ISSN 2305-994X

МАТЕРИАЛЫ
31-Й НАУЧНО-ТЕХНИЧЕСКОЙ КОНФЕРЕНЦИИ

МЕТОДЫ И ТЕХНИЧЕСКИЕ СРЕДСТВА ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ИНФОРМАЦИИ

27 – 30 ИЮНЯ 2022 ГОДА

Санкт-Петербург 2022

Методы и технические средства обеспечения безопасности информации:
Материалы 31-й научно-технической конференции **27 – 30 июня 2022** года. СПб: Изд-во Политехнического университета, 2022.

Приводятся тезисы докладов, отражающие теоретические и практические проблемы обеспечения безопасности информационных технологий, а также вопросы подготовки и переподготовки специалистов в этом направлении.

Предназначаются для преподавателей, научных сотрудников и инженерно-технических работников, занимающихся проблемами информационной безопасности. Ответственный за выпуск – доктор технических наук, профессор, член-корр. РАН Д.П. Зегжда

Сборник печатается без редакторских правок.

Ответственность за содержание тезисов возлагается на авторов.

ISSN 2305-994X

Введение:

31-й выпуск сборника «МАТЕРИАЛЫ НАУЧНО-ТЕХНИЧЕСКОЙ КОНФЕРЕНЦИИ «МЕТОДЫ И ТЕХНИЧЕСКИЕ СРЕДСТВА ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ИНФОРМАЦИИ» посвящается основателю и председателю Организационного комитета конференции, доктору технических наук, профессору, Заслуженному деятелю науки России, основателю и научному руководителю Института кибербезопасности и защиты информации Санкт-Петербургского политехнического университета Петра Великого Петру Дмитриевичу Зегжде.

Петр Дмитриевич посвятил свою жизнь развитию отечественной науки, занимаясь важнейшими фундаментальными и прикладными исследованиями в области информационной безопасности. Он был одним из первых отечественных специалистов в этой сфере. Труды и идеи Петра Дмитриевича стали основой для десятков уникальных разработок, в числе которых средства анализа управления безопасностью, криптографические системы защиты информации, антивирусные системы, защищенная операционная система и многие другие.

Зегжда П.Д. был не только выдающимся учёным, но и талантливым педагогом и организатором, под его началом была создана одна из первых в России кафедр в сфере информационной безопасности – кафедра «Информационная безопасность компьютерных систем» Санкт-Петербургского политехнического университета, которая стала основой для первого в России Института кибербезопасности и защиты информации.

Петр Дмитриевич активно выступал за популяризацию науки, по его инициативе в 1991 году впервые была проведена Всероссийская научно-техническая конференция «Методы и технические средства обеспечения безопасности информации», ставшая в дальнейшем одним из самых авторитетных мероприятий в сфере информационной безопасности России. Более 30 лет Пётр Дмитриевич руководил организацией конференции, лично занимался отбором докладов, созданием секций и непосредственным общением с делегатами, учёными и коллегами. Как председатель организационного комитета, он всегда стремился к разнообразию тем и научных идей, привлекая к участию самых интересных докладчиков, от его внимания не ускользали новые тенденции как теории, так и практических аспектов информационной безопасности. 30 выпусков сборника материалов конференции являются, по-сути, уникальной ретроспективой, справочником развития информационной безопасности в России.

Основные идеи организации конференции – диалог на пересечении теории и практики, науки и бизнеса, а также приоритет живого общения, открытой дискуссии были заложены Петром Дмитриевичем, и будут продолжаться и развиваться на всех последующих конференциях.

Организационный комитет
31-й научно-технической конференции
«Методы и технические средства обеспечения безопасности информации»
им. Петра Дмитриевича Зегжды.

1. Задачи информационной безопасности в эпоху цифровой трансформации

Полтавцева М.А., Зегжда Д.П.
Санкт-Петербургский политехнический университет Петра Великого

АНАЛИЗ ГЕТЕРОГЕННЫХ ПРЕЦЕДЕНТОВ В ЗАДАЧАХ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Развитие информационных технологий, цифровизация различных отраслей деятельности и применение цифровых решений в самых разных сферах от производства до государственных услуг приводит, с одной стороны, к увеличению числа и разнообразия атак на компьютерные системы, а с другой – к совершенствованию методов защиты от них.

Использование злоумышленниками социальной инженерии, информации из открытых источников, результатов пассивного и активного сбора данных об объекте и логического вывода (inference) при планировании и реализации APT-атак вызывает необходимость учета этих факторов при реализации мер и проектировании средств защиты, что нашло отражение, в частности в государственных стандартах [1-2]. Тестирование на проникновение (или тестирование проникновения) становится общепринятой практикой оценки защищенности [1] а оценка потока данных об угрозах в контексте постоянного анализа событий безопасности – обязательной практикой мониторинга информационной безопасности объекта защиты [2]. Это только часть подобного рода задач, так как оценка разнородных данных из гетерогенных источников, в том числе – открытых, возникает также при моделировании угроз (оценке информированности злоумышленника) и в некоторых других моментах. В области прогностической защиты от кибератак такой задачей является, например, построение ассоциативной памяти для поиска подобных прецедентов в «прошло» системы и выработки опережающей реакции с учетом возможного развития атаки. Сегодня ключевой аспект этих техник, анализ прецедентов, во многом реализуется вручную.

Существует множество инструментов, посвященных сбору и подготовке данных для решения указанных задач. Это, например, разнообразные средства работы с открытыми источниками, инструменты тестирования на проникновение, модули работы с задами данных угроз. Однако каждый такой инструмент предоставляет только набор слабо структурированных данных полученных из некоторого источника, который должен быть интегрирован с другими наборами, при необходимости – верифицирован, оценен и проанализирован специалистом. Отдельными проблемами являются:

— Гетерогенность источников данных, и в силу этого, синтаксическая и семантическая гетерогенность самих данных;

— Зависимость характеристик, значений и полноты данных от времени сбора;

— Зависимость полноты данных от настроек сканируемого объекта в целом, включая другие компоненты, а не только объекта характеризующегося конкретными данными;

— Невозможность подготовки наборов данных для применения популярных прецедентных техник машинного обучения.

Последнее является важным фактором, не позволяющим применить современные подходы для работы с разнородными данными обладающими неявными связями и зависимостями, такие как использование нейронных сетей глубокого обучения и др. Каждый объект анализа в данных задачах обладает высокой степенью уникальности, причем это разнообразие увеличивается и меняется по мере развития цифровых технологий, внедрения отраслевых требований, меняющих ландшафт получаемых при сборе данных и по другим причинам, включая эволюцию во времени уже существующих систем. В отношении оценки угроз, методы и средства, применяемые злоумышленниками, включая техники проведения многоступенчатых APT - атак также обладают высокой степенью изменчивости и эволюции, не позволяя полагаться только на имевшийся ранее опыт. Поэтому составление релевантной обучающей выборки, достаточно полно охватывающей предметную область, которая не устареет к моменту начала ее использования, не представляется сегодня возможным.

В этих условиях необходимо применять другие техники прецедентного анализа, позволяющие решать такие задачи, как:

— поиск подобных объектов и систем в базе знаний аналитической системы;

— поиск подобных сценариев в базе знаний;

а также обеспечить актуализацию базы знаний по мере эволюционных изменений как объекта защиты, так и внешней среды – источника угроз.

В докладе представлен базовый метод формализации гетерогенных данных согласно подходу “bag of objects” и представления прецедентов (на примере компьютерных атак и/или их имитации при тестировании проникновения) как графов объектов, показывающие хорошие результаты при решении задач поиска подобия. Дополнительные метрики, характеризующие свойства объектов, позволяют эффективно поддерживать релевантность базы знаний в динамических условиях.

Безусловно, построение автоматической системы анализе гетерогенных данных при решении рассматриваемых задач требует дополнительного расширения инструментария и специализации на конкретной задаче и объекте защиты для повышения точности анализа. Однако полученные результаты достаточны для построения эффективных автоматизированных систем, систем поддержки принятия решений в задачах информационной безопасности при анализе гетерогенных прецедентов для повышения скорости анализа, упрощения работы специалиста по информационной безопасности и тиражирования опыта путем развития и распространения базы знаний в этой области компетенций.

Список литературы:

1. ГОСТ Р 58143-2018 Методы и средства обеспечения безопасности. Детализация анализа уязвимостей программного обеспечения в соответствии с ГОСТ Р ИСО/МЭК 15408 и

ГОСТ Р ИСО/МЭК 18045. Часть 2. Тестирование проникновения. Утв. приказом Федерального агентства по техническому регулированию и метрологии от 24 мая 2018 г. N 274-ст.

2. ГОСТ Р 59547-2021 Защита информации. Мониторинг информационной безопасности. Общие положения. Утв. Приказом Федерального агентства по техническому регулированию и метрологии от 27 июля 2021 г. N 656-ст

3. Poltavtseva M.A., Semyanov P.V., Zaitzeva E.A. Heterogeneous semi-structured data analysis in information security // В сборнике: 2020 International Conference Engineering and Telecommunication, En and T – 2020. – 2020. – 1-5 С. doi: 10.1109/EnT50437.2020.9431309.

И.А.Трифаленков
ООО «НеоБИТ»

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ КАК ПОЛЕ ДЛЯ ИМПОРТОЗАМЕЩЕНИЯ

Принято считать, что с точки зрения импортозамещения информационная безопасность является одной из наиболее успешных областей ИТ, и это во многом справедливо. В области ИБ существует значительная номенклатура продуктов, производители этих продуктов ранее делали небезуспешные попытки позиционироваться на мировом рынке, у большинства из них есть развитая инфраструктура для создания, развития и поддержки линеек средств защиты. Эта позиция отражается в Указе Президента РФ №166 от 30.03.2022 предусматривающего отказ от импортных СЗИ к 2025 году.

Тем не менее представляется целесообразным проанализировать как имеющуюся номенклатуру, так и характеристики отечественных СЗИ и СКЗИ в сравнении с аналогами, имеющимися на мировом рынке.

При анализе средств защиты анализировать необходимо как основные характеристики (сервисы безопасности) так и эксплуатационные свойства продуктов. Эти свойства включают в себя:

— Производительность;

— Отказоустойчивость;

— Удобство эксплуатации;

— Интеграция в неоднородную инфраструктуру защиты;

Анализ номенклатуры средств защиты показывает наличие сегментов, не представленных отечественными разработчиками. Список таких сегментов включает в себя:

— Средства генерации и анализа траффика;

— Средства автоматического моделирования атак;

— Средства автоматизации проведения тестов на проникновение;

— Платформы управления политиками сетевых СЗИ;

— Защищенные платформы управления мобильными устройствами;

— Защищенные платформы контейнерных вычислений;

Разработка таких систем осложняется еще и тем, что стандарты и требования для них также необходимо разрабатывать и утверждать на уровне регуляторов.

Основными проблемами существующих отечественных решений практически в любом классе остаются вопросы удобства эксплуатации и интеграции со средствами других производителей, без которой создание полноценной комплексной системы информационной безопасности невозможно. Для решения этих проблем также необходима разработка соответствующих концептуальных подходов и стандартов.

Без решения указанных проблем попытка отказа от импортных средств защиты чревата рисками перехода на чисто формальную защиту информационных систем, которая оказывается уязвимой для современных видов атак, проводимых квалифицированными нарушителями.

Грушо А.А., Грушо Н.А., Тимонина Е.Е.
Федеральный исследовательский центр «Информатика и управление» РАН, Москва

МЕТАДАННЫЕ ДЛЯ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

Проблемы контроля и обеспечения информационной безопасности (ИБ) в сетецентрической системе являются актуальными. Для исследования этих проблем выбрана сетецентрическая система электронного документооборота (ЭДО) распределенного предприятия. ЭДО представляет собой защищенную распределенную информационную систему (РИС), в которой структура порождается организационной структурой предприятия, существует защищенная связь, обеспечивающая функционал ЭДО. Функционал ЭДО состоит из:

— защищенных функций взаимодействия с внешним окружением, не принадлежащим РИС (получение и отсылка документов);

— базы данных документов (архив документов);

— функций управления документооборотом;

— функций формирования поручений;

— функций доставки поручений исполнителям, контроля исполнения поручений и времени исполнения;

— функций идентификации и аутентификации, функций электронной подписи, удостоверяющего центра, центра генерации и распределения криптографических ключей;

— функций обеспечения целостности и неотказуемости движения документов.

Для реализации ЭДО в РИС необходима дополнительная информация и инструменты защиты ЭДО как распределенной информационной технологии (ИТ). Для обеспечения ИБ распределенных ИТ в работах [1, 2] разработан подход, основанный на метаданных (МД).

Суть подхода состоит в том, что управление соединениями и функции контроля выполнения ИТ реализуются централизованно на основе модели ИТ [3]. Такое управление организовано на знаниях о взаимодействиях субъектов ИТ и их расположении на хостах распределенной системы компьютеров. Эти знания формируются и хранятся в задаче М, которая изолирована от хостов и функций, которые на них выполняются. Такая изоляция является основой безопасности управления выполнением ИТ, использующей даже не безопасную сеть связи. Используя данные задачи М, следующая задача Н управляет соединениями и контролем развития ИТ. Передача информации от задачи М к задаче Н является однонаправленной и это не может нарушить безопасность управления [4].

В рассматриваемом случае отличие модели ИТ состоит в том, что сама ИТ ЭДО представляет собой динамически изменяемую структуру, решающую ограниченный набор задач с постоянно меняющимся набором исполнителей. Поэтому постоянно меняется модель маршрутов выполнения ИТ. Отсюда следует, что выделяется совокупность привилегированных субъектов РИС, которые обладают правом менять модель исполнения ИТ, находящуюся на более высоком уровне защиты. Причем это право должно ограничиваться определенными условиями движения управляющих документов вплоть до создания промежуточных управляющих документов и списков их исполнителей. Такая перенастройка маршрутов должна проводиться в ограниченное время, что требует использования элементов искусственного интеллекта.

Появление элементов автоматического принятия решений требует реализации мер доверия к этим решениям. Меры доверия относятся к функциям ИБ, которые предполагают построение модели угроз и модели нарушителей, которые формируются в специальной политике безопасности. Однако самое сложное в решении проблемы доверия состоит в том, чтобы построить модель обеспечения доверия и снабдить ее механизмами реализации.

Наличие субъектов управления изменениями модели нарушает важный принцип ИБ, запрещающий информационный поток с нижнего уровня защищенности к верхнему уровню. Эта проблема не снимается созданием изолированного безопасного контура, обслуживающего субъектов, участвующих в управлении ЭДО. Это связано с тем, что результаты выполнения поручений должны доставляться субъектам, давших поручения также в форме документов. Полностью изолировать ИТ ЭДО также невозможно, так как она взаимодействует с исполнителями поручений.

Детальный анализ взаимодействий ЭДО с субъектами, дающими поручения, позволяет модернизировать систему метаданных так, чтобы решить указанные проблемы. Основа этих изменений МД состоит в том, что для формирования маршрутов документооборота и контроля за движениями документов достаточно «бедного» языка, который не используется для формулирования самих поручений. «Бедный» язык при контроле синтаксиса может быть создан с таким условием, что не может использоваться для нанесения ущерба системе управления ЭДО.

Список литературы:

1. Grusho A., Grusho N., Zabezhailo M., Zatsarinny A., Timonina E. Information Security of SDN on the Basis of Meta Data // Lecture Notes in Computer Science, 2017. vol. 10446. P. 339-347.

2. Alexander Grusho, Nick Grusho, and Elena Timonina. Information Flow Control on the Basis of Meta // In: Vishnevskiy V., Samouylov K., Kozyrev D. (eds) Distributed Computer and Communication Networks. DCCN 2019. Lecture Notes in Computer Science, 2019. vol. 11965. P. 548-562.

3. Grusho A. A., Timonina E. E., Shorgin S. Y. Modelling for ensuring information security of the distributed information systems // Proceedings of 31th European Conference on Modelling and Simulation, 2017. P. 656-660.

4. Грушо А.А., Применко Э.А., Тимонина Е.Е. Теоретические основы компьютерной безопасности. – М.: Академия, 2009. 272 с.

Даниленко А.Ю., Акимова Г.П.
Федеральный исследовательский центр «Информатика и управление» РАН, Москва

ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ СОВРЕМЕННЫХ СИСТЕМ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

Системы электронного документооборота в настоящее время получают все большее распространение в связи с широким внедрением информационных технологий во все сферы жизни, которое принято называть цифровизацией. Практика использование СЭД в течение многих лет показала, что автоматизированные информационные системы (АИС) этой категории развиваются по нескольким направлениям. В большинстве своем они относятся к АИС в защищенном исполнении (АСЗИ), что связано с обработкой в них конфиденциальной информации, в том числе персональных данных (см. [1, 2]). А это, в свою очередь, накладывает определенные ограничения на операционные системы (ОС), под управлением которых работает АСЗИ. В частности, в последнее время появилось новое требование – использовать при разработке АИС отечественные ОС [3].

Наряду с реализацией функциональных возможностей, учитывающих особенности деловой логики предприятий и их подразделений, развивается и технологическая основа СЭД. В настоящей работе рассмотрены особенности обеспечения информационной безопасности при внедрении ряда технологических решений и совершенствовании функционала СЭД, варианты использования средств ОС для обеспечения безопасности рассмотрены в [4].

Электронный документооборот в настоящее время эффективно применяется как в рамках одного подразделения или предприятия, так и в организациях, имеющих разветвленной структуру, или, при наличии соответствующего соглашения, между различными организациями, что вводит дополнительные требования к обеспечению безопасности межструктурного взаимодействия. В последнем случае сразу встает вопрос подтверждения авторства и неизменности электронных документов и писем, то есть обеспечения целостности данных. Эта задача решается применением усиленной (квалифицированной или неквалифицированной) электронной подписи (ЭП), которая требует использования криптографических пакетов и инфраструктуры открытых и закрытых криптографических ключей. Авторство и неизменность документа могут быть подтверждены путем реализации подсистем контроля целостности, протоколирования и управления доступом в составе СЭД, которые являются компонентами подсистемы защиты информации. Второй вариант может рассматриваться как способ использования простых ЭП.

Передача электронных документов может осуществляться как по каналам связи, так и на внешних носителях. В обоих случаях СЭД должна обеспечивать внесение автоматизированным способом полученной информации в базу данных (БД) с минимальным участием пользователя и максимальной автоматизацией процесса. Отдельной проработки требует вопрос персонификации действий по вводу данных как при передаче на внешних носителях, так и по каналам связи, поскольку это действие должно быть зафиксировано в протоколах работы системы и протоколе безопасности как создание объекта защиты.

Системы электронного документооборота предназначены для работы с электронными документами, однако полностью бумажные документы будут заменены электронными нескоро. В связи с этим в СЭД до сих пор востребован функционал, синхронизирующий работу с электронными и бумажными документами и облегчающий их обработку.

Применение штрихкодирования – одна из технологий, применение которой активно развивается. Основное направление использования двумерных штрихкодов (QR-кодов) в СЭД – кодирование данных документа и нанесение штрихкода на сам документ. Эта информация позволяет выполнить автоматическую регистрацию пересылаемого документа путем заполнения значений реквизитов и уменьшает влияние человеческого фактора на качество ввода данных. Однако, следует учитывать, что оператор не может прочитать информацию непосредственно из QR-кода, поэтому единственный способ обеспечить достоверность вводимой информации – это контроль соответствия раскодированных данных и значений реквизитов, нанесенных на сам документ, что достигается организационными мерами, а именно участием сотрудника, регистрирующего входящие документы.

Логичным завершением жизненного цикла документа являются его уничтожение или списание в архив. Для документов на бумажных носителях уничтожение документа означает безвозвратное уничтожение содержащейся в нем информации. Что же касается электронных документов, то их уничтожение в контексте работы любой АИС крайне редко приводит к такому результату. Так, в зависимости от технологии работы системы, копии электронного документа могут храниться во временных директориях клиентских и серверных компьютеров, на серверах внешних подразделений и т.д., что может привести к нарушению конфиденциальности информации. В свою очередь, это означает, что работа должна быть построена так, что временные файлы и документы во внешних подразделениях своевременно уничтожаются, при этом информация все равно остается в резервных копиях баз данных. Задача поиска документов, подлежащих уничтожению, в резервных копиях может быть решена, но ее сложность столь велика, что она никогда не ставится. Вариантом решения может быть использование регламента, предусматривающего периодическое уничтожение резервных копий баз данных с таким расчетом, чтобы всегда оставалась возможность восстановления актуального состояния системы.

Архив СЭД представляет собой долговременное хранилище электронных документов, его БД и файловое хранилище могут располагаться на более надежных носителях информации со своим особенным регламентом резервного копирования. При помещении документа в архив изменяется состав его реквизитов, способы группировки документов в соответствии с номенклатурой дел предприятия, права доступа, логика работы с документами. При проектировании обеих БД – СЭД и архива – необходимо предусмотреть возможность автоматического преобразования значений реквизитов БД СЭД в значения реквизитов БД архива.

После помещения документа в архив права доступа к нему полностью пересматриваются. Так, запрещено редактирование как самого документа, так и всех материалов, внесенных в архив вместе с ним. Редактировать в дальнейшем можно только некоторые архивные реквизиты, такие как история работы с документом в архиве или срок его хранения, кроме того, для некоторых типов документов, например, приказов, разрешено добавлять новые файлы. Доступ к документу на чтение и редактирование некоторых реквизитов имеет ответственный за документ, который имеет право выдать документ для ознакомления другим пользователям. Аналогичные права должны быть у руководителя организации или подразделения.

При определении правил работы с документами в архиве СЭД необходимо решить, можно ли разрешать сотруднику, получившему документ, копировать полученные материалы. В случае запрета копирования это может быть реализовано как организационными мерами, так и программно. Также следует решить вопрос о правах доступа к архивным документам сотрудников, которые прежде с ними работали, в том числе автору документа. Представляется, что для этих пользователей возможно разрешить работу с документом без отдельного разрешения руководителя, то есть ответственный за документ или архивное дело выдает им документ по первому требованию. В любом случае все действия по выдаче

документов из архива протоколируются, их обоснованность может быть проверена администраторами системы и руководителями.

Список литературы:

1. О дополнительных мерах по обеспечению информационной безопасности Российской Федерации. Указ Президента Российской Федерации от 1 мая 2022 года № 250.

2. О персональных данных. Закон Российской федерации от 27 июля 2006 года N 152-ФЗ.

3. Об установлении запрета на допуск программного обеспечения, происходящего из иностранных государств, для целей осуществления закупок для обеспечения государственных и муниципальных нужд. Постановление Правительства России от 16 ноября 2015 г. № 1236.

4. Г.П. Акимова, А.Ю. Даниленко, Е.В. Пашкина, М.А. Пашкин, А.А. Подрабинович, И.В. Туманова. Об одном подходе к обеспечению безопасности данных в информационной системе средствами ОС и СУБД. // Информационные технологии и вычислительные системы. 2022, № 1, С.33-39. DOI 10.14357/20718632220104.

Балябин А.А., Новиков В.А., Петренко С.А.
ООО «Технологии Радиоконтроля»

МЕТОД ИММУННОГО ОТВЕТА НА РАНЕЕ НЕИЗВЕСТНЫЕ ВРЕДОНОСНЫЕ ВОЗДЕЙСТВИЯ

В условиях беспрецедентного роста количества угроз безопасности и кибератак злоумышленников становится очевидной недостаточная эффективность существующих классических алгоритмов, методов и средств защиты информации. Так критическая информационная инфраструктура Российской Федерации имеет сложную многоуровневую организацию, что снижает ее прозрачность и усложняет интеллектуальное управление. Как следствие, возникает потенциальная опасность наличия скрытых деструктивных программно-аппаратных закладок и недостатков программного обеспечения на различных уровнях системы.

Применяемые на сегодняшний день подходы к обеспечению надежности и отказоустойчивости (реконфигурация, n-кратное резервирование, сравнение с эталоном) не способны предотвратить возможные критические последствия для информационной инфраструктуры в случае реализации угроз безопасности. Кроме этого, постоянно появляются новые способы обхода средств защиты и способы атаки. В соответствии с результатами исследований, приведенными в [1], около 40% от общего количества кибератак являются новыми, ранее неизвестными кибератаками, которые не могут быть обнаружены.

Очевидно, что появление новых уязвимостей и способов их эксплуатации неизбежно по причине постоянного роста сложности программного и аппаратного обеспечения информационных систем. С другой стороны, очевидна необходимость обеспечения требуемой их устойчивости и надежности.

С учетом данного противоречия перспективной является идея использования биоинспирированных подходов, в частности, наделения информационных систем свойствами иммунитета по аналогии с иммунитетом живого организма [2], для эффективного противодействия как известным, так и ранее неизвестным кибератакам злоумышленников, и упреждения их последствий. Принципиальное отличие данного подхода от существующих заключается в наличии способности накапливать «иммунную память» к уже встречавшимся и

вновь появляющимся кибератакам, планировать «иммунный ответ» и осуществлять самовосстановление систем в реальном масштабе времени.

Общая схема предлагаемого метода представлена на рисунке 1. В исполняемый код программы на этапе трансляции встраиваются контрольные точки, сформированные на основе соотношений подобия [3], предназначенные для контроля целостности вычислений на критических участках программы. Информация о контрольных точках, эталонных соотношениях подобия (инвариантах), допустимых маршрутах выполнения составляет цифровой паспорт программы.

В процессе функционирования кибериммунной системы защиты, противодействия выявляемым кибератакам злоумышленников в системе появляется информация о типах и характеристиках воздействия. Для ее накопления с целью оперативного распознавания и реагирования на угрозы в будущем, в систему иммунной защиты входит подсистема хранения новых знаний кибериммунитета.

При обнаружении нарушения целостности вычислений осуществляется его классификация. В случае, если информации о вредоносном воздействии не содержится в базе данных кибериммунитета, запускается процедура самообучения и формирования новых знаний о выявленных нарушениях. По результатам анализа нарушения осуществляется синтез микропрограмм и запускается процедура восстановления искаженных вычислений.

