучения машины опорных векторов определяются опорные вектора, являющиеся векторами характеристик тех объектов x_i из экспериментальной выборки, для которых значения соответствующих им двойственных переменных λ_i отличны от нуля ($\lambda_i \neq 0$). Опорные вектора находятся ближе всего к гиперплоскости, разделяющей классы, и несут всю информацию о разделении классов. Так как задача квадратичного программирования решена, то классификация произвольного объекта λ будет выполнена по следующему правилу:

$$\alpha(z) = sign\sum_{i=1}^{m} \lambda_i \cdot y_i \cdot \exp(- < x_i - x_j, x_i - x_j > /(2 \cdot \sigma^2) - b),$$

где $b = < w, x_i > - y_i$;

λ_i – двойственная переменная;

x_i – объект из экспериментальной выборки;

y_i – число (-1 или 1), характеризующее классовую принадлежность объекта x_i из экспериментальной выборки;

$\exp\left(-\prec x_i - x_j, x_i - x_j \succ /(2\cdot\sigma^2)\right) = k(x_i, x_j)$ – радиальная базисная функция ядра;

C – параметр регуляризации (C>0);

m – количество объектов в экспериментальной выборке; $i = \overline{1,m}$.

$$w = \sum_{i=1}^{m} \lambda_i \cdot y_i \cdot x_i .$$

При этом суммирование в правиле выполняется только по опорным векторам.

3. Результат классификации и дополнительные признаки, заданные в виде функций, подаются блоку “система поддержки принятия решений”.

Система поддержки принятия решений основана на аппарате нечеткой логики, работающей по алгоритму Мамдани. Алгоритм работы данного компонента системы обнаружения вредоносных программ представлен на рисунке 6.

4. На основании правил производится анализ результатов.

5. Выводится результат в процентах.

В четвертой главе на основе проделанного в третьей главе эксперимента был разработан исследовательский прототип программы, реализующий предложенные методы: машину опорных векторов и систему поддержки принятия решений на основе нечеткой логики по алгоритму Мамдани. Исследовательский прототип позволяет выполнить анализ, а также оценку эффективности работы методов при различном уровне помех. Он реализован в виде программы на языке C++.

Программа представляет собой окно с тремя вкладками (рисунок 4): главная, статистика и журнал.