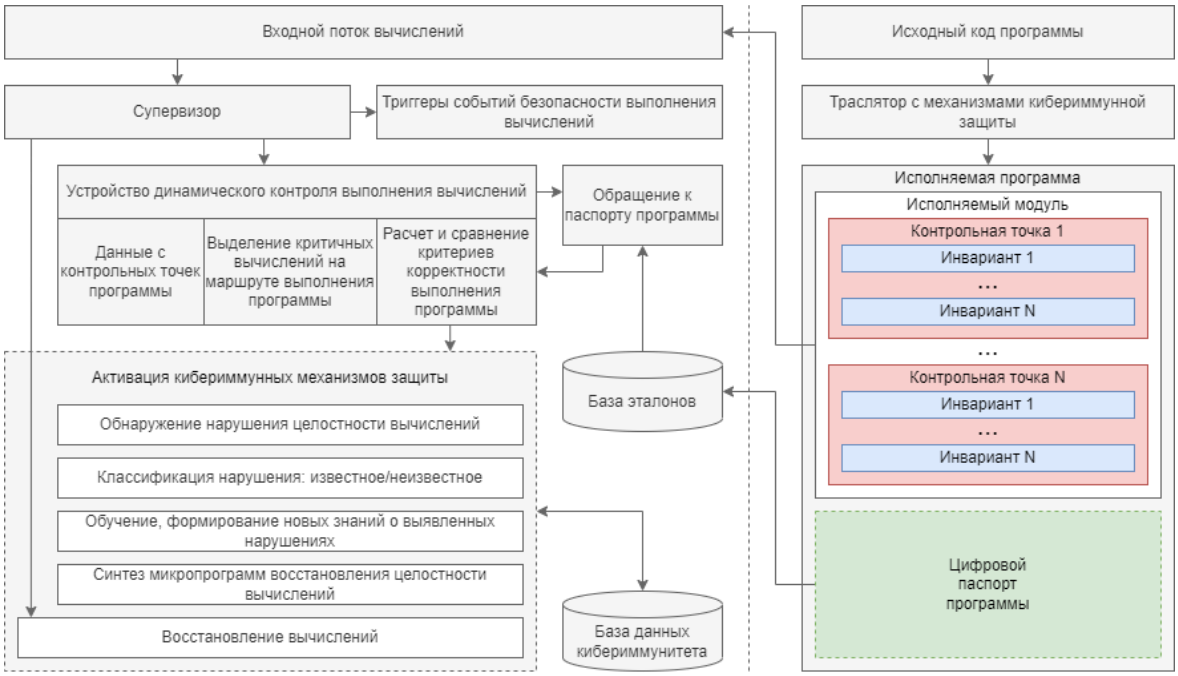

Рисунок 1 – Схема метода иммунной защиты

Предложенный в работе метод иммунного ответа позволит выявлять аномалии поведения систем, возникшие в результате деструктивных воздействий (в том числе и ранее неизвестных, за счет реализации механизмов иммунной защиты), противодействовать им, осуществлять самовосстановление параметров поведения, влияющих на киберустойчивость системы, а также накапливать знания о воздействиях для повышения эффективности реализации «иммунного ответа» на вторжения в будущем.

Список литературы

1. Петренко С. А. Кибериммунология: научная монография / Петренко С. А. – СПб: «Издательский Дом «Афина». 2021. 240 с.
2. Марчук Г. И. Математические модели в иммунологии: вычислительные методы и эксперименты. М.: Наука. 1991. 299 с.
3. Ковалев, В. В. Верификация программ на основе соотношений подобия / В. В. Ковалев, Р. И. Компаниец, В. А. Новиков // Труды СПИИРАН. – 2015. – № 1(38). – С. 233-245.

Завадский Е.В., Калинин М.О.
Санкт-Петербургский политехнический университет Петра Великого

МЕТОД СОЗДАНИЯ ЦИФРОВОГО ДВОЙНИКА ДЛЯ ПОВЫШЕНИЯ ЗАЩИЩЕННОСТИ СЕТЕЙ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

** Работа выполнена в рамках Государственного задания на проведение фундаментальных исследований код темы 0784-2020-0026*

Современная информационная система объекта КИИ, которая представляет собой единое киберпространство, является объединением множества различных цифровых сервисов, необходимых для решения поставленных задач. Структура данной сети динамична. Это обусловленно необходимостью непрерывной адаптацией к новым запросам – внедрение новых сервисов, выполнение реконфигурации существующих, построение новых взаимосвязей между ними. В условиях злонамеренных воздействий со стороны внешних и внутренних нарушителей, а также растущего числа уязвимостей [1] в программно-аппаратном обеспечении требуются новые комплексные подходы предотвращению кибератак.

В рамках данного подхода предложен метод реализации цифрового двойника сетевой инфраструктуры. На основе результатов анализа существующих решений [2-4] были определены их основные преимущества и недостатки, а также выделены наиболее актуальные характеристики – обеспечение высокой интерактивности узлов и возможность масштабирования.

В работе сформулирован ряд условий, позволяющих оптимизировать потребление вычислительных ресурсов цифровым двойником при сохранении полного соответствия его эксплуатационных свойств защищаемой сети. При помощи моделирования развития различных атак были получены параметры функционирования системы при использовании как каждого условиях в отдельности, так и совместно. На основе полученных результатов разработан метод построения виртуальной сетевой инфраструктуры цифрового двойника и осуществления динамического управления её топологией.

Предложенный метод позволяет обеспечить увеличение масштаба реализуемой виртуальной сети в сочетании с эффективным использованием ресурсов. Применение данного метода при реализации систем обнаружения вторжений и Honeypot-систем позволяет создать виртуальную сеть, неотличимую от реальной с точки зрения злоумышленника.

Предложенный метод позволяет обеспечить увеличение масштаба реализуемой виртуальной сети в сочетании с эффективным использованием вычислительных ресурсов. Он может быть использован для реализации следующих приложений безопасности:

1. Honeypot-система, виртуальная сеть которой неотличимую от реальной с точки зрения злоумышленника.

2. Система обнаружения вторжений, которая функционирует как «цифровая тень» реальной сети, в точности соответствуя её поведению.

Список литературы:

1. NIST National Vulnerability Database Statistics Results [Электронный ресурс]. URL: https://nvd.nist.gov/vuln/search/statistics?form_type=Basic&results_type=statistics&search_ type=all.

2. Franco J. et al. A survey of honeypots and honeynets for internet of things, industrial internet of things, and cyber-physical systems //IEEE Communications Surveys & Tutorials. – 2021. – Т. 23. – №. 4. – С. 2351-2383.

3. Ovasapyan T. D., Nikulkin V. A., Moskvin D. A. Applying Honeypot Technology with Adaptive Behavior to Internet-of-Things Networks Automatic Control and Computer Sciences 2021, 55(8), 1104–1110.

4. Piggin R., Buffey I. Active defence using an operational technology honeypot 11th International Conference on System Safety and Cyber-Security (SSCS 2016) 2016, 1–6.

Завадский Е.В., Зегжда Д.П., Калинин М.О.
Санкт-Петербургский политехнический университет Петра Великого

ПРЕДИКТИВНАЯ ЗАЩИТА ОТ КИБЕРАТАК ГИБКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ ПРОМЫШЛЕННЫХ ОБЪЕКТОВ НА БАЗЕ ТЕХНОЛОГИИ АДАПТИВНОЙ ОБМАННОЙ СИСТЕМЫ

** Работа выполнена в рамках Государственного задания на проведение фундаментальных исследований (код темы 0784-2020-0026).*

Информационные системы современных промышленных предприятий в рамках модели Индустрии 4.0 претерпевают значительные изменения, связанные с внедрением интеллектуальных технологий управления. Основной целью модернизации является повышения показателей эффективности: снижения простоя оборудования, оптимизации цепочек поставок, удаленного мониторинга производственных процессов и т.п.

При перманентных изменениях параметров протекающих процессов эффективность производства во многом зависит от уровня гибкости информационной системы и возможностей по её адаптации к новым условиям. Внедрение концепции IIoT подразумевает объединение операционных и информационных технологий. Возможность точной настройки параметров системы обеспечивается за счет сбора большого объёма аналитических данных с интегрированных во все производственные компоненты сенсоров, его анализа и передачи управляющих сигналов при помощи межмашинного взаимодействия (M2M).

Повышение автономности управления производственными процессами влечет повышение рисков, связанных с негативными последствиями от деструктивных воздействий со стороны злоумышленников. По данным Positive Technologies за 4 квартал 2021 50% кибератак привело к нарушению основной деятельности предприятий. Необходимы новые комплексные подходы к обеспечению безопасности производственных объектов, учитывающих разнородность программных и аппаратных характеристик их компонентов.

В рамках данного подхода предложен метод на базе технологии адаптивной обманной системы, которая позволяет повысить эффективность использования вычислительных ресурсов при развертывании виртуальной сети и сохранить высокую интерактивность (обеспечение возможности компрометации атакующим любого узла сети). Это достигается за счет применения моделируемого и виртуализируемого образа для каждого узла сети в сочетании с интеллектуальным алгоритмом их переключения, основанного на графовой структуре описания потенциальных кибератак.

Для обеспечения выявления злонамеренных воздействий на компоненты IIoT на каждом уровне сетевой топологии размещаются виртуальные узлы обманной системы, которые являются индикаторами аномального поведения и обеспечивают смещение внимания злоумышленника от ресурсов реальной сети за счет наличия предусмотренных недостатков безопасности. Таким образом достигается локализация кибератаки и перенаправление её дальнейшего развития в виртуальную сетевую инфраструктуру обманной системы.

Соболев Н.В., Зегжда Д.П.
Санкт-Петербургский политехнический университет Петра Великого

ПОСТРОЕНИЕ СЕТИ С HONEYPOT НА ОСНОВЕ КЛАССИФИКАЦИИ ТРАФИКА С ПОМОЩЬЮ LSTM

Системы обнаружения вторжений (IDS) являются важным элементом систем сетевой безопасности. Из-за больших объемов проходящего сетевого трафика существует необходимость быстрой и надежной фильтрации потенциально опасного трафика от безопасного.

IDS отслеживает трафик, сравнивая его с собственной базой данных возможных сетевых атак и нормальной сетевой активностью. Такой механизм работы позволяет обнаруживать:

- сетевые атаки;
- неавторизованный доступ к данным;
- действия вредоносных скриптов и программ;
- функционирование сканеров портов;
- нарушение политик безопасности;
- обращение к центрам управления бот-сетями и майнинг-пулам;
- аномальную активность.

Важно заметить, что IDS-система не отражает атаки, а только обнаруживает их и уведомляет о них администратора, помогая найти причину и устранить ее.

Обнаружить нарушения политик безопасности можно за счет написания своих собственных паттернов детектирования. Это помогает отслеживать определенное поведение в сети.

Предложенный в данном исследовании метод организации локальной сети предоставляет два уровня защиты хостовой машины от атак со стороны злоумышленника из сети Интернет. Где первый уровень – IDS, в основе которой лежит нейронная сеть LSTM,

реализующая обнаружение аномального поведения трафика, второй уровень – «обманный» ресурс honeypot, на который IDS переводит трафик в случае обнаружения аномалии.

Модель поведения нарушителя, а также используемые для проникновения инструменты, позволяет определить встроенный в локальную сеть honeypot, на который LSTM переводит все сетевые пакеты при обнаружении аномального трафика.

Общая схема сети с LSTM и honeypot показана на рисунке 1.

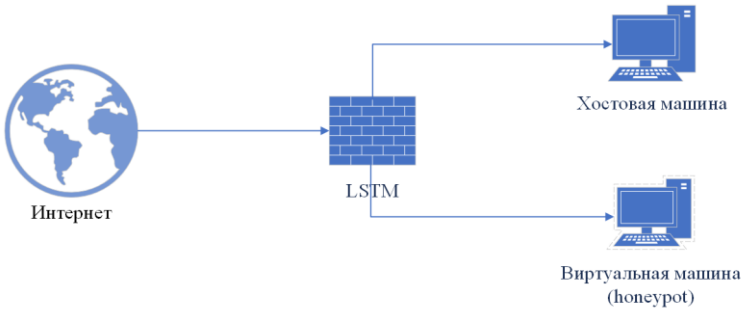

Рисунок 1 – Общая схема сети с LSTM и honeypot

Список литературы:

1. Wira Zanoramy Zakaria, ML Mat Kiah, A review on artificial intelligence techniques for developing intelligent honeypot. Доступ по ссылке: https://www.researchgate.net/publication/255995332_A_review_on_artificial_intelligence_techniques_for_developing_intelligent_honeypot.

2. Aidan Mitchell, An Intelligent Honeypot. Доступ по ссылке: https://www.researchgate.net/publication/326995569_An_Intelligent_Honeypot.

3. Merity S. et al. Pointer sentinel mixture models //arXiv preprint arXiv:1609.07843. – 2016. Доступ по ссылке: https://arxiv.org/abs/1609.07843

Кузинков А.М., Пилькевич С.В.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

ПРОБЛЕМА ИНТЕРОПЕРАБЕЛЬНОСТИ В СОВРЕМЕННЫХ АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ

Интенсивность роста количества информации достигает невероятных размеров. По данным американского сайта https://www.statista.com/ к 2025 году общемировой объем данных вырастет на порядок и достигнет 163 зетта байт, причем большая их часть будет создаваться не частными пользователями, а организациями.

Важным фактором, позволяющим достичь синергетического эффекта от совместного использования современных автоматизированных систем, является обеспечение интероперабельности. Согласно общепринятому определению, указанному в ГОСТ Р 55062-2012, интероперабельность –это способность двух или более информационных систем или компонентов к обмену информацией и к использованию информации, полученной в результате обмена. Интероперабельность служит ключевым принципом совместимости современных автоматизированных систем, как свойство бесшовной информационной интеграции её отдельных элементов и подсистем.

Проблема интероперабельности характерна для всех областей применения информационных технологий различных масштабов: от технологии умного дома до обработки огромных массивов данных при активной цифровизации общества.

Для регламентирования вопросов интероперабельности более чем в 30 странах существуют документы «e-Gov Interoperability Framework». В нашей стране разработан ГОСТ Р 55062-2012 «Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения». В котором отражено понятие интероперабельности в автоматизированных системах.

Интероперабельность в автоматизированных системах предполагает соблюдение определенных правил или привлечение дополнительных программных средств, обеспечивающих возможность взаимодействия модулей, подсистем или программных систем[1]. Правила и средства - это своего рода стандарты, которым должны удовлетворять интегрируемые системы.

Принято различать три уровня интероперабельности современных автоматизированных систем – техническую, синтаксическую и семантическую, которым соответствуют транспортная среда, формат сообщений и смысл данных[2].

На западе были разработаны три технологии обеспечения интероперабельности автоматизированных систем:

1. Модель уровней интероперабельности автоматизированных систем - LISI-модель (Levels of Information Systems Interoperability Model);

2.Модель оценки интероперабельности систем, возможностей, действий, программ и организаций - SCOPE-модель (Systems, Capabilities, Operations, Programs, and Enterprises model for interoperability assessment);

3.Концепцияпо архитектуре МО США – DODAF (DOD Architecture Framework)[3].

Назначение LISI - обеспечить министерство обороны соответствующей моделью зрелости и процессом для определения общих возможностей интероперабельности информационных систем, оценки способности систем соответствовать предъявляемым требованиям. USI - это процесс определения, измерения и оценки интероперабельности информационных систем. USI использует общие рамки и показатели результативности [4].

Подробнее остановимся на SCOPE-модели в соответствии с концепцией понятие сетецентрических операций применимо к любой прикладной области, в которой используется сетевая архитектура и реализуются преимущества человеко-машинного информационного взаимодействия. К таким «гражданским» областям применения SCOPE-модели можно отнести:

– государственное и корпоративное управление;

– логистические системы;

– аппаратные и программные комплексы, построенные на основе сервис-ориентированной архитектуры SOA (Service Oriented Architecture) [1].

В Российской Федерации на основе ГОСТ Р 55062-2012 «Информационные технологии. Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения» для формирования подуровней и большего числа параметров в проблемно-ориентированных моделях может использоваться международный опыт формализации интероперабельности, представленный в вышеописанных LISIи SCOPE-моделях.

Таким образом полученная интегрированная модель интероперабельности должна быть оформлена в виде национального стандарта ГОСТ Р, а ее использование унифицировано в отечественной практике построения автоматизированных систем в рамках реализации единого информационного пространства.

В заключение стоит отметить, что в при реализации и внедрении подходов к оцениванию степени интероперабельности автоматизированных систем и обмена данными будут достигнуты следующие положительные эффекты:

– повышение степени координации разнородных взаимодействующих автоматизированных систем и используемых ими информационных ресурсов создаст предпосылки к реализации интероперабельности указанных систем;

– последующее развитие описанных систем автоматически приведет к поддержке универсальной переносимости приложений на платформах различных типов;

– полученные оценки, описывающие и категоризирующие степени интероперабельности взаимодействующих автоматизированных систем, позволят количественно охарактеризовать уровни возрастающей сложности и возможностей систем. Каждая степень интероперабельности отражает определенные аспекты взаимодействия автоматизированных систем.

Список литературы

1. Макаренко, С.И. Модели интероперабельности информационных систем / С.И. Макаренко, А.Я. Олейников, Т.Е. Черницкая. — Текст: непосредственный // Системы управления, связи и безопасности. — 2019. — № 4. — С. 215-245.
2. Быстров, Р. П. Актуальное состояние проблемы интероперабельности / Р.П. Быстров, А.А. Каменщиков,А.Я. Олейников. — Текст: непосредственный // ИТ-СТАНДАРТ. — 2018. — № 3. — С. 23-28.
3. DOD Architecture Framework. Version 1.5. — Текст: электронный // Defenselink: [сайт]. — URL: http://www.defenselink.mil/cio-nii/docs/DoDAF_Volume_I.pdf (датаобращения: 21.05.2022).
4. Куприянов, А.А. Сетецентрические военные действия и вопросы интероперабельности автоматизированных систем / А.А. Куприянов. — Текст: непосредственный // Автоматизация процессов управления. — 2011. — № 3. — С. 82-97.

Шаваньгина О.В.
ООО «СЛ Технологии»

КИБЕРБЕЗОПАСНОСТЬ СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ

На данный момент в России установлено 13,5 мл камер.

1. Места уязвимости в системах видеонаблюдения

— Каналы и линии связи

— Архивные записи

— Устройства для записи видео и звука

2. От чего нужно защищаться

— От несанкционированного доступа к информации

— От умышленной порчи и искажения данных

— От захвата и управления информацией при несанкционированном просмотре

3. Меры защиты информации

— Идентификация и аутентификация пользователей

— Отказ от универсальных паролей, с использованием персональных данных

— Ограничение количества пользователей в системе

— Регулярный контроль защищенности информации

— Защита на уровне коммутационного оборудования

— Создание прошивки отечественного происхождения для устройств записи видео и звука

4. Тесты на проникновение

— Тестирование устройств для записи видео и звука

— Тестирование протокола

— Тестирование ПО

5. Отказ от пароля и логина по умолчанию

— Создание системы кибербезопасности, как мера защиты информации

6. Принцип минимальных прав

— Разграничение прав доступа и управления системой среди пользователей

7. Способы обеспечение кибербезопасности в видеонаблюдении

— Защита программного обеспечения

— Защита коммутационного оборудования

— Собственная защита устройств для записи видео и звука

Фатин А.Д., Павленко Е.Ю., Зегжда Д.П.
Санкт-Петербургский политехнический университет Петра Великого

ИММУНИЗАЦИЯ ДИНАМИЧЕСКИХ СЕТЕЙ В ЗАДАЧАХ КИБЕРБЕЗОПАСНОСТИ

** Исследование выполнено в рамках гранта Президента РФ для государственной поддержки молодых российских ученых – кандидатов наук МК-3861.2022.1.6.*

С ростом интереса научного сообщества к вопросам искусственного интеллекта, адаптивности киберфизических систем, динамической кластеризации и ситуационного управления, все большую популярность набирает направление иммунизации сложных компьютерных систем, представляющее собой пересечение вышеперечисленных областей знаний и практик.

В докладе рассматривается актуальность и применимость подходов иммунизации сложных сетей в задачах кибербезопасности, исследуется возможность существования вирусного равновесия в киберфизических системах, а также особое внимание уделяется иммунизации динамических сетей и способов ее реализации.

Также в перечень рассматриваемых и анализируемых тем входят следующие модели иммунизации сложных сетей:

— сегментная модель [1];

— модель цикличных переходов [2];

— модель растущей безмасштабной сети [3];

— P2P модель со статической топологией [4].

Выделяются основные преимущества и области применения описанных моделей, рассматриваются основные системы дифференциальных уравнений, применяемых для описания стохастических процессов иммунизации. Дополнительно количественно и качественно оцениваются наиболее распространенные методы иммунизации (таргетированная, случайная, по знакомству и прочие), проводится их сравнительный анализ; рассматриваются оптимальные способы реализации иммунизации с помощью жадных алгоритмов и моделей SIR (упрощенной и полной).

Список литературы:

1. Modeling the spread of malware with the influence of heterogeneous immunization / W. Liu, Ch. Liu, X. Liu, Sh. Cui, X. Huang // Applied Mathematical Modelling. – 2016. – Vol. 40. – Issue 4. – pp. 3141-3152. DOI: 10.1016/j.apm.2015.09.105.

2. 7 Virus Propagation on Time-Varying Networks: Theory and Immunization Algorithms / B. A. Prakash et al. // Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2010. Lecture Notes in Computer Science, Vol 6323, 2010. – pp. 99-114. DOI: 10.1007/978-3-642-15939-8_7.

3. Recoverable prevalence in growing scale-free networks and the effective immunization / Yu. Hayashi, M. Minoura, J. Matsukubo. DOI: 10.1103/PhysRevE.69.016112.

4. Bahashwan W. S., Al-Tuwairqi S. M. Modeling the Effect of External Computers and Removable Devices on a Computer Network with Heterogeneous Immunity / W. S. Bahashwan, S. M. Al-Tuwairqi // International Journal of Differential Equations. –2021. – Vol. 260. – pp. 1-13. DOI:10.1155/2021/6694098.

Сикарев И.А., Большаков В.А., Коринец Е.М.
Российский государственный гидрометеорологический университет

К ВОПРОСУ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ

В настоящее время роль кибербезопасности в работе геоинформационных систем продолжает расти по мере того, как все больше сфер охватывает цифровизация. Мощные географические информационные системы быстро преобразуют один или несколько слоев цифровых геопространственных данных в комплексные, подобные картам. Эти системы могут облегчить выполнение широкого спектра соответствующих геопространственных анализов в режиме, близком к реальному времени. Текущее состояние геопространственных

информационных технологий может предоставить лицам, принимающим решения, данные, необходимые им для уверенного противостояния широкому спектру угроз, включая стихийные бедствия, террористические атаки, саботаж и подобные кризисы. Однако текущая реализация этих технологий во всех федеральных, государственных и местных агентствах и юрисдикциях, необходимая для полной координации эффективного реагирования, практически отсутствует в конкретных областях. По мере того, как концепция национальной безопасности внедряется в повседневную работу правительства и повседневную жизнь общества, лица, принимающие решения, получат большую выгоду от «преимущества» кризисного управления, которое предоставляет ГИС. Руководители службы национальной безопасности должны понимать и внедрять изменения, необходимые для полной реализации возможностей этой технологии, и принимать управленческие решения, необходимые для ее организации на национальном уровне. Геоинформационные системы объединяют ресурсы, которые могут помочь организациям проанализировать потенциально скомпрометированные системы и разработать более надежную защиту. Эти системы можно использовать для доступа и обработки цифровых геопространственных данных практически в любом месте, поскольку они, в отличие от аналоговых данных, могут быть мгновенно переданы из любого места, где они хранятся, и передаваться в любое место, где они необходимы. Таким образом, данные системы можно применять для решения проблем по предотвращению растущего множества киберугроз. Перспективным направлением является комплексный подход к обеспечению безопасности, включающий в себя, создание геоинформационной системы обнаружения кибератак в режиме реального времени, что предполагает оперативный доступ к локализации угроз, а также применение шифрования данных с использованием криптографических методов. Оперативность получения информации позволяет своевременно применить обновленные меры противодействия такого вида угрозам. Пространственная информация также позволяет специалистам по безопасности обнаруживать несанкционированную активность на раннем этапе. Чтобы свести к минимуму последствия утечки данных или атаки вредоносного ПО, заинтересованные стороны должны координировать немедленные ответные действия. ГИС может обеспечить четкую визуализацию систем, вовлеченных в инцидент, и повысить ситуационную осведомленность нескольких отделов безопасности. Детальное представление о потоке данных через сеть любой организации позволяет получить оперативную информацию о любых сбоях устройств, которые могут помешать работе. Пространственная информация связывает возникновение киберугрозы с конкретными участками, позволяя специалистам проанализировать качество возникшей угрозы, ее связь с преднамеренной попыткой скомпрометировать систему, а также оценить предполагаемые последствия. Также специалисты по кибербезопасности и ИТ могут однозначно расставить приоритеты и решительно предотвратить несанкционированные вторжения. Эти характеристики делают географические информационные системы в сочетании с соответствующими наборами геопространственной информации бесценным инструментом для обработки, отображения и анализа информации, связанной со всеми аспектами национальной безопасности.

Фомичева С.Г., Беззатеев С.В.
Санкт-Петербургский университет аэрокосмического приборостроения, Санкт-Петербург

МЕТОД ОБЪЯСНЯЕМОГО ИЗВЛЕЧЕНИЯ ЗНАНИЙ SIEM-АГЕНТАМИ

Динамика развития SIEM (Security Information and Event Management) - решений во многом связана с масштабным применением интеллектуальных технологий и эффективных алгоритмов обработки больших данных. Сбор данных SIEM-системой, как правило, осуществляется с помощью специальных программных (автономных) агентов, которые

локально собирают журналы событий безопасности на конечных точках (endpoints) и по возможности передают их на сервер-коллектор. Процессы сбора и анализа событий безопасности ожидаются стать в дальнейшем практически полностью автоматическими за счет инкапсуляции механизмов мониторинга и реагирования на инциденты в базовый состав корпоративных IT-средств. (SCADA, MES, ERP и т.л.). Это в свою очередь приводит к возможности формирования встроенных искусственных иммунных систем и появлению новых сущностей – мобильных агентов безопасности, способных мигрировать в корпоративных информационно-телекоммуникационных сетях. Функционал мобильных агентов может значительно различаться – от выполнения простого транспортирования информации в SIEM-хранилища, до самостоятельного выявления аномалий на пути своей миграции. В последнем случае мобильный агент способен частично выполнять функционал автономных и управляющих агентов, и, в перспективе, функционал сервера корреляции. В то же время накладывается ряд требований к архитектуре мобильных агентов, способам анализа ими информации и извлечения знаний. К числу таких требований относятся – необходимость использования объяснимых методов извлечения знаний (в силу того, что агент должен оставаться обоснованно доверенным для своего владельца), самообучение (в силу частого появления раннее неизвестных аномалий), компактность (в силу потребности к миграции агента), унификация математических принципов, используемых как для извлечения знаний, так и для крипто- и имитозащиты знаний и данных мобильных агентов.

Нами предлагается само объясняемый метод извлечения знаний, основанный на модификации широко используемого в криптографии алгоритма Бэрлекэмпа-Месси. Он используется мобильным агентом для извлечения закономерностей как из внешних хранилищ, так и из переносимых им же данных, что позволяет в рамках выполнения целевой задачи, с одной стороны, выполнять функцию обобщения данных (задача уменьшения размерности), а с другой, при необходимости, освобождать собственные ресурсы памяти, делая выбор в пользу переноса знаний. Возможность применения модифицированного алгоритма Беэрмекэмпа-Месси базируется на доказательстве изоморфности нечетких продукционных и нечетких реляционных баз знаний, аппроксимация которых приведенными полиномами в полях Галуа доказана авторами в [1]. На сегодняшний день разработаны различные алгоритмы для объяснения систем искусственного интеллекта. Их подробный аналитический обзор представлен в [2]. В контексте решаемых нами задач основное внимание уделялось методам на основе само-объясняемых моделей (Self-explainable), в которых сами алгоритмы дают объяснение. Точность объяснения важна и для само-объясняемых моделей она часто может быть измерена, и даже может быть управляема. В частности, в работах [1,3] нами рассматриваются управляемость точностью объяснений с помощью алгоритмов квантования знаний мобильных агентов безопасности.

Агенты безопасности в SIEM-решениях выполняют свой функционал по крипто анализу, выполняя операции в конечных полях. Например, используют алгоритм Бэолекэмпа-Месси в своих задачах декодирования блоковых кодов и/или поиска эквивалентных регистров сдвига, генерирующих ключевые последовательности при симметричном крипто кодировании. В данной работе мы хотим показать, что данный функционал можно использовать и для извлечения знаний. Доказательство того, что нейро-нечеткие модели работоспособны в конечных полях приведено нами в [1]. Конструктивный результат доказанных в [1] теорем заключается в возможности применения математических принципов обработки информации в конечных полях в задачах машинного обучения, и, в частности, гибридных (нейро-нечетких) моделях.

В силу верности теорем, доказанных в [1] предлагается конструктивный метод извлечения знаний и консолидации информации на основе модифицированного алгоритма Берлекэмпа-Месси. Знания представляют собой искомые приведенные многочлены, определяющие обратные связи эквивалентного линейного регистра сдвига. Под консолидацией информации в данном контексте подразумевается последовательность

трансформации баз знаний: «мгновенная реакция» агента (нечеткие отношения) - оперативные знания агента (нечеткие продукции) – тактические знания агента (нечеткие тенденции) – стратегические знания агента (абстракции).

Основная идея предлагаемой нами модификации алгоритма Берлекэмпа-Месси заключается во включении в итеративный процесс вычисления невязок априорно заданного индекса эластичности ϕ (допустимой неточности). Индекс эластичности может быть проинтерпретирован как порог бдительности в ART-моделях [2]. Поскольку для выявления нечеткой тенденции не требуется абсолютного совпадения генерируемой последовательности с исходной, то модифицируя алгоритм Берлекэмпа-Месси при выявлении закономерностей в нечеткой реляционной базе знаний и очередную итерацию мы считаем законченной, если абсолютное значение невязки откланяется на величину $|\delta| \le \left\lfloor \frac{q}{\varphi} \right\rfloor$, где $q > 2$ - характеристика поля Галуа, в континууме которого функционирует мобильный агент безопасности. Полученные нами экспериментальные результаты [4] реализации агентов в SIEM-решениях позволяют отдать предпочтении к использованию модифицированного алгоритма Бэрлекэмпа-Месси при работе не только мобильных, но и при эксплуатации автономных и управляющих агентах в SIEM-системах. Работы по выбору метрик и оценки точности объяснений в настоящее время нами продолжаются.

Библиографический список

1. Fomicheva S. Soft quantization of the production's knowledgebases for multi-agent systems. "Proceedings of the 20th Conference of Open Innovations Association, FRUCT 2017" 2017. С. 69-76. DOI: 10.23919/FRUCT.2017.8071294

2. P. Jonathon Phillips, Carina A. Hahn. Peter C. Fontana, David A. Broniatowski, Mark A. Przybocki. Four Principles of Explainable Artificial Intelligence. https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8312-draft.pdf (Accessed 25.01.2022)

3. Фомичева С.Г. Теоретические аспекты квантования баз знаний в мультиагентных системах, Информационно-управляющие системы. 2017. Т. 88. № 3. С. 2-10. DOI:10.15217/issnl684-8853.2017.3.2

4. Bezzateev, S.V., Fomicheva, S.G., Zhemelev, G.A. Agent-based zerologon vulnerability detection. 2021 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2021 - Conference Proceedings, 2021, 9470548

Крундышев В.М., Калинин М.О.
Санкт-Петербургский Политехнический университет Петра Великого

ОПИСАНИЕ ДИНАМИКИ РАЗВИТИЯ КОМПЬЮТЕРНЫХ АТАК НА ОСНОВЕ РАСШИРЕНИЯ БАЗОВОЙ МОДЕЛИ ЛОТКИ-ВОЛЬТЕРРЫ

** Исследование выполнено в рамках стипендии Президента РФ для поддержки молодых ученых и аспирантов (СП-2714.2021.5)*

Несмотря на многообразие и доступность различных механизмов защиты критическим является вопрос соответствия скорости обнаружения и блокирования компьютерных атак скорости распространения вредоносного ПО в информационных системах (ИС) [1, 2]. Для оценки эффективности применяемых методов обнаружения и их адекватности изменяющимся

параметрам ИС и действующим киберугрозам возникает потребность в создании математической модели развития компьютерных атак.

В рамках решаемой задачи предлагается провести аналогию и транслировать модель Лотки-Вольтерры из экологических терминов в термины информационной безопасности [3]. Изменение популяций легитимных узлов и узлов-злоумышленников позволит отразить изменение состояния ИС под воздействием компьютерных атак. В качестве жертв будут выступать легитимные узлы, в качестве хищников – злоумышленники.

Для описания предлагаемой модели определены следующие понятия:

- жертва – узел, который выполняет свою целевую функцию и не осуществляет деструктивных воздействий по отношению к ИС;
- хищник – узел, который осуществляет целенаправленные компьютерные атаки на ИС, маршрутизацию и отдельные узлы-жертвы;
- инфицированный узел – устройство «зомби», которое осуществляет деструктивное воздействие на ИС и находится под управлением злоумышленника;
- вакцинированный узел – устройство, которое в период действия вакцины является более защищённым от компьютерных атак за счет использования дополнительных средств обеспечения безопасности;
- выведенный из строя узел – устройство, неспособное в полной мере выполнять целевую функцию (в терминах модели Лотки-Вольтерры – мертвая жертва).

Множество узлов ИС может быть представлено в виде совокупности пяти классов: Z и Q – хищники, X, W и Y – жертвы.

Узлы классов X и W успешно выполняют целевую функцию, при этом узлы класса X являются «вакцинированными». Узлы класса Y выведены из строя, при этом они не представляют угрозы для других узлов и могут быть либо восстановлены в класс W, либо со временем перейти в состояние D (терминальное состояние, описывающее выход узла из строя). Узлы класса Z находятся под прямым управлением нарушителей, реализуют компьютерные атаки и стремятся перевести узлы из классов X и W в классы Q и Y в зависимости от целей и мотивов. Узлы класса Q являются «узлами-зомби», реализуют компьютерные атаки на узлы классов X и W, стремясь перевести их в класс Y, при этом они могут быть восстановлены в класс W. Узлы класса Y переходят в это состояние, если они не были своевременно восстановлены в класс W, а узлы классов Q и Z при успешной работе системы обнаружения киберугроз.

Узлы классов Q и Y могут быть восстановлены в класс W с вероятностями r_1 и r_2. Узлы класса Z инфицируют узлы X и W с вероятностями τ_1 и τ_2. Под воздействием узлов из классов Q и Z жертвы из классов W и X могут перейти в класс Y с коэффициентами ε_1 и ε_2. μ, a, c – коэффициенты вымирания узлов класса Y, Q, Z. Узлы класса W становятся вакцинированными (переходят в класс X) с вероятностью β. Устойчивость к заражению (вакцинация) пропадает со скоростью прямо пропорциональной коэффициенту η. Узлы классов Q и Z могут вымирать только под воздействием системы обнаружения киберугроз – «охотника», которая напрямую влияет на значение коэффициентов a и c.

N_w, N_X, N_Y, N_Q, N_Z – количество узлов, принадлежащих классу W, X, Y, Q, Z соответственно. w_g, x_g, y_g, q_g, z_g – коэффициенты, характеризующие рост количества узлов в классах W, X, Y, Q, Z соответственно. m_X, m_Y, m_W – вероятность промаха охотника (попадания по узлам из классов X, Y, W).

Согласно предложенной модели, решаемой задачей является максимизация коэффициентов вымирания для классов Z и Q, при этом вероятности промаха должны стремится к нулю. Кроме того, учитывая, что узлы Q могут быть восстановлены в класс W, то относительно класса Q появляется также коэффициент r_1, который наравне с a, должен стремиться к своему максимуму.

Критерием адекватности применяемых методов обнаружения атак изменяющимся условиям функционирования ИС и действующим киберугрозам служит система (1) (скорость вымирания узлов-злоумышленников выше, чем скорость перехода узлов из нормального состояния в выведенное из строя, при этом количество ложных срабатываний не должно превышать λ).

$$\begin{cases} m_W \leq \lambda \\ m_X \leq \lambda \\ m_Y \leq \lambda \\ a > \eta \\ c > \eta \end{cases} \quad (1)$$

Таким образом, в соответствии с разработанной моделью система обнаружения и блокирования компьютерных атак в ИС должна обеспечивать выполнение критерия (1).

Список литературы:

1. Полтавцева, М. А. Формирование структур данных в задачах активного мониторинга безопасности / М. А. Полтавцева // Проблемы информационной безопасности. Компьютерные системы. – 2021. – № 1(45). – С. 9-19.
2. Zegzhda D., Lavrova D., Pavlenko E., Shtyrkina A. Cyber attack prevention based on evolutionary cybernetics approach. Symmetry. –2020. 12 (11), 1931.
3. Романов М. Ф. Математические модели в экологии : Учеб. пособие / М. Ф. Романов, М. П. Федоров. - 2-е изд, испр. и доп. - Санкт-Петербург : Иван Федоров, 2003. - 239 с.

Павленко Е. Ю.

Санкт-Петербургский политехнический университет Петра Великого (СПбПУ)

МОДЕЛИРОВАНИЕ АНТИЦИПАЦИОННЫХ МЕТОДОВ ПРОТИВОДЕЙСТВИЯ КИБЕРУГРОЗАМ ДЛЯ КРУПНОМАСШТАБНЫХ СИСТЕМ С АДАПТИВНОЙ СЕТЕВОЙ ТОПОЛОГИЕЙ

** Исследование выполнено за счет гранта Российского научного фонда № 22-21-20008, https://rscf.ru/project/22-21-20008/*

Развитие беспроводных и сенсорных технологий привело к расширению спектра сетевых топологий, используемых в различных крупномасштабных системах. Особенную популярность получили одноранговые сети, в том числе, с адаптивной топологией, и для них характерны киберугрозы, отличные от угроз традиционной клиент-серверной топологии. Практически все они направлены на несанкционированное воздействие на процесс маршрутизации в сети, что особенно опасно в случае промышленных систем, в которых процесс передачи данных между устройствами и реализация физических процессов интегрированы в единое целое.

Учитывая эпидемиологический характер распространения вредоносного воздействия в таких сетях, особенную актуальность представляет решение научной проблемы упреждающего

выявления киберугроз, характерных для сетей с адаптивной топологией, обеспечивающего запас времени, достаточный для противодействия киберугрозам.

В работе предлагается подход, базирующийся на антиципационном моделировании, где под термином «антиципация» следует понимать способность принимать решения с пространственным или временным упреждением в отношении ожидаемых событий. Таким образом, антиципационное моделирование в некотором смысле есть развитие методов и моделей теории прогнозного моделирования для противодействия киберугрозам крупномасштабным промышленным системам с адаптивной топологией, что обеспечит распознавание кибератак на ранней стадии и увеличит время на принятие решения о реагировании.

Моделирование сетевой инфраструктуры осуществляется с использованием динамических графов, обеспечивающих наглядное отражение изменений, происходящих в сети, в динамике – посредством задания теоретико-графовых операций для каждого перехода сети G_i в новое состояние G_{i+1}: $\varphi(G_i) = G_{i+1}$.

Для получения прогноза предлагается использовать адаптивные модели прогнозирования в сочетании с графовыми нейронными сетями, позволяющими обрабатывать масштабные участки сетевой инфраструктуры, представленные в виде набора графов [1].

Получение прогноза с использованием адаптивных моделей [2] обеспечит знание о состоянии каждого компонента сетевой топологии в следующий момент времени, временной промежуток будет варьироваться от нескольких секунд до нескольких часов, что позволит противодействовать быстрым и краткосрочным кибератакам [3]. Получение прогноза позволит не только зафиксировать кибератаку на ранней стадии, но и отследить точку проникновения злоумышленника в систему, выполнить эффективное расследование инцидентов и спрогнозировать возможные вектора кибератак на систему.

Список литературы:

1. Deep Learning on Dynamic Graphs [Электронный ресурс]. URL: https://towardsdatascience.com/deep-learning-on-dynamic-graphs-1b97c2fab0ab.
2. Сошникова, Л. А. "Адаптивные методы краткосрочного прогнозирования: учебная программа учреждения высшего образования по учебной дисциплине для специальности 1-25 80 10 «Статистика и анализ»." (2019).
3. Lamberty, Jose M. Short Term Cyber Attacks with Long Term Effects and Degradation of Supply Chain Capability. Naval Postgraduate School Monterey United States, 2016.

Павленко Е. Ю.

Санкт-Петербургский политехнический университет Петра Великого (СПбПУ)

ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ СЛОЖНЫХ СИСТЕМ НА ОСНОВЕ ИСКУССТВЕННОЙ ИММУНИЗАЦИИ

** Исследование выполнено в рамках гранта Президента РФ для государственной поддержки молодых российских ученых – кандидатов наук МК-3861.2022.1.6.*

Цифровизация технологической инфраструктуры оказала существенное влияние на информационно-технические системы, интегрированные с бизнесом, государственными сервисами и промышленностью. Такие системы функционируют как единое целое, представляя собой конгломерат баз и хранилищ данных, систем обработки информации,

систем SCADA (Supervisory Control And Data Acquisition) и т.д., взаимодействующих с другими подобными системами. Примерами таких крупномасштабных информационно-технических систем являются системы управления аэропортами, производственные системы, системы электронного правительства и банковские системы. Такие объекты подвержены постоянным целевым атакам, поэтому существует необходимость в разработке подхода, который обеспечит отражение различного рода киберугроз [1].

Предложенный в работе подход базируется на биоинспирированной концепции иммунизации, состоящей в переносе биологических свойств сохранения жизни на концепцию информационной безопасности [2, 3]. Иммунизация выступает в качестве механизма саморегуляции систем, защищающего их от определенного спектра киберугроз и снижающего риск повреждения критических узлов от новых киберугроз.

Ключевая идея предлагаемого иммунного подхода состоит из двух этапов: защита критических узлов системы и формирование иммунитета узлов, где под иммунитетом следует понимать сложность (алгоритмическую, вычислительную, организационную) для злоумышленника при реализации атаки на данный узел.

Атака на систему рассматривается как определенная схема действий, не зависящая от того, что произошло в системе. Действия оцениваются с точки зрения опасности для критических узлов системы, а защита заключается в ограничении распространения этих действий по всей системе (подобно тому, как ограничение коммуникаций является защитной мерой при пандемии) путем создания и активации защитных действий, симметричных действиям злоумышленника.

Предложенный подход направлен на решение следующих задач:

1. Предотвращение распространения кибератак.

2. Усиление существующей защиты системы за счет внесения параметрических / структурных изменений или дополнительных свойств.

3. Нейтрализация атак, направленных на конкретные (критические) узлы системы.

4. Ликвидация последствий атак.

5. Апостериорное обновление защитных мер для критических узлов. Важно отметить, что критические узлы могут быть разными, и для их защиты следует использовать разные методы.

Экспериментальные исследования, посвященные оценке адекватности и работоспособности предложенного подхода, состояли в сравнении эффективности защитных мер для моделируемой системы (состоящей из 600 узлов, из которых 200 были критическими, 350 - некритическими и 50 - резервными) в трех случаях: без иммунизации, с иммунизацией некритических узлов и с иммунизацией критических узлов.

Результаты исследований показали, что наилучший результат с точки зрения кибербезопасности достигается при иммунизации критических узлов. Так, в режиме без иммунизации 100 узлов были заражены уже в первые 5 виртуальных тактов модулируемой кибератаки, а на 32 такте система вышла из строя. При иммунизации некритических узлов отметка в 100 зараженных узлов была достигнута за 15 виртуальных тактов, через 40 тактов система была почти полностью неработоспособна, а полное заражение системы произошло через 55 тактов. При иммунизации критических узлов система смогла функционировать до 55 тактов, и важно отметить, что не все 600 узлов были отключены.

Список литературы:

1. Bécue, A., Praça, I. and Gama, J., 2021. Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. Artificial Intelligence Review, 54(5), pp.3849-3886.
2. Anokhin, P. K. Philosophical aspects of the theory of a functional system. Soviet Studies in Philosophy. 1971, 10(3), pp. 269-276.
3. Anokhin, P. K. Systemogenesis as a general regulator of brain development. Progress in brain research. 1964, 9, pp. 54-86.

Черкинский К.С.
Positive Technologies, Москва

КОМПЛЕКСНЫЕ СИСТЕМЫ РАСШИРЕННОГО ОБНАРУЖЕНИЯ КИБЕРАТАК И РЕАГИРОВАНИЯ НА НИХ С ИСПОЛЬЗОВАНИЕМ АГЕНТСКИХ РЕШЕНИЙ

1. Основные векторы развития атак на корпоративную инфраструктуру. Важность покрытия конечных точек.

2. Автоматизация обнаружения и реагирования. Обеспечение автономной отработки специфических сценариев.

3. Разбор сценариев (DLL Hijacking, Driver Abuse, атаки с использованием фишинговых вложений).

Сторожик В.С.
Арктический и антарктический научно-исследовательский институт, Санкт-Петербург

ОСОБЕННОСТИ РЕАЛИЗАЦИИ МЕР ПО ОБЕСПЕЧЕНИЮ БЕЗОПАСНОСТИ ЗНАЧИМЫХ ОБЪЕКТОВ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

В Стратегии национальной безопасности Российской Федерации отмечено, что увеличивается количество компьютерных атак на российские информационные ресурсы. Большая часть таких атак осуществляется с территорий иностранных государств. Использование иностранных информационных технологий и телекоммуникационного оборудования повышает уязвимость объектов критической информационной инфраструктуры (КИИ) [1].

В Доктрине информационной безопасности Российской Федерации к основным национальным интересам в информационной сфере отнесено обеспечение устойчивого и бесперебойного функционирования КИИ в условиях проведения компьютерных атак [2].

В целях повышения устойчивости и безопасности функционирования информационных ресурсов Российской Федерации Указом Президента Российской Федерации от 1 мая 2022 г. № 250 руководителям субъектов КИИ предписано обеспечивать незамедлительную реализацию организационных и технических мер, решения о необходимости осуществления которых принимаются ФСБ России и ФСТЭК России и направляются на регулярной основе в органы (организации) - субъекты КИИ с учетом меняющихся угроз в информационной сфере [3, 4].

В рамках реализации Федерального закона от 26 июля 2017 г. № 187-ФЗ ФСТЭК России определена как федеральный орган исполнительной власти, уполномоченный в области обеспечения безопасности КИИ [5, 6].

Приказом ФСТЭК России от 25 декабря 2017 г. № 239 утверждены требования по обеспечению безопасности значимых объектов КИИ, действие которых распространяется на информационные системы (ИС), автоматизированные системы управления (АСУ) и информационно-телекоммуникационные сети (ИТКС), которые отнесены к значимым объектам КИИ [7, 8]. Требованиями предусмотрено, что разработка и внедрение организационных и технических мер по обеспечению безопасности значимого объекта КИИ производится субъектом КИИ на стадиях (этапах) жизненного цикла в ходе создания (модернизации), эксплуатации и вывода из эксплуатации значимого объекта КИИ. При этом организационные и технические меры по обеспечению безопасности не должны оказывать негативного влияния на создание и функционирование значимого объекта КИИ. При разработке организационных и технических мер по обеспечению безопасности значимого объекта КИИ учитывается его информационное взаимодействие с иными объектами КИИ, ИС, АСУ или ИТКС [7].

Для обеспечения безопасности значимых объектов КИИ, являющихся ИС персональных данных, субъектом КИИ дополнительно учитываются требования, утвержденные постановлением Правительства Российской Федерации от 1 ноября 2012 г. № 1119, а также состав и содержание организационных и технических мер, утвержденных приказом ФСТЭК России от 18 февраля 2013 г. № 21 [9, 10].

Для обеспечения безопасности значимых объектов КИИ, являющихся государственными ИС, субъектом КИИ дополнительно учитываются требования, утвержденные приказом ФСТЭК России от 11 февраля 2013 г. № 17 [11, 12].

Для обеспечения безопасности значимых объектов КИИ, являющихся АСУ производственными и технологическими процессами, дополнительно учитываются требования, утвержденные приказом ФСТЭК России от 14 марта 2014 г. № 31 [13].

В случае если значимый объект является государственной ИС или ИС персональных данных, меры по обеспечению безопасности значимого объекта и меры защиты информации (персональных данных) принимаются в соответствии с более высокой категорией значимости, классом защищенности или уровнем защищенности персональных данных [7].

Список литературы:

1. Стратегия национальной безопасности Российской Федерации (утв. Указом Президента Российской Федерации от 2 июля 2021 г. № 400).

2. Доктрина информационной безопасности Российской Федерации (утв. Указом Президента Российской Федерации от 5 декабря 2016 г. № 646).

3. О дополнительных мерах по обеспечению информационной безопасности Российской Федерации (утв. Указом Президента Российской Федерации от 1 мая 2022 г. № 250).

4. Информационное сообщение ФСТЭК России от 24 марта 2022 г. № 240/22/1549 «О мерах по повышению защищенности информационной инфраструктуры».

5. Федеральный закон от 26 июня 2017 г. N 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации».

6. Указ Президента Российской Федерации от 25 ноября 2017 г. № 569 «О внесении изменений в Положение о Федеральной службе по техническому и экспортному контролю, утвержденное Указом Президента Российской Федерации от 16 августа 2004 г. № 1085».

7. Приказ ФСТЭК России от 25 декабря 2017 г. № 239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации».

8. Постановление Правительства Российской Федерации от 8 февраля 2018 г. № 127 «Об утверждении Правил категорирования объектов критической информационной инфраструктуры Российской Федерации, а также перечня показателей критериев значимости объектов критической информационной инфраструктуры Российской Федерации и их значений».

9. Постановление Правительства Российской Федерации от 1 ноября 2012 г. № 1119 «Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных».

10. Приказ ФСТЭК России от 18 февраля 2013 г. № 21 «Об утверждении состава и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных».

11. Приказ ФСТЭК России от 11 февраля 2013 г. № 17 «Об утверждении требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах».

12. Приказ ФСТЭК России от 15 февраля 2017 г. № 27 «О внесении изменений в Требования о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах, утвержденные приказом ФСТЭК России от 11 февраля 2013 г. № 17».

13. Приказ ФСТЭК России от 14 марта 2014 г. № 31 «Об утверждении Требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды».

Волошина Н.В.[(1)], Беззатеев С.В.[(2)]
[(1)]Национальный исследовательский университет ИТМО,
[(2)]Санкт-Петербургский государственный университет аэрокосмического приборостроения
Санкт Петербург

АЛГОРИТМ ШИРОКОВЕЩАТЕЛЬНОГО ВСТРАИВАНИЯ ИНФОРМАЦИИ В КОНТЕЙНЕР С ВЗВЕШЕННОЙ СТРУКТУРОЙ

В современном мире постоянно растут потоки обмена мультимедиа данными, которые, в свою очередь необходимо защищать от незаконного использования. Одним из эффективных подходов по обеспечению защиты цифровых мультимедиа данных, является стеганографическая защита в виде цифровых водяных знаков. Основными направлениями научных исследований в данной области являются поиск методов и алгоритмов по увеличению объема встраивания, незаметности встраиваемой информации и устойчивости к атакам. В данной работе представлен подход по увеличению объема передаваемой информации при сохранении исходного уровня незаметности вносимых искажений при широковещательной передаче данных с обеспечением разделения прав доступа к передаваемой в ЦВЗ информации.