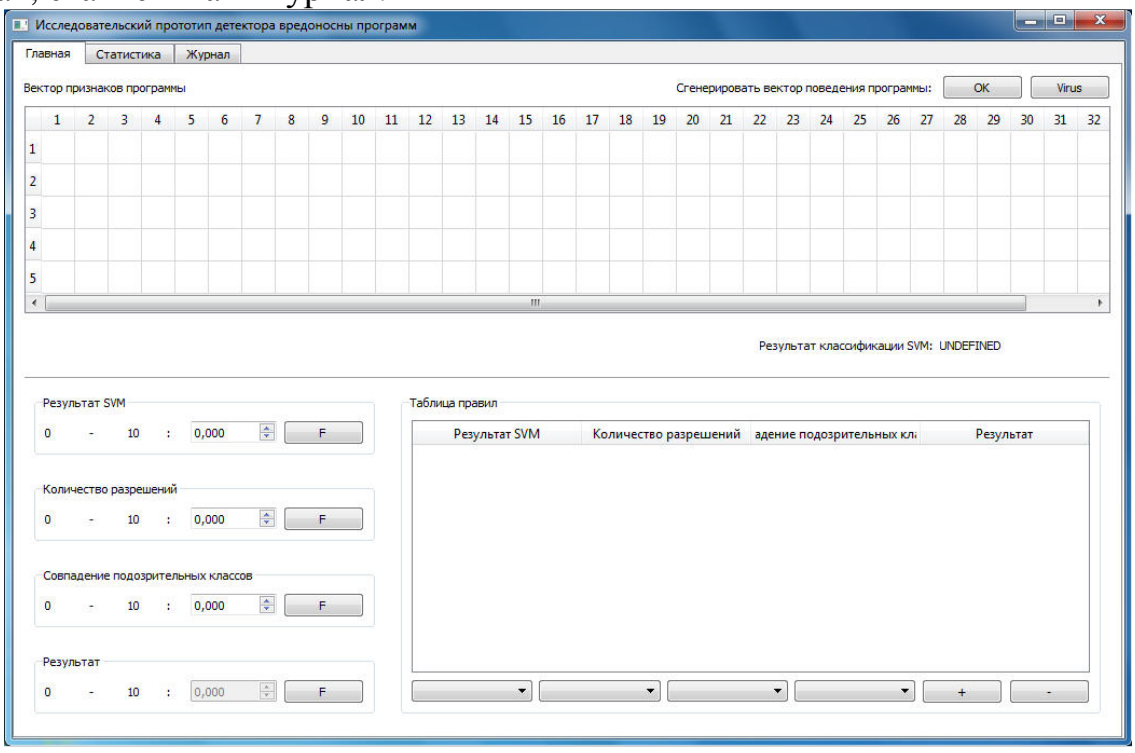

Рисунок 4 – Главное окно программы

Во вкладке “Главная” содержится таблица для ввода значений векторов, характеризующих поведение той или иной программы, что позволяет внести необходимое количество помех в рассматриваемый вектор. Также можно автоматически сгенерировать необходимый вектор программы: ok или virus, при этом машина опорных векторов не будет иметь представления о том, какой тип программы сгенерирован. Она будет выполнять классификацию и выводить результат.

Во вкладке “Статистика” содержится вся информации о проделанных экспериментах в виде таблицы с двумя столбцами: эталонная модель и результат работы машины опорных векторов.

Во вкладке “Журнал” ведется статистика всех действий, связанных с работой программы.

С целью оценки эффективности работы предлагаемого исследовательского прототипа модели системы обнаружения вредоносных программ с точки зрения практического применения выполнено сравнение результатов работы разработанной программы с результатами существующих антивирусных программ. Измерялся процент распознавания новых вредоносных программ на 35 образцах антивирусных программ, результаты приведены в таблице 3.

Таблица 3 – Результаты обнаружения вредоносных программ

Название	Процент обнаружения
Система обнаружения вредоносных программ	80
GData	60
Ikarus	55
Emsisoft	54
McAfee-GW-Edition	50
Panda, BitDefender	49
AntiVir (AVIRA GmbH), McAfee	47
F-Secure	46
NOD32 (ESET), VIPRE (GFI Software)	43
DrWeb	36
Norman	32
Sophos, Kaspersky	22
AVG	21
Avast	20
AhnLab-V3, Symantec	28
Microsoft	21
Comodo, TrendMicro, PCTools	20
nProtect	19
K7AntiVirus, TrendMicro-HouseCall	17
Prevx	16

Окончание таблицы 3

VBA32	11
F-Prot, Rising	8
Fortinet, VirusBuster, Commtouch, eSafe	6
SUPERAntiSpyware	4
CAT-QuickHeal, ClamAV	2
eTrust-Vet, Antiy-AVL, Jiangmin, TheHacker, ViRobot	1

Антивирусные программы обладают низкой эффективностью обнаружения новых вредоносных программ, результат лучшей составил 60 %, средний результат по всем рассматриваемым образцам составил 23,4 %. Разработанный исследовательский прототип модели системы обнаружения вредоносных программ показал лучший результат на уровне 80 % и сравнительно небольшое количество ложных срабатываний, равное 8 программам из 100 рассматриваемых образцов, что составило 19,35 %, но при этом не учитывался результат работы аппарата нечеткой логики, который выполнил коррекцию результата путем дополнительной классификации. Машина опорных векторов ошиблась, выдав вредоносную программу за безопасную, но нечеткая логика показала результат: подозрительная 45 % и опасная 65 %, таким образом дополнив работу машины опорных векторов и улучшив показатели эффективности обнаружения. Следовательно, разработанный метод обнаружения вредоносных программ позволяет выполнять их идентификацию путем анализа поведенческого характера программ, а также увеличить эффективность обнаружения, что положительно влияет на защиту информации в целом. Данный метод позволяет своевременно обнаружить и нейтрализовать угрозу со стороны как новых, так и уже имеющихся типов вредоносных программ, а также может использоваться в комплексе с классическими методами.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

1. Показано, что проблема защищенности ОС для мобильных устройств является актуальной темой на основе анализа сложившейся ситуации, согласно которому число вредоносных программ выросло более чем в 10 раз за период с 2012 по 2014 гг. и составило 12 миллионов в 2014 году. Анализ архитектуры мобильной ОС, структуры прикладных программ и уязвимостей (вредоносные программы, система разрешений, модифицированная мобильная ОС и т.д.) позволил сделать вывод, что существует необходимость исследований в данном направлении.

2. Предложена и обоснована экспериментальная выборка, которая позволяет описать поведенческий характер потенциальной вредоносной программы и на основе исследования множества прикладных программ классифицировать их.

3. Разработаны системные модели процесса функционирования системы обнаружения вредоносных программ для ОС Android с учетом различных факторов на основе SADT-методологии и IDFE-технологий, позволяющие

интегрировать систему обнаружения вредоносных программ с компонентами внутренних и внешних механизмов защиты ОС Android.

4. Показано, что машина опорных векторов выполняет классификацию разработанной экспериментальной выборки с большей точностью (процент правильно классифицированных составил 72,3 %, количество ошибок I рода – 0 %, ошибок II рода – 27,7 %), чем классические и нейросетевые методы, на основе проделанных экспериментов, позволяющих осуществить выбор наилучшего метода классификации. Установлено, что классические методы обладают невысокой точностью: ошибки I рода – 26,08 %, ошибки II рода – 22,7 %, а также, что ошибки классификации нейронными сетями присутствуют по всей выборке, но увеличиваются с ростом количества помех в экспериментальной выборке.

5. Разработан и реализован в виде программы комплекс алгоритмов обнаружения вредоносных программ в ОС Android на основе интеллектуальных технологий, позволяющий с высокой точностью выполнять обнаружение вредоносных программ, отличающийся динамическим анализом поведения вредоносной программы от классического сигнатурного метода.

6. Установлено, что исследовательский прототип модели системы обнаружения вредоносных программ с эффективностью 80 % выполняет обнаружение вредоносных программ на основе сравнительного анализа результатов обнаружения вредоносных программ: лучший результат антивирусными программами составил 60 %, что говорит о том, что эффективность обнаружения прототипа выше на 20 %.

Перспективы дальнейшей разработки темы. Разработанная методика обнаружения вредоносных программ в ОС для мобильных устройств (на примере Android) позволит увеличить эффективность обнаружения новых вредоносных программ. Данный алгоритм в связке с классическими методами анализа (сигнатурный метод) позволит значительно увеличить процент обнаружения как новых, так и уже известных вредоносных программ. Следовательно, это позволит укрепить существующие механизмы защиты информации в ОС для мобильных устройств, а также защитить личную информацию, которая располагается в мобильном устройстве.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

В рецензируемых журналах из списка ВАК

1. Обзор современного состояния защиты информации в мобильных системах / Жернаков С. В., Гаврилов Г. Н. // Научно-теоретический журнал «Вестник БГТУ им. В. Г. Шухова», №2, 2016 г. С. 171-176.

2. Детектирование вредоносного программного обеспечения с применением классических и нейросетевых методов классификации / Жернаков С. В., Гаврилов Г. Н. // Научно-теоретический журнал «Вестник воронежского государственного университета инженерных технологий», №4 (66), 2015 г. С. 85-93.

3. Malicious software detection in operating system (OS) for mobile devices (the case of Android OS) / Жернаков С. В., Гаврилов Г. Н. // XII международная

научная-техническая конференция «Актуальные проблемы электронного приборостроения» (Scopus), Том 1, Часть 2, 2016 г. С. 163-165.

4. Система обнаружения вредоносных программ в операционной системе (ОС) Android / Жернаков С. В., Гаврилов Г. Н. // Научный журнал «Вестник УГАТУ», Том 21, №2 (71), 2016 г. С. 117-122.

5. Применение интеллектуальных технологий для обнаружения вредоносных программ в операционной системе (ОС) Android / Жернаков С. В., Гаврилов Г. Н. // Международный научно-исследовательский журнал «International research journal», №5(47), 2016 г. С. 94-98.

6. Методика обнаружения вредоносных программ в операционной системе (ОС) для мобильных устройств (на примере ОС Android) / Жернаков С. В., Гаврилов Г. Н. // Теоретический и прикладной научно-технический журнал «Программная инженерия», №10(7), 2016 г. С. 455 – 463.

Объекты интеллектуальной собственности

7. Свидетельство о государственной регистрации программы для ЭВМ № 2016613207. Модель исследовательского прототипа детектора вредоносных программ / С. В. Жернаков, Г. Н. Гаврилов. Зарегистрирована 21.03.2016 г. – М.: Роспатент, 2016.