Предлагается алгоритм встраивания информации в контейнер с взвешенной структурой, позволяющий учесть взвешенную структуру контейнера и обеспечивающий широковещательный доступ [1, 3] к встраиваемой информации. Таким образом, из всего множества пользователей, имеющих персональные секретные ключи и имеющих доступ к заполненному контейнеру, корректную информацию получат лишь пользователи, выбранные перед выполнением процедуры вставки информации в контейнер. Алгоритм использует метод встраивания WF5, описанный ранее авторами в работах [2, 3] позволяющий оптимальным образом встраивать информацию в контейнер со взвешенной структурой. Оптимальность метода WF5 подразумевает встраивание информации в контейнер выбранного размера с минимально возможными искажениями. Однако, классический вариант этого метода позволяет встраивать информацию, доступную лишь одному конкретному пользователю, который извлекает ее из заполненного контейнера, используя свой секретный ключ, а именно – множество локаторов позиций L и многочлен Гоппы G(x) выбранного Г(L,G) кода, совершенного во взвешенной метрике Хэмминга.

Рассматриваемый алгоритм, обеспечивающий широковещательный доступ к встраиваемой информации, использует при встраивании информации набор ключей пользователей, которым предназначена встраиваемая информация. При этом основным достоинством предлагаемого алгоритма является то, что при встраивании обеспечивается минимальное искажение битового потока исходного контейнера.

Параметрами предлагаемого алгоритма являются общее число пользователей N, максимальный размер подмножества (коалиции) пользователей K, для которых может быть сформирована доступная им информация, встроенная в исходный контейнер, размер блока информации исходного контейнера n, в который будет встраиваться информация, число уровней t , используемых в методе WF5.

Список литературы:

1. Fiat Amos , Naor, Moni, Broadcast Encryption. CRYPTO 1993, pp. 480-491

2. Voloshina, N. , Bezzateev, S. , Prudanov, A. , Vasilev, M. , Gorbunov, A., Effectiveness of LSB and MLSB information embedding for BMP images, Conference of Open Innovation Association, FRUCT, 2016, pp. 378-384

3. Voloshina N., Bezzateev S., Zhidanov K., Weighted Digital Watermarking Approaches Comparison, Redundancy 2016, Saint Petersburg, 26-29 September, pp. 172-174, 2016

Хорев А.А

Национальный исследовательский университет «МИЭТ»

ОПЫТ ПРАКТИКО-ОРИЕНТИРОВАННОЙ ПОДГОТОВКИ СПЕЦИАЛИСТОВ ПО ТЕХНИЧЕСКОЙ ЗАЩИТЕ ИНФОРМАЦИИ В НИУ МИЭТ

Одним из направлений качества подготовки выпускников вузов в области защиты информации является реализация практико-ориентированного подхода, который основывается на реализации в рамках образовательной программы требований профессиональных стандартов в области информационной безопасности.

Профессиональный стандарт является новой формой определения квалификации работника по сравнению с единым тарифно-квалификационным справочником работ и профессий рабочих и единым квалификационным справочником должностей руководителей, специалистов и служащих.

Профессиональный стандарт содержит перечень знаний и умений, необходимых для реализации трудовых функций, выполняемых в рамках соответствующей квалификации.

Следовательно, образовательная организация, формируя профессиональные компетенции на основе трудовых функций, включенных в профессиональные стандарты соответствующих направлений профессиональной деятельности, должна включить в образовательную программу и индикаторы сформированности этих профессиональных компетенций (знания и умения).

Параллельно с разработкой профессиональных стандартов в России идет создание системы независимой оценки квалификации работников – процедуры подтверждения соответствия квалификации соискателя положениям профессионального стандарта.

Использование образовательной организацией для промежуточной аттестации оценочных средств, разработанных на основе оценочных средств, используемых при сдаче профессионального экзамена, позволит значительно повысить объективность оценки уровня подготовки выпускников различных вузов.

Одним из наиболее критичных показателей при организации сдачи профессионального экзамена, является наличие соответствующего материально-технического обеспечения оценочных мероприятий.

В России 130 вузов реализуют образовательные программы в области информационной безопасности и защиты информации. Однако, немногие из них имеют современную учебно-материальную базу, необходимую для подготовки специалистов в области технической защиты информации и, в частности, в области защиты информации от утечки по техническим каналам.

Рассмотрим представленную выше модель практико-ориентированной подготовки специалистов в области технической защиты информации на примере подготовки бакалавров и магистров в НИУ МИЭТ.

Перечень профессиональных компетенций, формируемых у выпускников в области технической защиты информации, приведен в табл. 1.

В бакалавриате формирование профессиональных компетенций происходит в ходе учебной (6 з.е.) и производственной (9 з.е.) практик, которые проводятся в 8-м семестре.

Таблица 1

Профессиональные компетенции, формируемые у выпускников НИУ МИЭТ, обучающихся по направлению «Информационная безопасность»

Направление направления подготовки	Код и наименование профессиональной компетенции выпускника	Трудовая функция из ПС, на основе которой сформулирована компетенция
10.03.01 Информационная безопасность (бакалавриат). Профиль «Техническая защита информации». Срок подготовки – 4 года	ПК-1. Способен проводить работы по установке, настройке и испытаниям защищенных технических средств обработки информации	B/6. Проведение работ по установке и техническому обслуживанию защищенных технических средств обработки информации.
	ПК-2. Способен проводить контроль эффективности защиты информации от утечки за счет побочных электромагнитных излучений и наводок	D/02.6. Проведение контроля защищенности информации от утечки за счет побочных электромагнитных излучений и наводок.

Направление направления подготовки	Код и наименование профессиональной компетенции выпускника	Трудовая функция из ПС, на основе которой сформулирована компетенция
	ПК-3. Способен проводить контроль эффективности защиты акустической речевой информации от утечки по техническим каналам	D/03.6. Проведение контроля защищенности акустической речевой информации от утечки по техническим каналам.
10.04.01 Информационная безопасность (магистратура. Программа «Аудит информационной безопасности» Срок подготовки – 2 года	ПК-1. Способен проводить аттестацию автоматизированных систем, средств обработки информации на соответствие требованиям безопасности информации	G/01.7. Проведение аттестации объектов вычислительной техники на соответствие требованиям по защите информации.
	ПК-2. Способен проводить аттестацию выделенных (защищаемых) помещений на соответствие требованиям безопасности информации	G/02.7. Проведение аттестации выделенных (защищаемых) помещений на соответствие требованиям по защите информации.

В магистратуре формирование профессиональных компетенций происходит при изучении дисциплин: технологии защиты информации от утечки по техническим каналам, технологии защиты информации от несанкционированного доступа, контроль защищенности информации от утечки по техническим каналам, а таже в ходе производственной практики. Производственная практика проходит в ходе 1-го (10 з.е.) и 2-го (12 з.е.) семестров.

Производственная практика проходит на базе проходит на базе НТЦ ТЗИ кафедры «Информационная безопасность» и ведущих предприятий г. Зеленограда и г. Москвы.

Заканчивается производственная практика выполнением практико-ориентированных заданий, разработанных на основе оценочных средств, используемых при сдаче профессиональных экзаменов для оценки соответствующих профессиональных квалификаций.

С целью формирования профессиональных компетенций на кафедре «Информационная безопасность» созданы четыре специализированные учебные лаборатории, оснащенные современной зарубежной и отечественной техникой и технологиями, стоимостью более 60 000 000 рублей.

2. Кибербезопасность и искусственный интеллект

Коваленко А.П.
Московский институт новых информационных технологий

ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ МНОГОСЛОЙНОГО ПЕРЦЕПТРОНА С КУСОЧНО-ЛИНЕЙНЫМИ ФУНКЦИЯМИ АКТИВАЦИИ

Аналитическое агентство Gartner включило разработку интерпретируемого искусственного интеллекта (eXplainable AI, XAI) в первую десятку наиболее актуальных задач в области анализа больших данных [1]. Ожидается, что набор методов интерпретации обученной модели ИИ позволит объяснить ее поведение в каждом конкретном случае, выявить потенциальные ошибки, повысить прозрачность и надежность предлагаемых решений. Особенно это важно в областях с высокой стоимостью ошибки (управление беспилотными транспортными средствами, медицинская диагностика и т.п.). Недостаточное понимание пользователями процесса принятия решений моделями ИИ, например, нейросетями, которые обычно сравнивают с «черным ящиком», представляет собой серьезное препятствие их внедрения в бизнес и повседневную жизнь [2].

Ряд моделей ИИ (нейронные сети, машины опорных векторов, ансамбли решающих деревьев) считаются «плохо» интерпретируемыми (в отличие от «хорошо» интерпретируемых деревьев решений и линейной регрессии), хотя все происходящие внутри них вычисления известны. Здесь имеется в виду то, что процесс принятия решения не удается представить в «прозрачной» форме:

1. Показать, какие признаки входных данных существенно влияют, или наоборот, не влияют на решение.

2. Представить алгоритм принятия решения в виде понятных шагов.

3. Объяснить смысл промежуточных результатов вычислений.

4. Представить результаты обучения модели ИИ в удобной для восприятия форме (с привлечением средств визуализации).

Обзор методов интерпретации моделей ИИ можно найти, например, в [3].

Однако в некоторых частных случаях «плохо» интерпретируемая модель может быть преобразована в эквивалентную ей «прозрачную» модель без существенных затрат вычислительных и временных ресурсов.

В докладе представлен новый метод геометрической интерпретации результатов обучения многослойного перцептрона, основанный на преобразовании полносвязной многослойной сети с широко применяемыми кусочно-линейными функциями активации

нейронов скрытых слоев (типа ReLU, LeakyReLU, ABS) в объясняющее двоичное дерево (eXplanatory Binary Tree, eXBTree).

Показано, что временная и пространственная сложность алгоритма построения eXBTree по обученной нейронной сети такого типа составляет O(ndK), где n – объем обучающей выборки, d – размерность входного пространства, K – общее число нейронов во внутренних слоях сети.

Основываясь на результатах работы [4], сформулированы асимптотические условия, при которых двоичный классификатор на основе eXBTree является строго состоятельным при увеличении объема обучающей выборки, то есть выборочная ошибка eXBTree-классификатора в пределе с вероятностью 1 совпадет с байесовской ошибкой.

В основе полученных результатов лежит тот факт, что многослойный перцептрон с кусочно-линейными функциями активации строит иерархическое (по слоям нейросети) рассечение компакта (например, гиперпараллепипеда) исходного d-мерного пространства, в котором решается задача нейросетевой аппроксимации функции, на выпуклые политопы (ячейки), формируя тем самым гистограмму, зависящую от данных.

Такая связь многослойного перцептрона с хорошо известным классификатором на основе гистограммной оценки плотности делает алгоритм нейросетевой классификации «прозрачным» как с содержательной, так и со статистической точек зрения. Однако, сразу возникает понятная для гистограммы, но «непрозрачная» для перцептрона проблема «проклятия размерности». В [4] показано, что число ячеек, на которое первый скрытый слой перцептрона рассекает гиперпараллепипед, пропорционально числу нейронов первого слоя в степени размерности пространства. Поскольку в современных нейросетевых моделях ИИ размерность пространства измеряется сотнями, а число нейронов в скрытых слоях достигает тысяч, получаемая «мегагистограмма» не сопоставима ни с каким объемом выборки! Поэтому ни о какой содержательной статистической постановке задачи нейросетевой классификации в общем случае речь идти не может.

Этот вывод означает, что нейросетевой классификатор решает задачу не статистической, а структурной классификации, особенности которой предполагается обсудить, с учетом сформулированных выше признаков «прозрачности», в завершение доклада.

Список литературы:

1. https://www.gartner.com/smarterwithgartner/gartner-top-10-data-analytics-trends/

2. https://rb.ru/opinion/uzhe-ne-black-box/

3. Li et al., Interpretable Deep Learning: Interpretation, Interpretability, Trustworthiness, and Beyond, 2021, arXiv:2103.10689.

4. Devroye, L et al., A Probabilistic Theory of Pattern Recognition, Springer, 1996

Беззатеев С.В., Елина Т.Н., Красников Н.С.
Санкт-Петербургский государственный университет аэрокосмического Приборостроения.

ВОПРОСЫ ОРГАНИЗАЦИИ КОМПЛЕКСНОЙ ЗАЩИТЫ СИСТЕМ МАШИННОГО ЗРЕНИЯ

Популярность нейронных сетей как средств реализации технологий «умного дома» в современном мире приводит к росту их уязвимости со стороны реализации угроз информационной безопасности. Разработка методов защиты нейронных сетей становится необходимым условием их надежного и безопасного функционирования. В первую очередь, структура системы комплексной защиты информации зависит от компонентов и модулей самого объекта защиты, их взаимного расположения и характера связей между ними.

Архитектура системы распознавания весовых товаров на кассе самообслуживания розничного магазина, основанной на применении сверточных нейронных сетей, представлена на рисунке 1.

Рисунок 1 – Архитектура информационной системы

Облачное хранилище содержит модель нейронной сети и набор данных. Выделенный сервер хранит базу данных доступа к хранилищу, моделей и изображений, модель нейронной сети и наборы данных для ее дообучения. Устройство сканирования расположено непосредственно на кассе и содержит обученную модель распознавания товаров. В соответствии с предложенной архитектурой информационной системы можно выделить перечень активов, подлежащих информационной защите:

- датасет с изображениями;
- нейронная сеть;
- облачное хранилище;
- нереляционная база данных.

В рамках разработки системы защиты предлагаемой информационной системы выделены и актуализированы следующие потенциальные угрозы обеспечения безопасности с отнесением их к конкретному информационному активу:

- УБИ.132: Угроза получения предварительной информации об объекте защиты (модель нейронной сети);

- УБИ.149: Угроза сбоя обработки специальным образом изменённых файлов (модель нейронной сети);
- УБИ.168: Угроза «кражи» учётной записи доступа к сетевым сервисам (модель нейронной сети, облачное хранилище, набор данных);
- УБИ.219: Угроза хищения обучающих данных (модель нейронной сети, набор данных);
- УБИ.220: Угроза нарушения функционирования («обхода») средств, реализующих технологии искусственного интеллекта (модель нейронной сети);
- УБИ.221: Угроза модификации модели машинного обучения путем искажения («отравления») обучающих данных (модель нейронной сети, набор данных);
- УБИ.222: Угроза подмены модели машинного обучения (модель нейронной сети).

На основе выделенных актуальных угроз безопасности информации определены следующие цели безопасности и удовлетворяющие им политики.

Цели безопасности:

1. Реализация мандатной модели разграничения доступа. (УБИ.132, УБИ.219, УБИ.221, УБИ.222).
2. Защита модели нейронной сети от попадания в неё некорректных данных и исследования кода программы (УБИ.149, УБИ.220).
3. Шифрование имен пользователей, паролей и данных приложений (УБИ.149, УБИ.168, УБИ.221, УБИ.222).
4. Обеспечение защиты от доступа к данным третьих лиц (УБИ.132, УБИ.222).

Политики безопасности:

1. Политика разграничения доступа. Доступ к различным активам должен быть строго разграничен в соответствии с привилегиями должности.
2. Политика работы с поступающей информацией. К обработке нейронной сетью допускаются только разрешенные данные, прошедшие предварительную обработку и проверку на наличие дефектов. Так сводится к минимуму возможность организации помех в корректной работе нейронной сети.
3. Политика создания рабочего аккаунта. Для получения доступа к активам необходим аккаунт, зарегистрированный на активно использующийся действительный адрес электронной почты. Пароль для рабочего аккаунта должен удовлетворять минимальным требованиям, установленным в технической политике компании.

Так как описываемая система является многопользовательской с разным уровнем доступа, при этом в ней не ведётся работа с секретными данными для нее был определен класс 1Г и выбран межсетевой экран 4-ого класса защиты. Профиль защиты такого межсетевого экрана приведен в документе ФСТЭК [1], а функциональные требования к классу защищенности 1Г указаны в руководящем документе [2].

Список литературы:

1. ФСТЭК. Профиль защиты межсетевых экранов типа «Г» четвертого класса защиты ИТ.МЭ.Г4.ПЗ.: методический документ, 2016.
2. Автоматизированные системы. Защита от несанкционированного доступа к информации. Классификация автоматизированных систем и требования по защите информации.: руководящий документ, 1992.

Крундышев В.М.
Санкт-Петербургский Политехнический университет Петра Великого

СИСТЕМА УПРАВЛЕНИЯ ОБНАРУЖЕНИЕМ КОМПЬЮТЕРНЫХ АТАК НА БАЗЕ НЕЙРО-НЕЧЕТКОЙ ЛОГИКИ В КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЕ

Работа выполнена в рамках Государственного задания на проведение фундаментальных исследований (код темы 0784-2020-0026)

Новые разновидности компьютерных атак, в т.ч. полиморфные целевые атаки, атаки на динамическую маршрутизацию, туннельные атаки, сложно определяются либо вовсе не определяются при помощи традиционных методов обнаружения и могут приводить не только к нарушениям конфиденциальности, целостности и доступности информационных ресурсов, но и к нарушениям непрерывного функционирования объектов критической информационной инфраструктуры (КИИ) [1, 2]. Критическим становится противоречие между вариативностью действующих атак, скоростью их реализации и инертностью механизмов обнаружения, что ослабляет защищенность значимых объектов КИИ, используемых в банковской сфере, в системах транспорта, связи, энергетики, атомной и химической промышленности, а также усиливает социальные и техногенные риски [3].

Для решения проблем, связанных с недостатком информации, необходимой для получения количественного описания протекающих в системе процессов, так и сложностью объекта управления, предлагается использовать аппарат нечеткой логики. Теория нечетких множеств позволяет строить нечеткие модели объектов с использованием лингвистических переменных и механизмов логического вывода [4]. В настоящее время модели нечеткого логического вывода используются при разработке нечетких экспертных систем, применяемых для решения задач диагностики, управления, поддержки принятия решений в различных предметных областях.

В данной работе представлена система управления обнаружением компьютерных атак на объекты КИИ, реализующая выбор детекторов атак в режиме реального времени, за счет комбинирования искусственных нейросетей и аппарата нечеткой логики. Также определена структура базы знаний, и предложен комплекс новых методов обнаружения компьютерных атак, покрывающих множество актуальных киберугроз объектам КИИ.

Для обеспечения гибкости при формировании базы знаний предлагается использовать модифицированные нечеткие продукционные правила, формулируемые в виде нечетких высказываний относительно значений тех или иных лингвистических переменных. В рамках решаемой задачи в качестве анализируемых данных предлагается набор характеристик объекта защиты: сетевые параметры, доступные вычислительные ресурсы для работы системы анализа киберугроз, экономические показатели, допустимое время реакции на компьютерную атаку, а также текущий уровень киберугроз. Собранные данные необходимо предобработать и преобразовать в базу методов обнаружения компьютерных атак следующей структуры (таблица 1). Для каждого из признаков определено терм-множество вида {Н (низкий), С (средний), В (высокий)}, а также задана функция принадлежности, ставящая в соответствие значению терм.

Таблица 1 – Структура базы методов обнаружения компьютерных атак

Входные переменные						**Выходная переменная**
Сетевые характеристики	*Вычислительные ресурсы*	*Экономические параметры*	*Время реакции*	*Уровень киберугроз*	*Параметры типов атак*	*Методы обнаружения атак*

Кол-во узлов	Скорость передачи данных	…	ЦП	ОП	Диск	Стоимость активов	…	Временные задержки	DoS-атака	Черная дыра	…	Активное воздействие	…	Нейросетевой метод
Н	Н	…	Н	Н	Н	Н	…	Н	Н	Н	…	Н	…	Метод машинного обучения
С	С	…	С	С	С	С	…	С	С	С	…	С	…	Роевой интеллект
В	В	…	В	В	В	В	…	В	В	В	…	В	…	Метод выравнивания

Предложенная структура поддерживает расширение характеристик объекта, обнаруживаемых компьютерных атак и детекторов. Добавление новой информации в базу знаний может осуществляться вручную оператором системы управления или за счет самообучения системы (приспособления к новым условиям или задачам).

Для принятия решений в условиях неполноты знаний об объекте защиты и актуальных киберугрозах реализована нейро-нечеткая модель, основанная на адаптивной системе нейро-нечеткого вывода ANFIS на базе алгоритма Такаги-Сугено-Канга, основным преимуществом которого является высокая производительность и точность. Данная нечеткая адаптивная сеть базируется на следующих положениях: входные переменные являются четкими, функции принадлежности всех перечисленных множеств определены функцией Гаусса.

Список литературы:

4. Cho H., Lim S., Belenko V., Kalinin M., Zegzhda D., Nuralieva E. Application and improvement of sequence alignment algorithms for intrusion detection in the Internet of Things. In 2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS). – 2020. Vol. 1. pp. 93-97.
5. Lim S., Kalinin M., Zegzhda P. Bioinspired Intrusion Detection in ITC Infrastructures. In International Conference on Technological Transformation: A New Role for Human, Machines and Management. Springer, Cham. – 2020. pp. 10-22.
6. Овасапян Т. Д. Применение нейронных сетей для выявления внутренних нарушителей в VANET-сетях / Т. Д. Овасапян, Д. А. Москвин, М. О. Калинин // Проблемы информационной безопасности. Компьютерные системы. – 2018. – № 1. – С. 68-73.
7. Катасев А. С. Методы и алгоритмы формирования нечетких моделей оценки состояния объектов в условиях неопределенности / А. С. Катасев // Вестник Технологического университета. – 2019. – Т. 22. – № 3. – С. 138-147.

Овасапян Т.Д., Москвин Д.А.
Санкт-Петербургский политехнический университет Петра Великого

АДАПТИВНАЯ СИСТЕМА УПРАВЛЕНИЯ НА БАЗЕ ОБУЧАЮЩЕГОСЯ АВТОМАТА ДЛЯ WSN-СЕТЕЙ

Начало текущего века характеризуется стремительным развитием крупномасштабной технологической инфраструктуры, в том числе повсеместным внедрением беспроводных сенсорных сетей, представляющих собой первые самоорганизующиеся сети [1], которые начали использоваться в сетях связи общего пользования во многих странах мира, включая и Российскую Федерацию.

Проблема подверженности беспроводных сенсорных сетей (БСС) атакующим воздействиям приобретает все более существенное значение, как из-за возрастающего

проникновения БСС в различные сферы жизни, так и в результате повышения структурно функциональной сложности таких сетей и предоставляемых ими сервисов [2, 3], что в свою очередь повышает риски нелегитимного использования БСС нарушителем информационной безопасности. Современные средства анализа и обеспечения безопасности не способны осуществлять комплексную защиту, поскольку они функционируют на одном-двух уровнях системы и не учитывают специфику угроз, характерных для каждого уровня в контексте беспроводных сенсорных сетей [4].

В рамках исследования предлагается система адаптивного управления работой БСС, использующая сочетание методов искусственного интеллекта и математического моделирования, а также позволяющая поддерживать защищенность и функциональную устойчивость беспроводной сенсорной сети. Сохранение функциональной устойчивости достигается адаптивным поведением на базе обучающегося автомата [5], благодаря которому узел изменяет правила взаимодействия со своими соседями. Работоспособность сети в условиях угроз безопасности и деструктивных воздействий обеспечивается за счет изменения поведения узла относительно вредоносного или неисправного узла на основе интеллектуального анализа входных сигналов. Например, узлу, который некорректно обрабатывает или не ретранслирует входящие данные, другие узлы пакеты отправляться не будут, при этом данные от датчиков узла все равно будут приниматься.

Сигналы, которые будут влиять на адаптивную работу предлагается разделить на две группы:

— показатели поведения (ретрансляция пакетов по сети, сохранение целостности пакетов, генерирование корректных данных и др.);

— показатели функционирования (оставшийся заряд батареи, память, нагрузка на микроконтроллер и др.).

Для экспериментальной оценки эффективности разработанной системы выполнен сравнительный анализ существующих средств имитационного моделирования работы распределенных сетей, внесены изменения в существующее средство для более корректного учета энергопотребления узлов сети при симулировании их работы. Проведена экспериментальная оценка, в рамках которой показано, что процент потерь пакетов не превышает 20% с учетом совершения атак. Время работы сети дольше на 30% по сравнению с сетью без системы защиты при совершении атак, направленных на исчерпание ресурсов узлов.

Список литературы:

1. Yick J., Mukherjee B., Ghosal D. Wireless sensor network survey //Computer networks. – 2008. – Т. 52. – №. 12. – С. 2292-2330.

2. Ognev R. A., Zhukovskii E. V., Zegzhda D. P. Clustering of malicious executable files based on the sequence analysis of system calls //Automatic Control and Computer Sciences. – 2019. – Т. 53. – №. 8. – С. 1045-1055.

3. Крундышев В., Калинин М. О., Зегжда Д. П. Моделирование и исследование свойств безопасности перемещающихся программно-конфигурируемых сетей VANET/MANET с использованием виртуальной среды суперкомпьютера //Информационная безопасность

регионов России (ИБРР-2017). И 74 Юбилейная X Санкт-Петербургская межрегиональная конференция. Санкт-Петербург, 1-3 ноября 2017 г.: Материалы конференции/СПОИСУ.–СПб., 2017.–580 с. – 2017. – С. 280.

4. Овасапян Т.Д. Обеспечение безопасности WSN-сетей на основе модели доверия / Овасапян Т.Д., Иванов Д.В. // Журнал «Проблемы информационной безопасности. Компьютерные системы». – СПб.: Изд-во Политехн. ун-та, 2017. - №4. - С. 64-72.

5. Цетлин М. Л. Конечные автоматы и моделирование простейших форм поведения //Успехи математических наук. – 1963. – Т. 18. – №. 4 112. – С. 3-28.

Гололобов Н. В., Зегжда Д. П. Павленко Е. Ю.
Санкт-Петербургский политехнический университет Петра Великого

РАСПОЗНАВАНИЕ КИБЕРУГРОЗ НА АДАПТИВНУЮ СЕТЕВУЮ ТОПОЛОГИЮ КРУПНОМАСШТАБНЫХ СИСТЕМ НА ОСНОВЕ РЕКУРРЕНТНОЙ НЕЙРОГЕНЕТИЧЕСКОЙ СЕТИ С ДОЛГОЙ КРАТКОСРОЧНОЙ ПАМЯТЬЮ

** Исследование выполнено за счет гранта Российского научного фонда № 22-21-20008, https://rscf.ru/project/22-21-20008/*

В современном мире широко применяются крупномасштабные системы для управления техническими процессами на производствах. Количество логически узлов в таких системах может достигать десятков тысяч, что создает необходимость их обеспечения адаптивной сетевой топологией для исключения каскадных нарушений функциональности в случае проведения атаки. Использование адаптивной сетевой топологии позволяет при выходе узла из строя оперативно обеспечить схождение сети таким образом, что уровень функциональности системы не изменится или несущественно снизится.

Обеспечение кибербезопасности адаптивных сетевых топологий является основной задачей при проектировании киберустойчивых крупномасштабных систем [1]. Использование всех доступных методов защиты от атак на адаптивную топологию зачастую приводит к снижению уровня удобства работы с системой, что, в свою очередь, неминуемо влечет финансовые издержки. Таким образом, приоритетным вопросом является разработка таких методов и способов, которые позволили бы реагировать на киберугрозы на ранних стадиях проведения атаки.

Наиболее перспективным направлением в создании автоматизированных средств превентивного реагирования является использование технологий машинного обучения. В частности, для таких целей применяется прогнозирование на основе полученных ранее показателей. Для прогнозирования используются нейрогенетические сети, позволяющие не только в значительной степени автоматизировать анализ данных, но также сократить вероятность антропогенной ошибки при проведении расчётов и формировании прогнозов.

Одной из современных тенденций развития технологии машинного обучения является сокращение времени обработки данных и повышение точности при минимизации мощностных затрат. В данном направлении широкое применение получили

нейрогенетические сети с долгой краткосрочной памятью, отличительной особенностью которых является сохранение результатов предыдущих вычислений – накопление опыта [2]. Кроме того, данные вид сетей позволяет отказаться от обратного распространения ошибки до первого слоя без значительных потерь в производительности.

В общем случае сетевая топология представляется в виде графа. Адаптивная составляющая добавляет возможность изменения топологии в момент времени T+1 вследствие нарушения работы одного узла или группы узлов, потери связи пограничных узлов с доменами и др. Для оценки состояния сетевой топологии в крупномасштабных системах могут использоваться телеметрические данные, которые могут быть использованы в качестве наборов данных для нейрогенетической сети.

Задача выявления киберугроз схожа с задачей выявления аномальных экземпляров в наборе данных [3]. Использование нейрогенетической сети с долгой краткосрочной памятью обусловлено особенностью данной архитектуры – повышенным уровнем запоминания контекста, что позволяет обнаруживать все типы аномалий [4]. В результате, выявленные коллективные, контекстуальные и точечные аномалии в показателях телеметрических данных могут сигнализировать об атаке на сеть и инициировать процедуры восстановления функциональности – адаптацию топологии с изоляцией атакуемых узлов.

Для определения киберугрозы посредством предлагаемого метода достаточно, чтобы предсказанное значение отличалось от фактического на некоторую величину, определяемую в ходе функционирования алгоритма. Таким образом, новизна заключается в использовании метода выявления киберугроз на основе машинного обучения, для которого отсутствует необходимость точного или близкого совпадения предсказанного и фактического значения экземпляра.

Полученные результаты могут быть использованы при обеспечении киберустойчивости крупномасштабных систем с адаптивной сетевой топологией. Кроме этого, результаты исследования позволят определить недостатки нейрогенетической сети с долгой краткосрочной памятью в задачах выявления аномалий для их последующего компенсирования.

Список литературы:

1. Adaptive Current Protection Technology for Distribution Network with Distributed Power Sources Based on Local Information / S. Cui, P. Zeng, Z. Wang, Y. Zuo // Mobile Information Systems. – 2021. – Vol. 2021. – P. 5137749. – DOI 10.1155/2021/5137749. – EDN ANBGLA.

2. Smagulova, K. A survey on LSTM memristive neural network architectures and applications / K. Smagulova, A. P. James // The European Physical Journal. Special Topics. – 2019. – Vol. 228. – No 10. – P. 2313-2324. – DOI 10.1140/epjst/e2019-900046-x. – EDN HRKIKB.

3. Gupta, S. Cyber security threat intelligence using data mining techniques and artificial intelligence / S. Gupta, A. S. Sabitha, R. Punhani // International Journal of Recent Technology and Engineering. – 2019. – Vol. 8. – No 3. – P. 6133-6140. – DOI 10.35940/ijrte.C5675.098319. – EDN QHKKNV.

4. Snorovikhina, V. Unsupervised Anomaly Detection for Discrete Sequence Healthcare Data / V. Snorovikhina, A. Zaytsev // Lecture Notes in Computer Science. – 2021. – Vol. 12602 LNCS. – P. 391-403. – DOI 10.1007/978-3-030-72610-2_30. – EDN FYZEWN.

Ломако А.Г., Менисов А.Б.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург

МОДЕЛЬ ЗЛОУМЫШЛЕННИКА ДЛЯ СУЩЕСТВУЮЩИХ ПРИКЛАДНЫХ СИСТЕМ, ИСПОЛЬЗУЮЩИХ ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Исследование выполнено в рамках гранта Президента РФ для государственной поддержки молодых российских ученых – кандидатов наук(МК-2485.2022.4).

Существуют различные группы нарушителей, которые могут реализовать угрозы безопасности информации системискусственного интеллекта (ИИ). Нарушители склонны эксплуатировать уязвимости в существующих и разрабатываемых системах ИИ.

Внутренние нарушители, в том числе сотрудники (разработчики программных, программно-аппаратных средств, лица, привлекаемые для установки, настройки, испытаний, пусконаладочных и иных видов работ, обеспечивающие функционирование систем и сетей или обеспечивающие системы оператора, авторизованные пользователи систем и сетей, системные администраторы и администраторы безопасности) и подрядчики (поставщики вычислительных услуг, услуг связи), имеющие доступ к объектам информационной инфраструктуры организации, могут украсть или повредить набор данных, используемый системами ИИ.

Специальные службы иностранных государств, как правило, продвинуты в области ИИ. Помимо разработки способов использования систем ИИ для защиты своих собственных объектов, они активно ищут уязвимости в системах ИИ, которые они могут использовать для атак, в том числе с применением технологий ИИ.

Другой вид нарушителей – это террористы, которые стремятся совершать террористические акты для нанесения ущерба отдельным сферам деятельности или секторам экономики государства. Террористы могут совершать атаки на системы ИИ, например, взламывать беспилотные автомобили, чтобы использовать их в качестве оружия.

Отдельные физические лица (хакеры), которые в основном руководствуются мотивом получения финансовой выгоды. Также они могут использовать ИИ, чтобы добиться известности.

Существуют также преступные группы, которые используют заранее написанные сценарии или программы для атаки на системы ИИ.

Помимо традиционных нарушителей, о которых говорилось выше, становится все более необходимым включать конкурентов в качестве субъектов угроз, поскольку некоторые компании все чаще демонстрируют намерение атаковать своих конкурентов, чтобы получить долю рынка.

Возможные негативные последствия от реализации (возникновения) угроз безопасности информации систем ИИ:

— компьютерный инцидент: факт нарушения и (или) прекращения функционирования информационного ресурса и (или) нарушения безопасности, обрабатываемой таким информационным ресурсом информации, в том числе произошедший в результате компьютерной атаки;

— несанкционированный доступ в систему (как категория компьютерного инцидента): факт доступа к информации или к информационному ресурсу, осуществляемого с нарушением установленных прав и/или правил доступа к информации или к информационному ресурсу с использованием штатных средств, предоставляемых средствами вычислительной техники, в том числе произошедший в результате компьютерной атаки.

— непреднамеренное (без злого умысла) отключение информационного ресурса (как тип компьютерного инцидента): факт нарушения работоспособности информационного ресурса, произошедшего в результате непреднамеренных (без злого умысла) действий или обстоятельств.

— нарушение или замедление работы информационного ресурса (как категория компьютерного инцидента): факт приведения информационного ресурса в состояние, при котором он не способен выполнять возложенную на него функцию должным образом, в том числе произошедший в результате компьютерной атаки.

— юридический действия: факт действия третьих лиц (по договору или иным образом) с целью запрещения действий или компенсации убытков на основании применимого законодательства.

Некоторые события безопасности ИИ могут проявляться как в виде непреднамеренного ущерба, так и в виде злонамеренных действий. Следует отметить, что перечисленные события относятся к системам ИИ, а компьютерные инциденты для других элементов экосистемы вынесены в ограничения.

Определение сценариев предусматривает установление последовательности возможных тактик и соответствующих им техник, а также доступности интерфейсов для использования соответствующих способов реализации угроз безопасности информации. В большинстве сценариев злоумышленники демонстрируют многоэтапную реализацию угроз ИИ, которую можно разделить на следующие этапы: сбор информации, получение первоначального доступа, внедрение и использование вредоносного кода, закрепление в системе и сети, управление вредоносным кодом и компонентом, повышение привилегий, сокрытие действий, получение доступа к другим компонентам, сбор и вывод информации и неправомерный доступ или воздействие.

Для описания способов реализации угроз наиболее хорошо зарекомендовали себя следующие подходы к представлению угроз безопасности информации:

— деревья атак [1], представленные в 1999 году Брюсом Шнайером;

— цепочки KillChain [2], разработанная в 2011 году компанией LockheedMartin для обнаружения нарушителей на протяжении всего жизненного цикла компьютерной атаки,

— набор тактик и техник поведения нарушителей MITREATT&CK [3];

— модель жизненного цикла компьютерной атаки [4], в которой особое внимание уделяется моделированию APT-атак, демонстрируя повторяющийся характер нарушителей для дальнейшего повышения привилегий.

В данном исследовании введено ограничение, которое заключается в том, что нарушитель демонстрирует различные возможности для выполнения своей вредоносной деятельности против систем ИИ в виде многоэтапного вторжения в несколько объектов инфраструктуры.

Существует достаточное количество исследований по определению возникновения угроз безопасности информации с помощью традиционных средств [5-8]. Стоит отметить, что ведущие организации в области информационной безопасности [9] отмечают растущее использование возможностей ИИ в текущем ландшафте угроз:

— расширение существующих угроз, которое связано с сложными компьютерными атаками на большое количество потенциальных целей и низкой стоимостью атак;

— введение новых угроз, связанных с задачами, которые были бы невыполнимы для человека;

— изменение типичного характера угроз, которое включает в себя новые атрибуты автоматизированных, высокоэффективных, трудно определяемых и крупномасштабных атак в ландшафте угроз.

Список литературы:

1. Schneier B. Attack trees //Dr. Dobb's journal. – 1999. – Т. 24. – №. 12. – С. 21-29.

2. Hutchins E. M. et al. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains //Leading Issues in Information Warfare & Security Research. – 2011. – Т. 1. – №. 1. – С. 80.

3. MITRE. 2017. ATT&CK Matrix for Enterprise. ». [Электронный ресурс]: https://attack.mitre.org/matrices/enterprise/ (дата обращения 01.02.2022).

4. Bu Z. Zero-day attacks are not the same as zero-day vulnerabilities //Fireye Blog. – 2014. – Т. 24.

5. Саенко И. Б. и др. Модель угроз ресурсам ИТКС как ключевому активу критически важного объекта инфраструктуры //Электросвязь. – 2021. – №. 1. – С. 36-44.

6. Рыбкина О. В. Построение модели угроз безопасности информации на основе математической модели Ланкастера //Научно-техническое и экономическое сотрудничество стран АТР в XXI веке. – 2021. – Т. 1. – С. 252-257.

7. Бражук А. И. Методика моделирования угроз компьютерных систем на основе предметно-ориентированных моделей //Информационные технологии в моделировании и управлении: подходы, методы, решения. – 2021. – С. 94-101.

8. Суханов И. Д., Рыбкина О. В. Новые подходы к моделированию угроз безопасности информации //Научно-техническое и экономическое сотрудничество стран АТР в XXI веке. – 2021. – Т. 1. – С. 277-282.

9. Caldwell M. et al. AI-enabled future crime //Crime Science. – 2020. – Т. 9. – №. 1. – С. 1-13.

Данилов В.Д., Овасапян Т.Д.
Санкт-Петербургский политехнический университет Петра Великого

АНАЛИЗ МЕТОДОВ ГЕНЕРИРОВАНИЯ СИНТЕТИЧЕСКИХ ДАННЫХ В КОНТЕКСТЕ СОЗДАНИЯ HONEYPOT-СИСТЕМ

Актуальность исследования и создания методов, осуществляющих анализ сетевых атак и противодействие им, трудно переоценить. Наряду с этим наблюдается бурный рост сетевых атак. Так по данным компании Check Point Software Technologies за 2021 год количество кибератак увеличилось на 40% по сравнению с 2020 годом [1]. Основной мотив атак – получение данных. Наибольшую угрозу представляют сетевые атаки, основанные на неизвестных уязвимостях нулевого дня (0-day) [2]. Это свидетельствует о необходимости создания методов для анализа и противодействия сетевым атакам.

Одним из методов анализа действий злоумышленников при эксплуатации неизвестных уязвимостей является использование honeypot-систем, которые позволяют изучить поведение атакующих и в дальнейшем предугадать наиболее потенциальные сценарии для атак [3]. Несмотря на успешное использование данных систем существует актуальная проблема, когда злоумышленник при атаке понимает, что система ненастоящая. Поскольку использование реальных данных для заполнения honeypot-систем небезопасно, в рамках доклада исследуется возможность генерирования синтетических данных с помощью методов глубокого обучения для последующего применения в honeypot-системах.

В качестве генерируемых типов данных в докладе рассматриваются наиболее актуальные целевые объекты сетевых атак в контексте honeypot-систем:

— персональные данные;

— медицинская информация;

— сетевой трафик автоматизированной системы управления технологическим процессом (АСУ ТП).

Основными задачами сгенерированных данных в контексте honeypot-систем являются привлечение потенциальных нарушителей для совершения нелегитимных действий, а также невозможность раскрытия исходных данных, из которых были получены синтетические. Для оценки качества синтетических данных с точки зрения решения этих задач в докладе анализируются возможные атаки, направленные на получение реальных данных, а также исследуются метрики, помогающие определить схожесть поддельных данных и реальных по их свойствам [4].

В качестве результата в докладе рассматривается разработанный макет автоматизированной системы генерирования синтетических данных для honeypot-систем и производится оценка эффективности его работы согласно раннее введенным показателям качества.

Список литературы:

1. Check Point Research: Cyber Attacks Increased 50% Year over Year [Электронный ресурс]. – URL: https://blog.checkpoint.com/2022/01/10/check-point-research-cyber-attacks-increased-50-year-over-year/. – (дата обращения 14.04.2022)

2. Маршев И. И., Жуковский Е. В., Александрова Е. Б. Использование ассемблерного кода программ для защиты средств обнаружения вредоносных программного обеспечения от состязательных атак //Методы и технические средства обеспечения безопасности информации. – 2020. – №. 29. – С. 85-86.

3. Овасапян Т. Д., Никулкин В. А., Москвин Д. А. ПРИМЕНЕНИЕ ТЕХНОЛОГИИ HONEYPOT С АДАПТИВНЫМ ПОВЕДЕНИЕМ ДЛЯ СЕТЕЙ ИНТЕРНЕТА ВЕЩЕЙ //Проблемы информационной безопасности. Компьютерные системы. – 2021. – №. 2. – С. 135-144.

4. Hayes J. et al. Logan: Membership inference attacks against generative models //Proceedings on Privacy Enhancing Technologies (PoPETs). – De Gruyter, 2019. – Т. 2019. – №. 1. – С. 133-152 - Bellovin S. M., Dutta P. K., Reitinger N. Privacy and synthetic datasets //Stan. Tech. L. Rev. – 2019. – Т. 22. – С. 1.

Гетьман А.И., Иконникова М.К., Степанов И.А.
Институт системного программирования им. В.П. Иванникова РАН

ПРОБЛЕМЫ ПОДГОТОВКИ НАБОРОВ ДАННЫХ ДЛЯ КЛАССИФИКАЦИИ СЕТЕВОГО ТРАФИКА

Одной из важных задач анализа сетевого трафика является классификация потоков данных. Примерами конечных классов являются:

— используемые протоколы прикладного уровня;

— приложения, генерирующие трафик;

— действия, производимые пользователем и т.д.

Для решения этой задачи часто используются методы машинного обучения, так как они обладают существенными преимуществами перед другими подходами. Рассматриваемые в работе методы являются представителями класса методов обучения с учителем, то есть для обучения модели требуется размеченный набор данных, содержащий примеры всех используемых классов. Для самых популярных задач машинного обучения существуют "классические" наборы данных, которые можно использовать для обучения моделей и сравнения полученных результатов с другими исследователями. Однако, в задаче

классификации интернет-трафика в целом нельзя выделить такой набор (наборы). В данном исследовании рассматриваются проблемы подготовки наборов данных для классификации сетевого трафика и причины, почему не существует универсальных общедоступных наборов для этой задачи.

Способы получения трафика для классификации могут отличаться от исследования к исследованию. Самым простым можно считать перехват (зеркалирование) реального трафика какой-то сети, но здесь возникает проблема разметки таких данных. Кроме того, полученные наборы данных являются несбалансированными по количеству примеров каждого класса. Второй подход - это контролируемый сбор данных, когда записывается только трафик, генерируемый целевыми приложениями, но тут сложность представляет организация фильтрации фонового трафика. Для получения большого числа примеров без привлечения реальных пользователей могут создаваться генераторы искусственных данных, имитирующие реальный трафик на основе анализа образцов из сети. Однако, в этом случае нужно следить, чтобы получаемые таким образом данные не были излишне однообразными и соответствовали реальному положению дел в сети. Отдельной проблемой встаёт постоянное появление новых протоколов и приложений, требующих поддержки наборов данных в актуальном состоянии и умения определять ранее неизвестные классы.

Для разметки перехваченного из сети трафика инструменты DPI сопоставляют содержимое пакетов с базой данных имеющихся сигнатур. Такие инструменты требуют больших временных затрат и затрат по памяти, а также могут испытывать сложности при работе с шифрованным трафиком. В [1] рассматриваются примеры наиболее популярных инструментов DPI и сравниваются получаемые с их помощью результаты. Авторы приходят к выводу, что разные инструменты могут выдавать разные метки классификации, что не позволяет гарантировать правильность разметки обучающего и тестового набора данных. Таким образом, даже если предположить, что некоторые выбросы в обучающей выборке могут быть сглажены в процессе обучения, получаемую оценку работы классификатора можно считать лишь приблизительной. Указанная работа не рассматривала инструмент разбора пакетов Wireshark, хотя он тоже может использоваться для указанных целей. Нами было проведено сравнение результатов работы Wireshark и nDPI на тестовом pcap файле, которое подтвердило выводы авторов статьи - расхождения присутствуют, хоть и в небольшом количестве.

Поскольку трассы, полученные из сетевого трафика, могут содержать в себе конфиденциальные данные пользователей сети и данные, собранные без их согласия, обмен наборами данных без ограничений не всегда возможен. Также, нельзя распространять только наборы признаков, так как они могут быть недостаточны для предлагаемых другими исследователями методов или содержать ошибки вычисления (как наборы данных, собранные с использованием достаточно популярного инструмента CICFlowmeter: ISCX2012, CICIDS2017, CICIDS2018 и др.).

Соблюсти баланс между конфиденциальностью и открытостью может помочь анонимизация распространяемых данных, скрывающая содержащуюся в них информацию о конкретных пользователях, но сохраняющая полезные для исследований признаки. Был проведён обзор найденных инструментов анонимизации, не все из которых, к сожалению, удалось протестировать. Большинство параметров шифрования в них предназначено для транспортного и сетевого уровней, для более ранних из них уже были разработаны схемы

взлома метода анонимизации. Не было найдено инструментов более широкого профиля, содержащих все желаемые функции.

Список литературы:

1. Carela-Español V., Bujlow T., Barlet-Ros P. Is our ground-truth for traffic classification reliable? // International Conference on Passive and Active Network Measurement. Springer, Cham, 2014, pp. 98-108.

Ломако А.Г., Менисов А.Б.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург

ЛАНДШАФТ УГРОЗ БЕЗОПАСНОСТИ ИНФОРМАЦИИ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

** Исследование выполнено в рамках гранта Президента РФ для государственной поддержки молодых российских ученых – кандидатов наук(МК-2485.2022.4).*

Технологии искусственного интеллекта (ИИ) снова набрали популярность в последнее десятилетие, повышая эффективность принятия решений в различных сферах [1]. На сегодняшний день происходит цифровая трансформация, результатами которой являются конвергенции различных технологий (например, Интернета вещей, робототехники, квантовых технологий и т. д.) и растущего объема и разнообразия данных, а также их новых характеристик (например, распределенных данных) для использования ИИ [2]. В контексте информационной безопасности ИИ можно рассматривать как новый подход, и, соответственно, методы ИИ используются для поддержки и автоматизации соответствующих операций, например, фильтрации трафика [3-5], исследования инцидентов [6-10] и т.д. Не смотря на полезность ИИ, не следует игнорировать факт, что ИИ и его применение может подвергнуть организации новым, а иногда и непредсказуемым угрозам, и может открыть новые возможности для злоумышленников [11, 12]. Это особенно важно в вопросах безопасности объектов критической информационной инфраструктуре, таких как информационные системы, информационно-телекоммуникационные сети и автоматизированные системы управления военного назначения, интеллектуального производства, электронного здравоохранения и т. д. [13-15].

В настоящее время перечень угроз ИИ не является полным [16]. Создание доверенной системы, использующей ИИ, невозможно без знания специфики применяемых ИИ, без знания поверхности компьютерных атаки, целей и возможностей нарушителей. Для составления полного перечня угроз, специфичных для ИИ, необходимо проведение комплексного исследования прикладных задач, используемых в них технологий ИИ и определение моделей нарушителя.

При рассмотрении вопросов информационной безопасности ИИ необходимо знать, что методы и системы, использующие ИИ, могут привести к неожиданным результатам или могут быть изменены для манипулирования ожидаемыми результатами [17]. Это особенно актуально при разработке программного обеспечения ИИ, которое часто основано на моделях

черного ящика [18]. Стоит отметить, что ИИ может использоваться злоумышленниками, например, ИИ как средство усиления киберпреступности и облегчения процессов компьютерных атак.

Поэтому оценивания угрозы ИИ следует исследовать следующие вопросы [19]:

- определение возможных частей ИИ и процессов применения ИИ, на которые реализуют (возникают) угрозы безопасности информации;
- определение источников угроз безопасности информации и оценка возможностей нарушителей по реализации угроз безопасности информации ИИ;
- определение негативных последствий, которые могут наступить от реализации (возникновения) угроз безопасности информации ИИ;
- определение способов реализации (возникновения) угроз безопасности информации ИИ.

Сложность и обширность негативных последствий требуют создания основы для общего понимания соответствующих угроз безопасности информации ИИ. Это имеет особое значение, учитывая долгосрочные цели в отношении ИИ. Безопасная экосистема ИИ должна поставить информационную безопасность и защиту информации на первое место и способствовать соответствующим инновациям, наращиванию потенциала, повышению осведомленности и инициативам в области исследований и разработок.

Список литературы:

1. Попов В. Г., Галиаскаров Д. Ф., Болябкин М. В. Оценка актуальности и эффективности интеграции систем искусственного интеллекта в сегменте информационной безопасности //Научный электронный журнал Меридиан. – 2021. – №. 3. – С. 130-132.
2. Петренко С. А. Киберустойчивость индустрии 4.0. – 2020.
3. Khudoyarova A., Burlakov M., Kupriyashin M. Using Machine Learning to Analyze Network Traffic Anomalies //2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). – IEEE, 2021. – С. 2344-2348.
4. Cvitić I. et al. Ensemble machine learning approach for classification of IoT devices in smart home //International Journal of Machine Learning and Cybernetics. – 2021. – С. 1-24.
5. Alshammari A., Aldribi A. Apply machine learning techniques to detect malicious network traffic in cloud computing //Journal of Big Data. – 2021. – Т. 8. – №. 1. – С. 1-24.
6. Shakeel P. M. et al. Internet of things forensic data analysis using machine learning to identify roots of data scavenging //Future Generation Computer Systems. – 2021. – Т. 115. – С. 756-768.
7. Sharma S., Krishna C. R., Kumar R. RansomDroid: Forensic analysis and detection of Android Ransomware using unsupervised machine learning technique //Forensic Science International: Digital Investigation. – 2021. – Т. 37. – С. 301168.
8. Hina M. et al. Sefaced: Semantic-based forensic analysis and classification of e-mail data using deep learning //IEEE Access. – 2021. – Т. 9. – С. 98398-98411.

9. Manzoor N. et al. Role of machine learning techniques in digital forensic investigation of botnet attacks //International Journal of Management (IJM). – 2021. – Т. 12. – №. 2.

10. Iqbal S., Alharbi S. A. Advancing automation in digital forensic investigations using machine learning forensics //Digital Forensic Science. – 2020. – С. 3.

11. Andrade R., Torres J., Flores P. Management of information security indicators under a cognitive security model //2018 IEEE 8th annual computing and communication workshop and conference (CCWC). – IEEE, 2018. – С. 478-483.