В трудах международных и всероссийских конференций

8. Анализ угроз информационной безопасности современных мобильных систем / Жернаков С. В., Гаврилов Г. Н. // Интеграционные процессы науки XXI века: сборник статей Международной научно-практической конференции. Стерлитамак: РИЦ АМИ, 2015. С. 54 – 60.

9. Взгляд на защиту информации в мобильной системе / Жернаков С.В., Гаврилов Г.Н. // Научные перспективы XXI века: материалы XIII Международной (заочной) научно-практической. Нефтекамск: РИО ООО «Наука и образование», 2015. С. 46-52.

10. Выявление вредоносных программ с использованием современного интеллектуального метода на этапе установки / Гаврилов Г. Н., Жернаков С. В. // XIII Международная научно-практическая конференция «Научные перспективы XXI века. Достижения и перспективы нового столетия», г. Новосибирск, 2015 г.;

11. Об одном подходе защиты информации в средствах мобильной связи / Жернаков С. В., Гаврилов Г. Н. // XV Международная научно-практическая конференция: Научное обозрение физико-математических и технических наук в XXI веке. Москва: Ежемес. науч. журнал «Prospero», 2015.С. 9-13.

12. Реализация метода опорных векторов в системе Android для классификации и обнаружения вредоносных программ / Жернаков С. В., Гаврилов Г. Н. // Актуальные вопросы развития инновационной деятельности в новом тысячелетии: XVII Международная научно-практическая конференция. Новосибирск: Ежемес. науч. журнал «МИС» №6 (17), 2015. С. 10-14.

13. Применение искусственной иммунной системы для обнаружения вредоносных программ в мобильных устройствах Android / Жернаков С. В., Гаврилов Г. Н. // XXII Туполевские чтения (школа молодых ученных):

Международная молодежная научная конференция. Том IV. Казань: Изд-во «Фолиант», 2015. С. 70-78.

14. Реализация методов анализа данных для классификации вредоносного программного обеспечения / Жернаков С. В., Гаврилов Г. Н. // Проблемы техники и технологии телекоммуникаций: XVI Международная научно-техническая конференция. Том 3. Уфа: УГАТУ, 2015. С. 130-134.

15. Детектирование вредоносного программного обеспечения в мобильной операционной системе на базе Android на основе разрешений с применением метода опорных векторов / Жернаков С. В., Гаврилов Г. Н. // Научно периодическое издание Ceteris Paribus: Европейский фонд инновационного развития. Ежемес. науч. журнал «Ceteris Paribus», 2015. С. 10-14.

16. Применение классических методов кластеризации для классификации программ по их свойствам / Жернаков С. В., Гаврилов Г. Н. // Мавлютовские чтения: IX Всероссийская молодежная научная конференция, 2015. С. 36-43.

17. Применение Марковской цепи для моделирования функционирования мобильной операционной системы (ОС), типа Android с учетом воздействия вредоносных программ / Гаврилов Г. Н., Жернаков С. В. // Перспективные направления развития современной науки: XV Международная научная конференция. г. Москва: Сборник научных работ, 2016. С. 45-48.

18. Оценка риска информационной безопасности, а также эффективности внедрения дополнительной системы обнаружения вредоносных программ в мобильной операционной системе (ОС) типа Android по средствам нечетких когнитивных карт (НКК) / Гаврилов Г. Н., Жернаков С. В. // Актуальные проблемы науки XXI века: VIII Международная научно-практическая конференция. г. Москва: Сборник статей часть 1, 2016. С. 83-90.

Диссертант Г. Н. Гаврилов

ГАВРИЛОВ Григорий Николаевич

РАЗРАБОТКА ДЕТЕКТОРА ВРЕДОНОСНЫХ ПРОГРАММ В МОБИЛЬНОЙ ОПЕРАЦИОННОЙ СИСТЕМЕ (ОС) ТИПА ANDROID С ПРИМЕНЕНИЕМ СОВРЕМЕННЫХ ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ

Специальность 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 00.00.2017. Формат 60x84 1/16.
Бумага офсетная. Печать плоская. Гарнитура Times New Roman.
Усл. печ. л. 1,2. Уч.-изд. л. 1,1.
Тираж 100 экз. Заказ № .
ФГБОУ ВО «Уфимский государственный авиационный технический университет»
Редакционно-издательский комплекс УГАТУ
450008, г. Уфа, ул. К. Маркса, 12.