12. Kaloudi N., Li J. The ai-based cyber threat landscape: A survey //ACM Computing Surveys (CSUR). – 2020. – Т. 53. – №. 1. – С. 1-34.

13. Федеральный закон от 26 июля 2017 года № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации».

14. Приказ ФСТЭК России от 21 декабря 2017 года № 235 «Об утверждении Требований к созданию систем безопасности значимых объектов критической информационной инфраструктуры Российской Федерации и обеспечению их функционирования».

15. Приказ Федеральной службы по техническому и экспортному контролю от 25 декабря 2017 года № 239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации».

16. База угроз безопасности информации ФСТЭК [Электронный ресурс]: https://bdu.fstec.ru/threat (дата обращения 01.02.2022).

17. Володин И. В. и др. Классификация механизмов атак и исследование методов защиты систем с использованием алгоритмов машинного обучения и искусственного интеллекта //Прикаспийский журнал: управление и высокие технологии. – 2021. – №. 2 (54). – С. 91-98.

18. Papernot N. et al. Practical black-box attacks against machine learning //Proceedings of the 2017 ACM on Asia conference on computer and communications security. – 2017. – С. 506-519.

19. Методический документ ФСТЭК России от 5 февраля 2021 г. «Методика оценки угроз безопасности информации».

Огнев Р.А., Зегжда Д.П., Жуковский Е.В.
Санкт-Петербургский политехнический университет Петра Великого

ОЦЕНКА УСТОЙЧИВОСТИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ К СОСТЯЗАТЕЛЬНЫМ АТАКАМ В УСЛОВИЯХ НЕПОЛНОЙ ИНФОРМАЦИИ

Развитие бизнес-модели «машинное обучение как сервис» (MLaaS), в которой пользователь может загрузить свой собственный набор данных для обучения на сервер и после чего получит возможность отправлять запросы к обученной модели, привело к дополнительному вниманию к атакам в условиях неполных знаний («черный ящик», black-box), когда злоумышленник имеет доступ только ко входу и выходу обученной модели. На данный момент выделяют три класса атак методом «черного ящика»:

1. Неадаптивные атаки: злоумышленник имеет доступ к распределению обучающих данных и переносит атаки «белого ящика» на суррогатную модель, аппроксимирующую целевую. В частности, можно обучить локальную модель с архитектурой f' на исходном распределении данных μ. f' ведет себя аналогично исходной модели f, и, следовательно, можно проводить состязательные атаки на f' и переносить их на f.

2. Адаптивные атаки: у злоумышленника нет информации об обучающем наборе данных или модели. Таким образом, он должен обрабатывать целевую модель как оракул, адаптивно запрашивая модель с данными X, не связанных с обучением, для получения меток Y. Атакующий обучает суррогатную модель с архитектурой f' по кортежам (X, Y). Он проводит атаки белого ящика на f' и переносит их на целевую модель f. В данном подходе запросы X выбираются в зависимости от реакции базовой модели.

3. Строгие атаки: злоумышленник также может собирать кортежи (X, Y), опрашивая целевую модель. Однако, в отличие от адаптивных атак, эти модели не могут изменять входные данные.

Как только локальная модель обучена с высокой степенью достоверности, любая из стратегий атаки «белого ящика» может быть применена к локальной модели для создания состязательных примеров, которые в конечном итоге могут быть использованы для обмана целевой модели из-за свойства переносимости нейронной сети (transferability). Однако в случае адаптивного сценария «черного ящика» злоумышленник не имеет никаких сведений об обучающей выборке, что делает ее наиболее применимой в реальной жизни.

В работе [1] представлена одна из первых атак на классификаторы глубоких нейронных сетей, которую авторы обобщили и на другие модели машинного обучения в киберпространстве в условиях реального мира. Авторы создают модель-копию целевого классификатора путем генерации синтетических данных, реализуют на нее известные состязательные атаки «белого ящика» и затем переносят их на MLaaS нейронные сети MetaMind, Amazon и Google. Они смогли показать, что предоставленные облачные модели неправильно классифицировали 84,24%, 96,19% и 88,94% состязательных примеров, сгенерированных их методом.

Основная идея предложенного метода заключалась к генерации синтетических данных при помощи техники Jacobian-based Dataset Augmentation ($\tilde{O}-$ оракул, J_F- матрица Якоби суррогатной модели, S_i- обучающая выборка на i итерации):

$$S_{p+1} \leftarrow \{\vec{x} + \lambda * sign(J_F[\tilde{O}(\vec{x})]) : \vec{x} \in S_p\} \cup S_p$$

Однако, данный подход требует большого количества запросов к оракулу, что может быть обнаружено системами проактивной защиты. В ответ на это злоумышленник может применить распределенную атаку, что позволит ему обмануть систему защиты. Следующим шагом для защиты может быть внедрение на этап тестирования «временных бомб», которые замедляют генерацию синтетических данных, но в тоже время они не должны существенным образом увеличивать время ответа системы на запросы пользователей.

Идея итеративной генерации новых входных данных из старых на основе ответа оракула, используемая в адаптивных атаках, может использоваться для их обнаружения. В работе предложен алгоритм, внедряемый в работу системы машинного обучения на этап,

предшествующий тестированию входных данных. Основными задачами, решаемыми предложенным алгоритмом, являются:

1. Выявление показателя зависимости между текущими и историческими входными данными
2. В зависимости от показателя, то есть обнаружения попытки атаки, в работу системы добавляется временная задержка, значение которой увеличивается экспоненциально в зависимости от степени корреляции.

При вычислении показателя зависимости между данными основной задачей является обнаружение самых незначительных изменений. Для реализации данной идеи используется двухэтапный подход. На первой стадии используется fuzzy-хеширование для быстрой фильтрации данных и построения множества исторических данных, которые близки к тестируемым. На втором этапе используется ресурсоемкая фильтрация на основе матрицы ковариации, которая позволяет обнаруживать пары данных с минимальной разницей. Предложенный подход позволяет обнаруживать попытки адаптивных атак, в основе которых лежит идея итерационной генерации новых входных данных.

Список литературы:

1. Papernot N. et al. Practical black-box attacks against machine learning //Proceedings of the 2017 ACM on Asia conference on computer and communications security. – 2017. – С. 506-519.

Югай П.Э, Жуковский Е.В.
ООО «НеоБИТ»

ОБНАРУЖЕНИЕ И КЛАССИФИКАЦИЯ ВРЕДОНОСНЫХ УСТАНОВОЧНЫХ ФАЙЛОВ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Внедрение вредоносного программного обеспечения (ВПО) в замаскированные под легитимные установочные файлы является распространенной техникой атак на компьютерные системы. В особенности данное явление широко распространено для установочных файлов операционной системы Windows.

Так, например, в Интернете были обнаружены экземпляры ВПО RedLine Stealer, замаскированные под обновление Windows 11, которые осуществляют кражу учетных данных и номеров банковских карт из браузеров, логинов и паролей электронных почт, мессенджеров и прочей конфиденциальной информации.

Другим примером, который был обнаружен осенью 2021 года, является внедрение вредоносного функционала в установочные файлы редактора текста Notepad++ группой злоумышленников, известных как StrongPity. Эта группа также известна в распространении зараженных установщиков WinRAR в рамках целенаправленных атак в период с 2016 по 2018 годы. Перечисленные примеры внедрения вредоносного функционала в замаскированные под легитимные установочные файлы говорят об актуальности данной проблемы, которая требует эффективных решений по ее устранению.

Одним из наиболее актуальных подходов к выявлению подобных вредоносных установочных файлов является применение машинного обучения, которое позволяет обеспечивать высокую точность обнаружения вредоносных файлов и может быть внедрено в компоненты защиты.

В докладе будет представлен метод выявления вредоносных исполняемых файлов, основанный на использовании машинного обучения, а также приведены результаты исследования особенностей легитимных и вредоносных установочных файлов, выделены особенности троянов-установщиков и троянов-загрузчиков. Приведен сравнительный анализ, применимости для решения указанной задачи различных методов машинного обучения: наивный байесовский классификатор, случайный лес и алгоритм C.45. В результате применения метода взаимной информации удалось сократить количество используемых признаков до 49 наиболее значимых атрибутов исполняемых файлов, которые дают положительные результаты при классификации легитимных установочных файлов и троянских программ.

Вавилова А.С., Волошина Н.В.
Университет ИТМО, г. Санкт-Петербург

МЕТОДИКА ОБНАРУЖЕНИЯ УЯЗВИМОСТЕЙ, СВЯЗАННЫХ С ПАРАМЕТРАМИ НЕЙРОННОЙ СЕТИ, В АЛГОРИТМАХ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Технология машинного обучения применяется в различных сферах жизни общества: фильтрация контента в социальных сетях, формирование рекомендаций на сайтах коммерческих организаций, товаров для конечного потребителя (камеры, смартфоны). Системы на основе методов глубокого обучения используются для идентификации объектов на изображениях [1], распознавании естественной речи [2], соотнесения новостей, публикаций или продуктов с интересами пользователей и выявления релевантных результатов поиска.

О наличии множественных недостатков и уязвимостей в работе алгоритмов машинного обучения свидетельствует большое количество публикаций, основная тематика которых касается вопросов обеспечения безопасности данных в сложноорганизованных системах на основе нейронных сетей. Существует множество опубликованных исследований об атаках, способах защиты информации и методах обеспечения конфиденциальности данных в алгоритмах машинного обучения. В качестве механизмов обеспечения безопасности в таких системах были предложены различные методы, такие как состязательное обучение [3], генеративная состязательная сеть, статистический подход и повторяющаяся нейронная сеть.

Источником уязвимости в нейронной сети являются не только данные, но и параметры модели. В виртуальной среде существуют различные причины искажения параметров нейронной сети, например отравление обучающих данных [4], сжатие [5] и квантование параметров [6]. Для аппаратных нейронных сетей [7] искажения параметров возникают из-за износа оборудования или фонового шума.

Исследование уязвимостей, основанных на искажении параметров модели машинного обучения, является одним из наиболее важных направлений в изучении аспектов

информационной безопасности нейросетевых технологий. Нахождение техник, позволяющих противодействовать атакам, базирующимся на искажении параметров, может привести не только к повышению надежности параметров нейронной сети, но и к улучшению показателя точности работы модели. Основной целью проводимого исследования является повышения уровня обеспечения безопасности алгоритмов на основе нейронных сетей с помощью разрабатываемой методики определения уязвимых к искажению параметров нейронных сетей.

Для исследования уязвимостей параметров нейронной сети и оценки надежности параметров предлагается показатель, который измеряет максимальное изменение потерь, вызванное искажениями (под искажениями понимаются искусственные искажения параметров) параметров модели в случае возникновения неблагоприятного сценария работы нейронной сети. Определение показателя производится на основе бесконечно малого градиента, что является наиболее эффективной оценкой по сравнению с случайными искажениями, применение которых не вызывает оптимальное снижение значений функции потерь.

Применение показателя дает возможность проведения систематического анализа надежности параметров и исследования уязвимостей различных компонентов в глубокой нейронной сети при ухудшении точности распознавания после реализации атак на основе искажения параметров. Сравнение методов искажения на основе градиента и случайного искажения параметров с помощью введенного параметра показывает, что окрестности параметров на поверхности функции потерь плоские, за исключением некоторых крутых участков функции, при возможности выбора параметров на значительном расстоянии от крутых участков функции потерь надежность выбранных параметров при работе алгоритмов на основе нейронных сетей значительно улучшится. Основным предложением в рамках разрабатываемой методики является проведение состязательно устойчивого к искажению обучения, процесс которого включает искажения параметров для поиска параметров, «далеких» от окрестностей крутых участков функции потерь, для достижения большего уровня надежности параметров нейронной сети.

В рамках доклада будут представлены эмпирические данные по использованию предложенного показателя в задачах улучшения стойкости к искажению алгоритмов на основе нейронных сетей.

Таким образом, в рамках разрабатываемой методики предполагается применение специального параметра – показателя, характеризующего максимальную высоту подъема в пределах определенного расстояния от текущего параметра функции потерь. Для повышения надежности модели в части параметров нейронной сети предлагается улучшение процесса обучения глубоких нейронных сетей с учетом уязвимых параметров в части внедрения состязательного обучения, устойчивого к искажению, которое позволит повысить точность распознавания и производительность обобщения в глубоких нейронных сетях.

Список литературы:

1. Szegedy C. et al. Rethinking the inception architecture for computer vision //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 2818-282.

2. Xu B. et al. NADAQ: natural language database querying based on deep learning //IEEE Access. – 2019. – Т. 7. – С. 35012-35017.

3. Ilyas A. et al. Adversarial examples are not bugs, they are features //arXiv preprint arXiv:1905.02175. – 2019.
4. Gu T. et al. Badnets: Evaluating backdooring attacks on deep neural networks //IEEE Access. – 2019. – Т. 7. – С. 47230-47244.
5. Arora S. et al. Stronger generalization bounds for deep nets via a compression approach //International Conference on Machine Learning. – PMLR, 2018. – С. 254-263.
6. Nagel M. et al. Data-free quantization through weight equalization and bias correction //Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2019. – С. 1325-1334.
7. Salimi-Nezhad N. et al. A digital hardware system for spiking network of tactile afferents //Frontiers in neuroscience. – 2020. – С. 1330.

3. Социальные коммуникации цифрового общества: доверие и безопасность

Соловей Р.С., Дахнович А.Д., Москвин Д.А.
Санкт-Петербургский политехнический университет Петра Великого

СОВРЕМЕННЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ АВТОМАТИЗИРОВАННЫХ АККАУНТОВ В СОЦИАЛЬНЫХ СЕТЯХ

В современном мире социальные сети имеют большое влияние. По последним оценкам 58.4% населения пользуются социальными сетями, а среднее время использования социальных сетей составляет 2 часа 27 минут в сутки [1]. При этом социальные сети используется не только для общения и получения образовательного и развлекательного контента, но и для получения новостей. Согласно исследованию фонда «Общественное мнение», 19-% россиян предпочитают получать новости из социальных сетей [2].

С ростом количества пользователей растет и количество автоматизированных аккаунтов (ботов) в социальных сетях. Основными целями создания автоматизированных аккаунтов как правило являются:

1. Автоматизированный сбор данных для использования на сторонних ресурсах.
2. Использование аккаунтов для влияния на общественное мнение.
3. Использование аккаунтов для мошенничества.

Присутствие автоматизированных аккаунтов в социальных сетях угрожает и компаниям, владеющими социальным сетям, и их пользователям.

Компании несут как имиджевые потери - удешевление рекламных контрактов и снижение потока инвестиций, так и технические – автоматизированные аккаунты расходуют мощности, предназначенные для легитимных пользователей.

Пользователи, в свою очередь, подвергаются как точечному мошенничеству, так и находятся под воздействием масштабных информационных кампаний, призванных повлиять на общественное мнение. Согласно исследованию [3], 45% публикаций о пандемии коронавируса COVID-19 были созданы аккаунтами, которые больше похожи на ботов, чем на реальных пользователей.

Для обнаружения и борьбы с автоматизированными аккаунтами исследователи используют различные подходы, методы и алгоритмы. В данном докладе приведено исследование актуальных угроз от ботов, исследованы, приведены признаки ботов, по которым их можно обнаружить и описаны наиболее актуальные методы обнаружения ботов в различных источниках.

Список литературы:

1 Global social media statistics research summary 2022. // URL: https://www.smartinsights.com/social-media-marketing/social-media- strategy/new-global-social-media-research/. (дата обращения 27.05.2022)

2 Источники информации: интернет. Востребованные источники информации в интернете. // URL: https://fom.ru/SMI-i-internet/14538. (дата обращения 27.05.2022)

3 Uyheng J., Carley K. Bots and online hate during the COVID-19 pandemic: Case studies in the United States and the Philippines // Journal of Computational Social Science – 2020. – № 3. P. 445-468. DOI: https://doi.org/10.1007/s42001-020-00087-4

Тельбух В.В.
Военно-космическая академия им. А.Ф. Можайского, г. Санкт-Петербург

ПРОГНОЗИРОВАНИЕ ПОДВЕРЖЕННОСТИ АУДИТОРИИ СОЦИАЛЬНЫХ СЕТЕЙ ЦЕЛЕНАПРАВЛЕННОМУ НЕГАТИВНОМУ ИНФОРМАЦИОННОМУ ВОЗДЕЙСТВИЮ

Согласно проведенного исследования [1], социальные сети являются основным инструментом манипуляции общественным мнением. При этом, в работе [2] отмечается, что управление массовым мнением в деструктивных целях, которыми могут быть,

как подмена культурно-исторических ценностей, так и формирования недоверия к военно-политическому руководству государства. США ведут активную работу в направлении создании инструментов и механизмов формирования выгодного общественного мнения на территории своих геополитических противников [3], что требует выстраивание собственной системы защиты от подобных негативных воздействий.

В этих условияхна первый план выходит выявление и прогнозирование подверженности негативному влиянию аудитории сетевых сообществ.Решение данной задачи позволит выявить потенциальные очаги негативной напряженности в сети и провести упреждающий комплекс мероприятий, позволяющий нивелировать последствия деструктивного воздействия еще на ранних этапах его проведения.

Изучение степени исследования задачи показал, что в работе [4] авторы проводили эксперимент по прогнозированию реакции пользователей социальных сетей методами машинного обучения: опорных векторов, методом градиентного бустинга, случайного леса и многослойной парцептрона. Векторное представление информационных сообщений осуществлялось при помощи методов «BagofWords», TF-IDF, Word2Vec. Для прогнозирования брали статистические данные числа отметок «мне нравится» и «рассказать друзьям». В результате проведенного эксперимента авторы пришли к выводу, что лучший прогноз показал метод многослойного перцептрона с коэффициентом детерминации 0,57. Таким образом, можно сделать вывод, что методы, основанные на статистическом подходе малоэффективны.

В работе [5] оценивали степень подверженности пользователей антиководным публикациям. Авторы создали алгоритм под названием AVAXTAR, при помощи которого был собран Dataset текстовых сообщений, опубликованных в социальной сети «Twitter», имеющих хештеги против вакцинации. Полученные данные были обработаны при помощи модели векторного представления Sent2Vec. Далее сформирован 600-мерный векторный словарь признаков для каждого негативного образца сообщений. На основе сформированного словаря алгоритм классифицировал публикации, подаваемые на вход. В результате эксперимента точность классификации составила 0,5938. Низкий показатель классификации связан большим разнообразием словоформ и их комбинациями, которые могут иметь разный контекст.

Авторы статьи [6] в своей работе сосредоточили свое внимание на эволюционной классификации настроений. В ходе исследования были собраны корпуса текстов 160 тыс. сообщений из сети «Twitter». При помощи моделей BoW и N-Gram определили наиболее часто обсуждаемые темы. Используя анализ настроений NLTK ранжировали темы по популярности среди пользователей, классифицировав их на два класса: положительные и отрицательные. Полученный набор данных с помощью word2vec преобразовали векторное представление слов с размерностью 200 признаков. На полученных данных обучили пятислойную модель LSTM. Далее провели эксперимент, обучив LSTM на двух наборах данных: 1 – без учета рейтинга настроений и 2 – с учетом рейтинга настроений. В результате проведенного эксперимента, прогнозирование популярности определенных тем среди пользователей по заданным метрикам улучшилась с 84,46% до 91,67%.

Из приведенных экспериментов следует, что прогнозирование популярности новостных событий, а равно подверженности целевой аудитории информационно-психологическому воздействию, путем применения методов машинного обучения может быть эффективным при условии улучшения количества и качества обработки обучающих данных. Однако, следует отметить, что данные подходы не учитывают внутренние и внешние факторы, воздействующие на целевую аудиторию (конкретного пользователя), которые могут изменить их предпочтения в результате целенаправленного информационного воздействия. Также не берется во внимание текущая динамика популярности тем в информационных потоках, т.е. модели обучаются на ретроспективных данных.

Таким образом, методы машинного обучения могут применяться в качестве определения одного из признаков подверженности аудитории целенаправленному информационно-психологическому воздействию. Видится, что в рамках решаемой задачи необходимо учитывать связи между субъектом (агент сетевого влияния) и объектами (пользователь и сообщества в сети) воздействия – теория сетевых графов. Также следует взять во внимание особенности «вирусного» распространения контента, данные технологии широко применяются в маркетинге и спецслужбами иностранных государств в ходе информационных операций, и могут быть описаны закономерностями, выявленными в эпидемиологии.

Список литературы:

1. Строганов, В. Б. Технологии политической манипуляции в интернете: специальность 23.00.02 «Политические институты, процессы и технологии» : диссертация на соискание ученой степени кандидата политехнических наук / Строганов Вадим Борисович ; Уральский

гуманитарный институт. - Екатеринбург, 2019. - 361 с. - Библиогр.: с. 18–23. - Текст : непосредственный.

2. Володенков С.В. Киберсимулякры как инструмент виртуализации современной массовой политической коммуникации // Информационные войны. 2014, №4, С. 19 с.18-21 - Текст : непосредственный.

3. U.S. Air Force. Persona Management Software. URL: http://seankerrigan.com/docs/PersonaManagementSoftware.pdf (дата обращения 20.05.2022).

4. Прогнозирование реакции пользователей в социальных сетях методами машинного обучения / Е. П. Попова, В. Н. Леоненко, Н. Г. - Текст : непосредственный // Научно-технический вестник информационных технологий, механики и оптики. - 2020. -

№ 1. - С. 118-124.

5. Matheus S., Goran M., Keith B., A Python Package to Detect Anti-Vaccine Users on Twitter. Availableat: https://doi.org/10.48550/arXiv.2110.11333 (Accessed 22 may 2022).

6. Arunava K., Sourav D., Anup K., Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. In RAAI 2020. Advances in Intelligent Systems and Computing, vol 1355 (2021). Available at: https://doi.org/10.48550/arXiv.2106.06910 (Accessed 22may 2022).

Бессольцев В.Е., Гильмуллин Р.М.
Военно-космическая академия им. А.Ф. Можайского, г. Санкт-Петербург

ПОДХОД К АВТОМАТИЗИРОВАННОМУ МОНИТОРИНГУ TELEGRAM-КАНАЛОВ

Мировая военно-политическая обстановка всегда является предметом споров, дискуссий и множества разнообразных точек зрения различных геополитических организаций, политиков, военных специалистов, а также рядовых граждан по всему миру. В свете интеграции цифровых технологий в различные сферы человеческой деятельности результатом таких споров становятся публикации, зачастую провокационные, в различных источниках СМИ, в частности в веб-пространстве, размещаемые противостоящими сторонами с целью достижения превосходства путём пропаганды идеологии насилия, терроризма и неприкрытого вмешательства во внутренние дела других государств [1], что нередко приводит к возникновению военных конфликтов. При этом складывающаяся в мире обстановка формирует определенные условия и факторы, создающие прямую возможность нанесения ущерба национальным интересам и способность влияния на состояние национальной безопасности нашей страны.

В этом ключе, в рамках проводимого авторами статьи исследования, в целях реализации подхода к автоматизированному мониторингу Телеграм-каналов используется веб-версия приложения Телеграм, для которой на языке программирования JavaScriptразработано специализированное расширение zScan (далее по тексту – zScan), общий принцип которого представлен на рисунке 1.

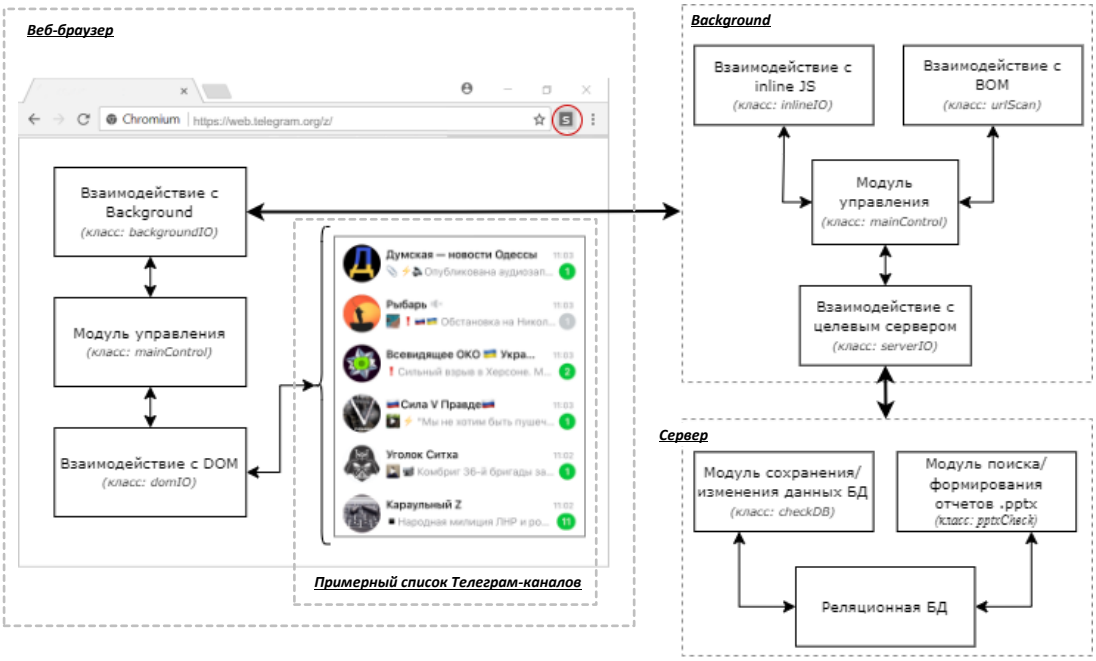

Рисунок 1 - Общий принцип работы zScan

Телеграмм используют технологию идентификации на предмет того, кто выполняет действия на сайте, человек или поисковый робот (webcrawler), поэтому применение других языков программирования, например,PHP, Python, Ruby, приведет к разработке программного модуля, который будет однозначно идентифицирован как поисковый робот, что приведет к мгновенной блокировке его функционирования.

Основным преимуществом реализуемого подхода к автоматизированному мониторингу телеграмм-каналов, схема алгоритма сканирования которого представлена на рисунке 2, являетсявозможность выполнения действий от лица пользователя с имеющейся легитимной метаинформацией об автоматизированном рабочем месте пользователя (например, отправляемые заголовки –header), в том числе имитация действий пользователя по взаимодействию с веб-версией Телеграм (прокрутка содержимого веб-страницы, клики манипулятором типа «мышь»).

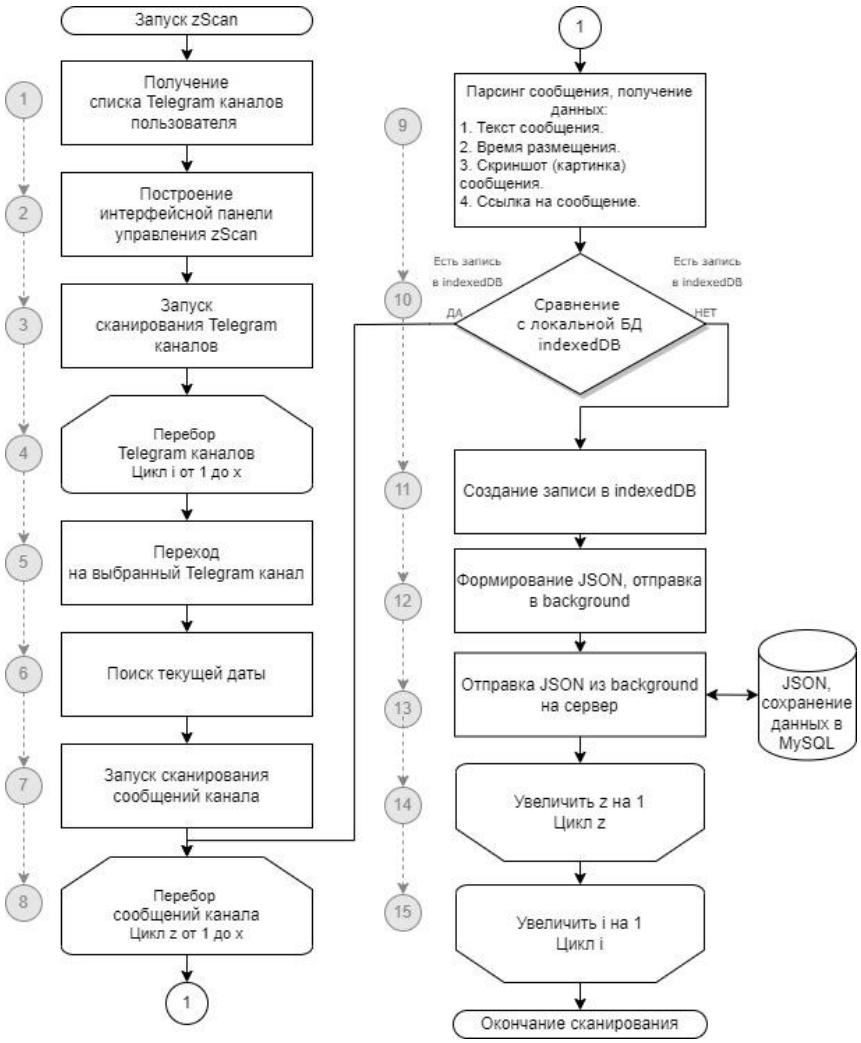

Рисунок 2 – Схема алгоритма сканированияz Scan

Для хранения собранной информации с выбранных Телеграм-каналов авторами предложенного подхода используется веб-сервер на основе Node.js [2] и сервер баз данных MySQL [3]. Собранная и сохраненная информация может использоваться для проведения дальнейших исследований по анализу семантических связей сообщений, поиска наиболее информативных по содержанию каналов, каналов, которые оперативно размещают информацию по требуемым ключевым запросам и т.д.

Список литературы:

1. Фёдоров Е.А.: [сайт] / Депутат ГД РФ. Действительный государственный советник РФ.URL: http://eafedorov.ru/publicatsii/publicatsii-1760.html (дата обращения: 17.05.2022).
2. Nodejs.org: проект JavaScript с открытым исходным кодом. URL: https://nodejs.org/en (дата обращения: 13.05.2022).
3. MySQL.com: СУБД. URL: https://www.mysql.com (дата обращения: 17.05.2022).

Киселёв А.Н., Дворянов Н.А., Криковцев А.С., Петров А.М.
Военно-космическая академия им. А.Ф. Можайского, г. Санкт-Петербург

ПОДХОД К ТЕСТИРОВАНИЮ ЗАЩИЩЕНОСТИ WEB-СЕРВЕРА ОТ ПЕРСПЕКТИВНЫХ DOS И DDOS АТАК С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННЫХ МЕЖСЕТЕВЫХ ЭКРАНОВ

В настоящее времяв сети Интернет стремительно увеличилось число DDoS-атак.Исследование «Лаборатории Касперского» показывает, что количество DDoS-атак на российские компании в марте 2022 года выросло в восемь раз. При этом их средняя длительность увеличилась с 12 минут до 30 часов. На сегодняшний день, проблема защиты от DoS и DDoS-атак стоит очень остро, а если учитывать постоянно растущие пропускные способности сетей и быстро меняющуюся обстановку в мире, то в ближайшее время можно ожидать значительного роста таких сетевых атак. Так, например, в апреле 2022 года неназванная криптовалютная платформа подверглась одной из крупнейших DDoS-атак в истории, чья мощность достигала 15,3 млн запросов в секунду. Таким образом, разработка перспективных методик и средств защиты от масштабных DoS и DDoS-атак на сетевую инфраструктуру является крайне актуальной.

DoS (DenialofService «отказ в обслуживании») – сетевая атака на вычислительную систему с целью довести её до отказа, то есть создание таких условий, при которых добросовестные пользователи системы не смогут получить доступ к предоставляемым системным ресурсам (серверам), либо этот доступ будет затруднён.DDoS (Distributed Denial of Service) – разновидностьDoS-атаки. Главным отличием DoS-атаки от DDoS-атаки является то, что DoS проводится с одной машины (одного IP-адреса), а DDoS с множества машин (множества разных IP-адресов).

При проведении на DoS и DDoS-атак на web-сервер, производительность работы и время отклика сервера значительно ухудшается,переводя сервер в нерабочее состояние и делая его недоступным для пользователей.

В рамках исследований были смоделированы актуальные DDoS-атаки, с помощью различного ПО тестирования web-серверов – GoldenEye; MHDDoS; LOIC,которые осуществляют генерацию вредоносного трафика путем перемешивания запросов с различных браузеров, операционных систем и рефералов с применением различных протоколов и на разных уровнях модели OSI (Рисунок 1). Пример содержания черного списка Iptables до начала атаки приведен на рисунке 2.

Рисунок 1 – Разработка сценарияDDoS-атаки с помощью MHDDoS

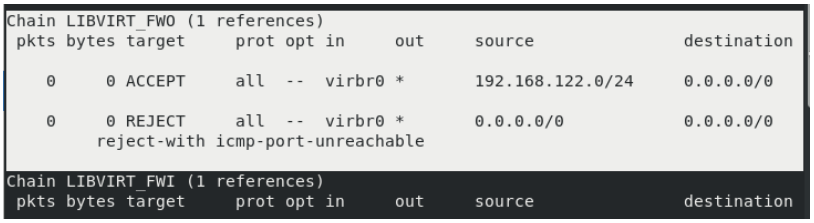

Рисунок 2 - Содержание черного списка Iptablesдо начала атаки

Для защиты от DoS и DDoS-атак,в рамках проведенных исследований, был разработан скрипт, который автоматически блокирует всеIP-адреса, с которых поступило превышенное количество запросов, заданное администратором безопасности от одного пользователя за короткий промежуток времени (Рисунок 3). Данные IP-адреса помещаются в черный список межсетевого экранаIptablesи находятся там до момента ручной разблокировки. Для быстрой разблокировки требуемогоIP адреса также был разработан дополнительный скрипт.

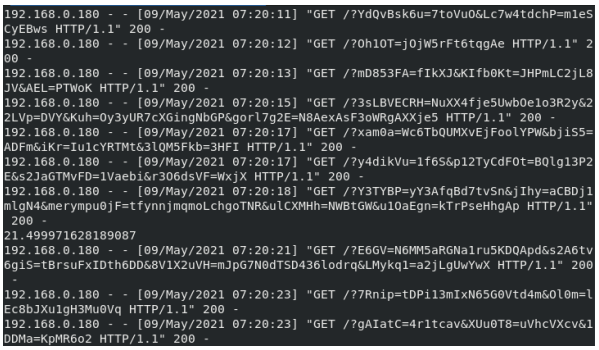

Рисунок 3 – Полученные сервером запросы при совершении атаки

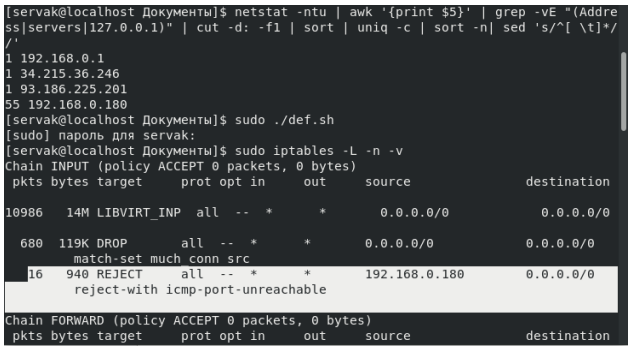

Рисунок 4 - Содержание черного списка межсетевого экрана Iptables после запуска скрипта автоматической блокировки

На рисунке 4показано как с IP-адреса 192.168.0.180 было отправлено 55 HTTP запросов, после чего был запущен скрипт def.sh. В результате выполнения скрипта IP-адрес был автоматически добавлен в черный список Iptables, после чего атака прекратилась.

Таким образом, в рамках проведенных исследованийбыл разработан подход для защиты от DoS и DDoS-атак с использованием автоматизированных средств конфигурирования утилиты управления работой межсетевого экранаIptables. Данный подход целесообразно использовать для защиты как от DoS-атак (совершаемых с одногоIP-адреса), так и от DDoS-атак(совершаемых с нескольких IP-адресов) с помощью зацикливания запуска скрипта. Представленный способ можно использовать в совокупности с другими методами защиты, например, применением нейронной сети, распознающей и нейтрализующей основные виды DDoS-атак.

Список литературы

1. Cloudflare о рекордной DDoS атаке [Электронный ресурс]. – Режим доступа: https://xakep.ru/2022/04/29/new-ddos-record-3/
2. Число DDoS-атак на российские организации [Электронный ресурс]. – Режим доступа: https://habr.com/ru/news/t/658565/
3. DDoSattacksin Q3 2021[Электронный ресурс]. – Режим доступа: https://securelist.com/ddos-attacks-in-q3-2021/104796/
4. Глоссарий ТЦИ [Электронный ресурс]. – Режим доступа: https://tcinet.ru/press-centre/glossary/article.php?ELEMENT_ID=5224

Кочетков И.В., Гильмуллин Р.М., Калинкин И.Д.
Военно-космическая академия им. А.Ф. Можайского, г. Санкт-Петербург

АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ ИДЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ С ПРИМЕНЕНИЕМ ВЕБ-ОБОЗРЕВАТЕЛЕЙ

Современный мир, современные тенденции экономического развития мира, социальные взаимодействия, научные исследования немыслимы без использования сети интернет. С целью обработки данных и обеспечения взаимодействия между информационными ресурсами сети интернет и человеком, как правило, используют специальное программное обеспечение, такое как веб-обозреватель. Последние десятилетия показывают, что наряду с легитимной деятельностью в интернете процветают незаконные виды деятельности, пропагандирующие противоправные действия, такие как распространение информации порнографического характера, информации, направленной на пропаганду войны, разжигание национальной, расовой или религиозной ненависти и вражды, размещение персональных данных граждан Российской Федерации, клевета, торговля оружием и наркотическими веществами.

В этой связи идентификация пользователя сети интернет остаётся актуальной в целях поддержания способности влияния на состояние национальной безопасности нашей страны, что позволит оперативно реагировать на возникающие угрозы безопасности и своевременно нейтрализовать источник угрозы, с учётом использования нарушителями методов обхода блокировок.

Идентификация — это присвоение субъектам и объектам доступа идентификатора и (или) сравнение предъявляемого идентификатора с перечнемприсвоенных идентификаторов.Результаты сравнительного анализа методов присвоения уникального идентификатора приведены в таблице 1.

Таблица 1. Сравнительные характеристики методов идентификации.

Характеристики: Возможность идентификации…	Методы			
	canvasFingerprinting	adobeFlash и JavaScript	Cookie	HTTP - заголовки
типа используемого устройства на стороне клиента	Да	Нет	Да	Да

графического драйвера устройства на стороне клиента	Да	Да	Нет	Нет
процессора устройства на стороне клиента	Да	Нет	Нет	Нет
графического адаптера устройства на стороне клиента	Да	Да	Нет	Нет
операционной системы на устройстве клиента	Да	Да	Да	Нет
используемого клиентом веб-браузера	Да	Да	Да	Нет
Возможность сохранения настроек профиля (язык, шрифты, время)	Да	Да	Да	Да

Приоритетным методом в присвоении fingerprint (отпечатка браузера) является использование технологии canvas Fingerprinting, поскольку позволяет собирать параметры аппаратной части устройства, которая в свою очередь редко подвергается замене, и как следствие, снижение вероятности изменения fingerprint пользователя.

Подход к установления уникального идентификатора заключается в считывании параметров, в качестве которых могут использоваться следующие виды информации, такие как User-Agent (используемое программное обеспечение, вид операционной системы, версии ПО и многое другое), наличие настройки DoNotTrack, язык браузера, часовой пояс, плагины и их характеристики, разрешение экрана и его цветовые параметры (глубина цвета), поддерживаемые технологии HTML5, шрифты, их прорисовка.

Другим методом идентификации веб-обозревателя могут быть использованы куки-файлы. Куки (англ. Cookie) - фрагмент данных, сформированный веб – сервером, хранимый на устройстве пользователя. Обычно веб-браузер при попытке открыть страницу сайта пересылает этот фрагмент данных веб-серверу в составе HTTP-запроса. Куки применяются для сохранения данных на стороне пользователя, обычно используется для сбора статистики, аутентификации, хранения персональных данных, отслеживания сеанса доступа;

Cookie используются веб-серверами для идентификации пользователей и хранения данных о них. Благодаря им пользователь может не вводить пароль каждый раз, когда идентифицирует себя на сайте. Недостатком Cookie является то, что они хранятся на устройстве пользователя их легко удалить тем самым стерев информацию о себе в сети.

Еще одним метода установления отпечатка веб – обозревателя является использование заголовков HTTP — это строки в HTTP-сообщении, содержащие разделённую двоеточием пару имя-значение. Информация в заголовке позволяет определить, используется ли на стороне клиента персональный компьютер или мобильное устройство, в каком часовом поясе расположен пользователь, и даже расположение в сети. К сожалению, именно из этих данных складываются массивы параметров в последствии, из которых и формируются Fingerprint.

Применение же AdobeFlash и JavaScript позволяет передавать больше информации для дальнейшего формирования цифрового следа. Так, если на стороне активирован JavaScript, то вовне передаются данные о плагинах или спецификациях АРМ пользователя. Если установлен и активирован Flash, то это предоставляет стороннему «наблюдателю» еще больше информации.

Таким образом наиболее совершенной технологией присвоения Fingerprint является применение Canvas HTML5.

Список литературы:

1. Л. А. Бояркина, В. В. Бояркин. Цифровой след и цифровая тень как производные персональных данных.. - М.,: Просвещение, 2018. - 45 с.

2. Бессольцев В.Е. Метод идентификации абонентских терминалов информационно-телекоммуникационных сетей в условиях априорной неопределенности относительно адресно-коммутационной информации. // Современная наука. Актуальные проблемы теории и практики. - 2018. - №4. - С. 117-131.

3. Гильмуллин Р.М., Бессольцев В.Е.Цифровой идентификатор веб-обозревателя на основе анализа времени исполнения javascript кода // Современная наука. Актуальные проблемы теории и практики. - 2020. - №5. - С. 73-79.

4. Фингерпринт (отпечаток браузера): что это такое и как его скрыть. // ИД "Комитет" URL: https://vc.ru/u/738105-adspower-browser/212732-fingerprint-otpechatok-brauzera-chto-eto-takoe-i-kak-ego-skryt (дата обращения: 25.05.2021).

Кочетков И.В., Пилькевич С.В.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург

НОВЫЕ АСПЕКТЫ ЗАДАЧИ ИДЕНТИФИКАЦИИ ВЕБ-ОБОЗРЕВАТЕЛЕЙ

В настоящее время глобальная компьютерная сеть Интернет является одним из самых популярных источников информации, который неминуемо сместил фокус влияния традиционных средств массовой коммуникации и изменил традиционные схемы информирования общественности о происходящих в мире событиях. Тем не менее, на ряду с расширением функционала, повышением оперативности доведения новостных сообщений и развитием обратной связи с аудиторией, появились и серьезные проблемы в области кибербезопасности [1].Провоцирование деструктивных действий, массовое распространение вредоносного контента, вызывающего недовольство или панику среди населения, реализуемое через социальные медиа ресурсы сети Интернет, стало представлять собой одну из актуальных угроз национальной безопасности в информационной сфере[2].

В свете изложенногоидентификация пользователя при посещении им удаленного ресурса в сети Интернет представляется одной из важнейших задач современного общества знаний.

Согласно ГОСТ Р ИСО/МЭК 19795-1-2007 идентификация пользователя - процесс при котором осуществляется поиск в регистрационной базе данных и предоставляется список кандидатов, содержащий от нуля до одного или более идентификаторов [3]. Иными словами, присвоение идентификатора – есть деанонимизация пользователя в сети, а именно создание уникального отпечатка Браузера (веб-обозревателя).

Посещение современных удаленных ресурсов сети Интернет, в большинстве случаев,сопряжено с использованием технологии JavaScript. Отключение JavaScript приведет к значительным изменениям итогового документа, и как следствие к полной невозможности восприятия пользователем полученной информации. Поэтому JavaScript является единственным пассивным способом сбора информации в интересах дальнейшей идентификации пользователя.

Результаты сравнительного анализа методов присвоения уникального идентификатора позволили сформировать классификационную схему параметров, участвующих в формировании цифрового идентификатора веб-обозревателя на основе JavaScript.Исследование которой, в свою очередь, показало, что приоритетным аспектом для формирования цифрового идентификатора веб-обозревателя на основе JavaScriptпри анализе работы оборудования персонального компьютераявляется использование технологии вывода примитивов на CanvasFingerprint. Это обусловлено тем обстоятельством, что позволяет

собирать параметры аппаратной части устройства, которая в свою очередь редко подвергается замене, и как следствие, снижается вероятность изменения цифрового и пользователя [4].

В основе формирования цифрового идентификатора лежит сбор множества параметров $A\langle i\rangle$ таких, что

$$A\langle i\rangle = \{a_1,\ a_2,\ a_3,\ \dots\ a_i\},\ i = (1,\ j),$$

где i – множество параметров, участвующих в формировании Fingerprint. Далее, путём сериализации, происходит преобразование множества параметров в множество строк:

$$S\langle i\rangle = \{s_1,\ s_2,\ s_3,\ \dots\ s_i\},\ i = (1,\ j).$$

Затем в процессе конкатенации образуется итоговая строка $S_{\text{итог}}$, эквивалентная кортежу $K\langle |S_{\text{итог}}| \rangle$, из которого используя хеш-функцию типа Locality-sensitive hashing (LSH), получаем искомый цифровой идентификатор в виде хеша. Алгоритм LSH реализуется путем генерирования множества кортежей размерностью *r*, размер которых равен весу кортежа $S\langle m\rangle$:

$$C = \{c_1,\ c_2,\ c_3,\dots,c\ |S_{\text{итог}}\},\ i = (1,\ S_{\text{итог}}),$$

После генерации множества кортежей, выполняется скалярное произведение кортежа $K\langle |S_{\text{итог}}| \rangle$ с каждым кортежем множества *C*.

Хеш формируется по правилу: если скалярное произведение больше нуля, то в соответствующий бит хеша записывается единица, иначе – ноль.

Таким образом наиболее совершенной технологией присвоения Fingerprint является применение Canvas HTML5.

Canvas является средством программирования, которое позволяет рисовать простейшие графические примитивы - линии, фигуры, текст, и создавать различные эффекты мультимедиа, такие как игры, картографические инструменты и динамические графики, музыкально-световые представления и эмуляторы физических процессов [5]. Отрисовка изображений в Canvas происходит благодаря наличию веб-обозревателе поддержки кроссплатформенного API для работы с графикой - WebGL.

Так как пользователь в редких случаях прибегает к смене операционной системы или графического устройства по причине сложности подбора параметров их конфигурации с другими аппаратными или программными компонентами, а в случае использования мобильного терминала и вовсе невозможности. Таким образом использование Canvas HTML5 становится лучшим методом для формирования Fingerprint, поскольку используются аппаратные и программные параметры устройства пользователя, по которым создается итоговый хеш. Следовательно, применение прокси и VPN является нецелесообразным.

Список литературы:

1. Пилькевич С.В., Кузьмичев В.А. Подход к разработке модели распространения информации в социальных медиа ресурсах. // Методы и технические средства обеспечения безопасности информации: Материалы 30-й научно-технической конференции 22–25 сентября 2021 года. СПб.: Изд-во Политехн. ун-та, 2021. – С. 59-61.

2. Проблемы национальной безопасности России в интернет-пространстве.–URL: https://cyberleninka.ru/article/v/problemy-natsionalnoy-bezopasnosti-rossii-v-internet-prostranstve (дата обращения: 08.03.2022).

3. ГОСТ Р ИСО/МЭК 19795-1-2007. Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Ч. 1. Принципы и структура. – URL: http://docs.cntd.ru/document/1200067413 (дата обращения: 12.05.2022)

4. Фингерпринтинг конкретного ПК с точностью 99,24%. – URL: https://geektimes.com/post/284604/ (дата обращения: 13.05.2022).

5. How does Canvas Fingerprinting work – FingerprintJS. URL: – https://Fingerprintjs.com/blog/canvas-fingerprinting/(дата обращения: 10.04.2022).

4. Интеллектуальные методы анализа безопасности программного и аппаратного обеспечения

Падарян В.А., Тихонов А.Ю.
Институт системного программирования им. В.П. Иванникова Российской академии наук

ОТЕЧЕСТВЕННОЕ БЕЗОПАСНОЕ ПО: АКТУАЛЬНЫЕ ПРОБЛЕМЫ И ПУТИ РЕШЕНИЯ

Сложность ПО непрерывно возрастает в силу ряда факторов, взаимно усиливающих друг друга: эскалации размеров ПО, сложности систем сборки, усложнения внутренней структуры и связей с внешним окружением. Сложность современного ПО не позволяет гарантировать отсутствия в нем программных дефектов. Эксплуатация критических дефектов реализует наиболее разрушительные угрозы: несанкционированный доступ к данным и их порчу, потерю работоспособности или деструктивные воздействия на физические объекты.

Значительный вклад в сложность ПО вносят средства разработки и сборки. Крупные программные системы нередко собираются распределенными сборочными конвейерами. Средства компиляции в результате агрессивных оптимизаций способны генерировать исполняемый код, поведение которого среднестатистический разработчик изначально не предполагал. Применение интерпретируемых языков приводит к появлению дополнительного слоя программного обеспечения (виртуальной машины, библиотек времени выполнения, стандартных библиотек), который с одной стороны может повысить безопасность, а с другой – усложнит выявление привнесенных ошибок.

Усложнились внешние связи, произошел переход от компьютерных сетей к распределенным вычислениям, обработке и хранению данных. Компьютерные систем, нуждающиеся в масштабировании, строятся поверх облачных платформ, в состав компьютерных систем включаются мобильные устройства и устройства Интернета вещей. Сокращаются локально изолированные системы. Необходимость добавления новых, актуальных возможностей приводит к открытости систем, расширению поверхности атаки – тех интерфейсов, через которые прямо или опосредовано реализуются атаки на компьютерные системы.

Вследствие перечисленных факторов произошло существенное усложнение анализа. Фактически не существует средств полного контроля за поведением современной компьютерной системы. Ограничиваясь только «защитой по периметру», организационными мерами и навесной защитой, у компьютерных систем невозможно достичь требуемого уровня доверия.

В условиях нового уровня угроз, обеспечить безопасность отечественного ПО возможно лишь всесторонним комплексом мер, таких как развитие технологий анализа программного кода, учебных программ, актуализация требований ведомств-регуляторов и национальных стандартов, создание центров компетенций.

Как реакция на качественные изменения в области разработки ПО и компьютерной безопасности, в 2018 году принято решение Президиума РАН о необходимости создания нового научного направления «Анализ, трансформация программ и кибербезопасность». Для

исполнения этого решения в 2021 году приказом Минобрнауки №118 [1] утверждена новая научная специальность ВАК «Кибербезопасность», нацеленная на развитие представленных новых научных направлений. Специальность концентрируется на вопросах анализа и трансформации различных представлений программ и проектной документации программно-аппаратных комплексов и не включает исследования в области криптографии, алгоритмов и методов криптографической защиты информации, управления рисками безопасности информации, а также общих методов и систем защиты информации и информационной безопасности.

В 2021 году запущен проект «Технологический центр исследования безопасности операционных систем, созданных на базе ядра Linux». Главная задача центра – повышение уровня безопасности отечественных Linux-систем. В рамках работы Центра запланирован ряд взаимосвязанных активностей, в числе которых:

- применение лучших практик разработки безопасного ПО (статический анализ исходного кода, фаззинг-тестирование, полносистемный динамический анализ помеченных данных и другие);
- разработка исправлений, устраняющих ошибки и уязвимости в ядре Linux;
- наполнение БДУ ФСТЭК России сведениями об уязвимостях ядра Linux.

В работу по формированию требований к результатам деятельности Технологического центра вовлечены эксперты отечественных компаний, таких как «Алладин Р.Д.», «Базальт СПО», «Байкал электроникс», ИВК, «ИнфоТекс», «Код безопасности», «МЦСТ», «НТЦ ИТ РОСА», «НТЦ «Модуль», «Открытая мобильная платформа», «РусБИТех-Астра», «ЯНДЕКС.ОБЛАКО».

В 2022 году начатые работы получили развитие. Были запущены проекты по анализу критических компонент и созданию унифицированной среды разработки безопасного ПО. Модель Linux-центра применена для налаживания систематического анализа базового системного ПО: сред выполнения интерпретируемых языков (nodejs, LUA, .Net), сетевых сервисов (Nginx, suricata, squid), фреймворков разработки (Qt) и др. совместными силами специалистов ИСП РАН и представителей индустрии. Начата работа по созданию облачного решения – среды разработки безопасного ПО доступной широкому кругу отечественных разработчиков ПО. Среда предоставляет возможность размещать в ее инфраструктуре разрабатываемое ПО и обеспечивать применение инструментов анализа кода с требуемой регулярностью. Еще одним важным результатом создания унифицированной среды должен стать пополняемый набор тестов, который позволит объективно сравнить различные инструменты статического и динамического анализа по таким характеристикам, как доли пропусков и ложных срабатываний, скорость работы.

Список литературы:

1. Приказ Министерства науки и высшего образования РФ от 24 февраля 2021 г. № 118 "Об утверждении номенклатуры научных специальностей, по которым присуждаются ученые степени, и внесении изменения в Положение о совете по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, утвержденное приказом Министерства образования и науки Российской Федерации от 10 ноября 2017 г. № 1093". http://publication.pravo.gov.ru/Document/View/0001202104060043

Козачок А. В., Николаев Д.А., Ерохина Н. С.
Академия ФСО России, г. Орел

О НЕКОТОРЫХ ПОДХОДАХ К ОЦЕНКЕ ПОВЕРХНОСТИ АТАКИ И ФАЗЗИНГУ ВЕБ-БРАУЗЕРОВ

По статистике веб-браузер является одним из основных инструментов доставки вредоносных программ на компьютеры пользователей. Статический и динамический анализ кода являются наиболее эффективными методами обнаружения уязвимостей в программном обеспечении (ПО). В рамках проведения динамического анализа особое место занимает проведение фаззинг-тестирования [1].

В последние годы наблюдается тенденция внедрения фаззинга в процессы разработки и становится стандартом де-факто для такого большого и сложного программного обеспечения, как веб-браузеры и их ключевых компонент – графических и JavaScript движков. Современные веб-браузеры подвергаются обширному фаззингу для обнаружения уязвимостей.

Подходы к фаззингу веб-браузеров определяются следующими особенностями:

– веб-браузер имеет большую кодовую базу и сложную внутреннюю структуру, а также множество дополнительных модулей;
– входными данными веб-браузеров в основном являются сложноструктурированные текстовые данные;
– существует множество различных веб-браузеров, имеющих значительные архитектурные особенности;
– не существует универсального инструментария фаззинга, совместимого со всеми веб-браузерами.

Поиск дефектов в движках JavaScript является важной задачей, учитывая широкий перечень приложений, которые могут быть затронуты этими ошибками. Также это является сложной задачей для тестировщика, так как спецификации изначально неполны для обеспечения гибкости разработки.

Основная сложность при фаззинге движков JavaScript состоит в генерации синтаксически и семантически корректных входных данных, чтобы можно было исследовать различные функциональные возможности. Однако из-за динамической природы JavaScript и особенностей различных движков это сделать довольно сложно. Еще одна проблема заключается в том, что существующие фаззеры не могут генерировать вызовы новых методов, которые не включены в первоначальный исходный корпус или заранее определенные правила, что ограничивает возможности поиска ошибок [2].

Важной задачей является определение поверхности атаки (ПА) для проведения исследований веб-браузера. Так как кодовая база очень обширна и в документации на браузеры зачастую не выделены интерфейсы функций безопасности, данная задача является нетривиальной. Под ПА понимается совокупность интерфейсов и реализующих их модулей ПО, через которые могут реализовываться угрозы безопасному функционированию ПО.

В качестве тестируемого ПО был выбран веб-браузер FireFox и были рассмотрены следующие тесты:

1) запуск FireFox и его завершение без осуществления каких-либо действий;

2) запуск FireFox, открытие веб-страницы http://ifconfig.me/ip (без активного контента), завершение FireFox.

Целью являлось определение кода, вызываемого при рендеринге веб-страницы и последующая модификация кода для выявления аварийного завершения и подтверждения его достижимости.

На первом этапе была осуществлена компиляция приложения с возможностью получения покрытия кода, были выполнены тесты, для сформированных отчетов рассчитана разность покрытия кода и сгенерированы отчеты по покрытию.

На рисунке 1 представлен пример применения разработанного скрипта для расчета разности покрытий: содержание трех фалов a.json, b.json, c.json с указанием названия и количеством вызовов функций и строк кода. Файл c.json является результатом расчета разности покрытия для a.json и b.json.

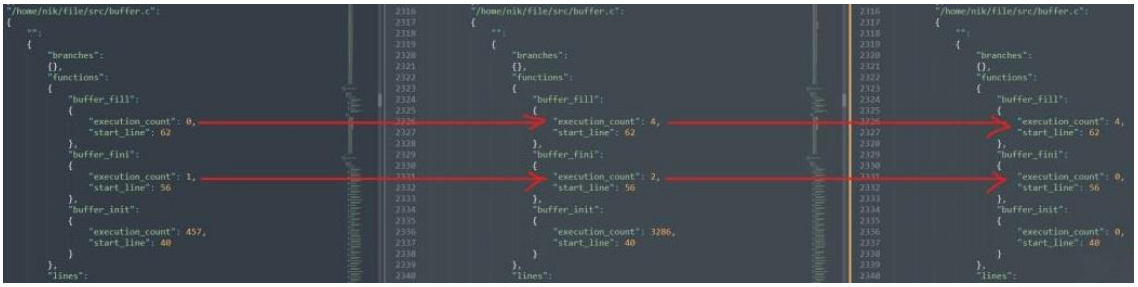

Рис. 1. Пример расчет разности покрытия кода

На втором этапе из сформированной поверхности атаки была выбрана функция *NewLinkAttributes()* файла *nsHtml5PlainTextUtils.cpp*, вызываемая при открытии страницы с *plaintext.css.* Далее в код был интегрирован вызов функции *abort()* и осуществлена сборка приложения.

На третьем этапе выполнен второй тест, в результате чего произошло аварийное завершение вкладки веб-браузера.

Преимущество предложенного подхода состоит в простоте его реализации при наличии исходного кода анализируемого ПО. Он также позволяет сформировать перечень потенциальных целей для фаззинга, однако у такого подхода есть недостатки:

– он не позволяет анализировать связь параметров функций с входными параметрами программы;

– он требует подготовки набора входных данных программы, обеспечивающих выполнение специфичных блоков кода.

Указанные недостатки частично устраняются подходом на основе отслеживания помеченных данных, позволяющим определять функции, обращающиеся к помеченным данным, тем самым формируя поверхность атаки [3] – инструмент Natch, разработанный ИСП им. В.П. Иванникова РАН.

Список литературы:

1. Козачок, А. В., Козачок, В. И., Осипова, Н. С., Пономарев, Д. В. Обзор исследований по применению методов машинного обучения для повышения эффективности фаззинг-тестирования // Вестник ВГУ. Серия: Системный анализ и информационные технологии, 2021 (4), С. 83-106, DOI: 10.17308/sait.2021.4/3800.

2. Tian Y., Qin X., Gan S. Research on Fuzzing Technology for JavaScript Engines // The 5th International Conference on Computer Science and Application Engineering, 2021, pp. 1-7, DOI: 10.1145/3487075.3487107.

3. Климушенкова М.А., Бакулин М.Г., Падарян В.А., Довгалюк П.М., Фурсова Н.И., Васильев И.А. О некоторых ограничениях полносистемного анализа помеченных

данных // Труды Института системного программирования РАН, 2016, 28(6), 11-26, https://doi.org/10.15514/ISPRAS-2016-28(6)-1.

Вишняков А.В., Кобрин И.А., Федотов А.Н.
Институт системного программирования им. В.П. Иванникова Российской академии наук

СИМВОЛЬНЫЕ ПРЕДИКАТЫ БЕЗОПАСНОСТИ В ГИБРИДНОМ ФАЗЗИНГЕ

Современное программное обеспечение стремительно развивается. Кодовая база продуктов постоянно увеличивается. Однако новый код неизбежно приносит с собой ошибки и уязвимости. В настоящее время распространен подход для предупреждения программных дефектов непосредственно во время процесса разработки программного обеспечения – безопасный цикл разработки ПО (SDL), который становится стандартом индустрии. Следуя стандартам SDL, разработчики обязаны применять различные инструменты анализа кода для повышения качества их продукта и защиты от злонамеренных атак. Таким образом, многие ошибки обнаруживаются во время разработки до того, как программа была внедрена в реальный мир.

Во время безопасного цикла разработки ПО непрерывно применяется фаззинг-тестирование [1, 2]. Фаззинг – это динамический метод анализа программы за счет порождения новых входных данных. Фаззер мутирует входные данные программы и наблюдает за ее выполнением. Таким образом, могут быть получены входные данные программы, на которых она зависает или аварийно завершается. Повсеместно распространен фаззинг с обратной связью по покрытию [1, 2]. При таком фаззинге не только осуществляется наблюдение за результатом выполнения программы, но и собирается информация о покрытом коде. Для организации обратной связи используются генетические алгоритмы. Наиболее «приспособленными» считаются входные файлы, открывающие как можно больше нового кода. Входные данные «скрещиваются» между собой и оставляются наиболее «приспособленные» с точки зрения покрытия.

Более продвинутые методы гибридного фаззинга [3-5] помимо обратной связи по покрытию извлекают пользу из семантики программы, которую учитывают методы динамической символьной интерпретации (DSE). DSE позволяет обнаруживать сложные состояния программы, труднодоступные обычному фаззингу. Для этого строится математическая модель программы, которая позволяет учитывать ее семантику при генерации новых входных данных. Гибридный фаззинг, благодаря динамической символьной интерпретации, решает две задачи: (1) генерация новых входных данных для расширения тестового покрытия программы и (2) обнаружение ошибок. Процесс фаззинга может выглядеть следующим образом. Динамическая символьная интерпретация позволяет исследовать новые состояния программы благодаря инвертированию условных переходов, встреченных на пути выполнения. Более того, DSE позволяет накладывать дополнительные ограничения (предикаты безопасности) для генерации новых входных данных, активирующих дефекты.

Предикаты безопасности – это дополнительные условия на предикат пути, которые позволяют обнаружить ошибки в программе. Мы будем искать ошибки работы с памятью и неопределенное поведение: выход за границу массива, целочисленное переполнение, деление на нуль и другие [6]. Для построения предиката безопасности во время символьной интерпретации программы анализируются определенные ее инструкции и места, которые считаются опасными. Конъюнкция предиката безопасности и предиката пути передается на вход SMT-решателю [7]. Решение такой системы уравнений представляет собой набор входных данных, которые обеспечивают как прохождение потока управления по

исследуемому пути, так и проявление конкретной программной ошибки. Таким образом, запуск программы на этих входных данных воспроизводит найденные ошибки.

Для применения символьных предикатов безопасности был разработан следующий метод автоматизированного поиска ошибок. Сначала производится гибридный фаззинг выбранного проекта, при котором совместно работают фаззер libFuzzer [1] и инструмент динамической символьной интерпретации Sydr [8]. После этого производится минимизация корпуса входных данных, полученных в результате фаззинга. Затем идет этап проверки предикатов безопасности на всех файлах из минимизированного корпуса. Срабатывания предикатов безопасности верифицируются с помощью исполняемого файла, собранного с санитайзерами. Верифицированные результаты оцениваются человеком на критичность и корректность.

В результате применения предложенного метода в 5 проектах с открытым исходным кодом было обнаружено 11 новых ошибок: 1 деление на нуль, 1 выход за границу массива, 7 целочисленных переполнений, 1 целочисленное переполнение, повлекшее за собой аварийное завершение программы, и 1 целочисленное переполнение, приводящее к переполнению буфера на куче.

Список литературы:

1. K. Serebryany. Continuous fuzzing with libFuzzer and AddressSanitizer. 2016 IEEE Cybersecurity Development (SecDev), page 157. IEEE, 2016.

2. A. Fioraldi, D. Maier, H. Eißfeldt, and M. Heuse. AFL++: combining incremental steps of fuzzing research. 14th USENIX Workshop on Offensive Technologies (WOOT 20), 2020.

3. S. Poeplau and A. Francillon. Symbolic execution with SymCC: don’t interpret, compile! 29th USENIX Security Symposium (USENIX Security 20), pages 181–198, 2020.

4. S. Poeplau and A. Francillon. SymQEMU: compilation-based symbolic execution for binaries. Proceedings of the 2021 Network and Distributed System Security Symposium, 2021.

5. L. Borzacchiello, E. Coppa, and C. Demetrescu. FUZZOLIC: mixing fuzzing and concolic execution. Computers & Security, 108:102368, 2021.

6. A. Vishnyakov, V. Logunova, E. Kobrin, D. Kuts, D. Parygina, A. Fedotov. Symbolic security predicates: hunt program weaknesses. 2021 Ivannikov ISPRAS Open Conference (ISPRAS), pages 76–85. IEEE, 2021.

7. A. Niemetz and M. Preiner. Bitwuzla at the SMT-COMP 2020. arXiv preprint arXiv:2006.01621, 2020.

8. A. Vishnyakov, A. Fedotov, D. Kuts, A. Novikov, D. Parygina, E. Kobrin, V. Logunova, P. Belecky, and S. Kurmangaleev. Sydr: cutting edge dynamic symbolic execution. 2020 Ivannikov ISPRAS Open Conference (ISPRAS), pages 46–54. IEEE, 2020.

УДК 004.056

Кубрин Г.С., Зегжда Д.П.
Санкт-Петербургский политехнический университет Петра Великого

ПОИСК УЯЗВИМОСТЕЙ НА ОСНОВЕ ПРИМЕНЕНИЯ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ К ГРАФОВОМУ ПРЕДСТАВЛЕНИЮ КОДА

В связи с внедрением программного обеспечения в процессы управления критической инфраструктурой актуальной является задача анализа ПО на наличие уязвимостей, использование которых злоумышленниками может привести к серьёзным финансовым убыткам. На протяжение нескольких последних лет наблюдается рост числа выявляемых уязвимостей в ПО. По данным базы уязвимостей cvedetails только за 2021 год было зарегистрировано 20141 новых CVE записей [1].

В последние годы получили развития модели нейронных сетей, предназначенные для обработки строго структурированных данных, в том числе графовые модели, такие как GGNN [2]. Существующие исследования в области применения подобных моделей нейронных сетей к задаче поиска уязвимостей делают акцент на поиск ошибок некорректной работы с памятью (бинарных уязвимостей) [3-5]. При этом по статистике с портала cvedetails, значительная часть критических уязвимостей (с оценкой CVSS от 7) связана с логическими ошибками (уязвимости инъекции команд, создания произвольных файлов и другие). Существующие подходы поиска уязвимостей данного типа в большинстве случаев опираются на наличие исходного кода и опираются на использование ограниченного множества правил [6, 7].

В данной работе предлагается модель описания сложных уязвимостей, вызванных наличием двух и более компонентов (примитивов). На базе предложенной модели выделяется набор подзадач по поиску сложных логических уязвимостей. Для выделенных подзадач выполняется сравнительный анализ эффективности применения методов на базе глубоких нейронных сетей с альтернативными методами анализа кода (points-to анализ, символьное исполнение и т.д.). Предлагается метод обнаружения примитивов логических уязвимостей в языках программирования с промежуточным представлением кода за счёт применения глубоких графовых нейронных сетей к модели кода, учитывающей зависимости по управлению и по передаче данных. Для ПО, написанного на языках программирования с промежуточным представлением (C#, Java, PHP) описывается алгоритм построения графового представления для применения метода обнаружения примитив уязвимостей.

Предложенный метод был применён к задаче поиска уязвимостей в ПО на языке программирования с промежуточным представлением – PHP. Подход к поиску сложных уязвимостей на базе примитивов работы с файловой системой был применён к ПО класса веб-серверов и позволил выявить уязвимость семейства XSS в одной из тестовых программ.

Список литературы:

1. Список уязвимостей, обнаруженных в 2021 году [Электронный ресурс]. URL: https://www.cvedetails.com/vulnerability-list.php?page=1 &year=2021.

2. Li Y. et al. Gated graph sequence neural networks //arXiv preprint arXiv:1511.05493. – 2015.

3. Demidov R., Pechenkin A. Application of Siamese Neural Networks for Fast Vulnerability Detection in MIPS Executable Code //Proceedings of the Future Technologies Conference. – Springer, Cham, 2019. – С. 454-466.

4. Zhou Y. et al. Devign: Effective vulnerability identification by learning comprehensive program semantics via graph neural networks //Advances in neural information processing systems. – 2019. – Т. 32.

5. Cheng X. et al. DeepWukong: Statically detecting software vulnerabilities using deep graph neural network //ACM Transactions on Software Engineering and Methodology (TOSEM). – 2021. – Т. 30. – №. 3. – С. 1-33.

6. Patil S. S. Automated Vulnerability Detection in Java Source Code using J-CPG and Graph Neural Network : дис. – University of Twente, 2021.

7. Wang X. et al. A Machine Learning Approach to Classify Security Patches into Vulnerability Types //2020 IEEE Conference on Communications and Network Security (CNS). – IEEE, 2020. – С. 1-9.

Куракин А.С.
ООО «Специальный Технологический Центр», Санкт-Петербург

ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ ГРУППОЙ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Одним из наглядных примеров решения задачи интеллектуального управления группой беспилотных летательных аппаратов (ГБЛА) является технология swarm («рой»). «Рой», как правило, представляет собой объединение большого количества миниатюрных однотипных БЛА. Лидерами в использовании данной технологии являются американское Управление перспективных научно-исследовательских проектов в сфере обороны *DARPA* (*Defense Advanced Research Project Agency*), Управление военно-морских исследований *ONR* (*Office of Naval Research*) Минобороны США и китайская компания «Норинко» (*China North Industries Corp., Norinco*).

В свою очередь, отечественные роботизированные комплексы в большинстве случаев оснащены БЛА, отличающимися по размеру, типу поставленных к выполнению задач и соответственно установленным на них полезным нагрузкам. Данное обстоятельство делает классическую технологию «роя» сложно применимой для отечественных ГБЛА. Сложность применения обусловлена тем, что БЛА в группе разнородны и выполняют разные функции, в том числе один БЛА из группы может выполнять несколько ролей одновременно в зависимости от установленных на него типов нагрузок. В данном случае совокупность ролей, аналогично ролям операторов наземного пункта управления, формирует «виртуальный экипаж», разграничение прав доступа внутри которого не должно препятствовать выполнению задач предназначения. Разграничение прав доступа между членами «виртуального экипажа» должно строиться с учетом современных принципов нейросетевых систем, обеспечивая безусловное выполнение поставленных задач и минимизацию времени на принятие решения, в том числе с учетом перехода на новый сценарий действий БЛА.

Также является актуальной проблема оперативного и автоматического перераспределения ролей внутри ГБЛА, так как потеря БЛА, оснащенного необходимой для выполнения задачи нагрузкой, может привести к угрозе срыва выполнения единого полетного задания всей ГБЛА. Вероятность возникновения сложно прогнозируемых ситуаций, а также большое количество возможных сценариев делают трудновыполнимой разработку полного множества вариантов развития событий, которое могло бы обеспечить все БЛА заготовленными инструкциями на каждый из этих вариантов. Исходя из этого, целесообразно

говорить о необходимости наличия искусственного интеллекта у членов «виртуального экипажа» для обеспечения принятия решения в автоматическом режиме работы БЛА.

Таким образом, возникает необходимость разработки механизма трансформации технологии «роя» в технологию «виртуального экипажа». При этом выработка решающих правил построения «виртуального экипажа» является задачей многофакторного прогнозирования разноуровневого взаимодействия ролей «виртуального экипажа» в составе группировки БЛА. Необходимость соблюдения требований по разграничению прав доступа между ними еще больше усложняет поиск решения. Ввиду того, что процесс построения «виртуального экипажа» требует обработки и создания новых данных, то целесообразно воспользоваться принципами нейронных сетей для получения полного набора решающих правил, управляющих «виртуальным экипажем».

На основе разработанной нейросетевой модели взаимодействия БЛА (рисунок 1) сформулируем следующие требования построения группировки БЛА:

- ГБЛА формируется с целью выполнения единого полетного задания;
- единое полетное задание ГБЛА является совокупностью частных полетных заданий каждого БЛА группы;
- требования к членам ГБЛА определяются целями и задачами единого полетного задания;
- количество ролей внутри ГБЛА должно обеспечивать выполнение целей и задач единого полетного задания;
- комплект полезных нагрузок определяется ролью БЛА, то есть частным полетным заданием;
- набор сценариев, определяющих действие БЛА в тех или иных обстоятельствах, в том числе учитывая роль, определенную частным полетным заданием, устанавливается перед запуском БЛА.

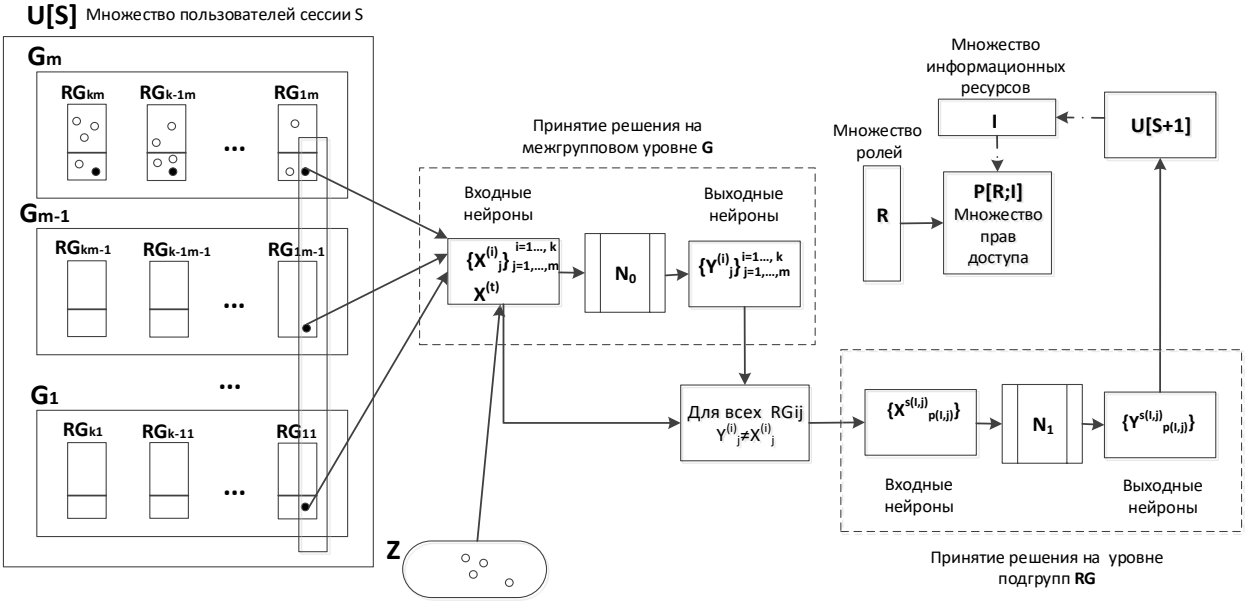

Рисунок 1 – Модель нейросетевого взаимодействия ГБЛА

Таким образом, интеллектуальное управление ГБЛА включает в себя две стадии:

1. Предполетная стадия – формирование (конфигурирование) состава ГБЛА и комплектов полезных нагрузок для БЛА, исходя из целей и задач, установленных единым полетным заданием, а также определение сил и средств, необходимых для их выполнения.

2. Полетная стадия – формирование и динамическое изменение (смена сценариев) «виртуального экипажа», выполняющего свои роли в рамках единого полетного задания ГБЛА.

Макаров А.С.
Санкт-Петербургский политехнический университет Петра Великого

ЗАЩИТА ВСТРАИВАЕМЫХ СИСТЕМ ОТ УГРОЗ БЕЗОПАСНОСТИ НА ОСНОВЕ ПОВЕДЕНЧЕСКОГО АНАЛИЗА АППАРАТНЫХ КОМПОНЕНТОВ

Информационные технологии давно используют в критичных областях человеческой деятельности, в которых цена сбоя очень велика. Вычислительные системы применяются в медицине, энергетике, логистике, промышленности, в системах контроля доступа и в других современных отраслях, требующих автоматизацию. Множество датчиков и систем управления находятся в одной сети, пересылают информацию и принимают разного рода решения, а также могут иметь доступ к сети интернет. Успешная атака на подобные системы может предоставить злоумышленнику полный контроль над киберфизической системой, а его действия могут привести к серьёзным последствиям. Поэтому необходимо уделять особое внимание к обеспечению безопасности каждого устройства.

Наряду с увеличением масштабов автоматизированных систем увеличивается и сложность их компонентов. Вводятся в эксплуатацию новые стеки протоколов либо расширяются уже имеющиеся, внедряются новые алгоритмы и механизмы взаимодействия между узлами системы. Узлом системы как правило является микроконтроллерное либо микропроцессорное устройство. Безопасность встраиваемых систем, входящих в информационную систему, определяет безопасность самой информационной системы.

Одной из проблем, приводящей к выходу из строя устройств встраиваемых систем, является низкая киберустойчивость устройства [1]. А именно свойство устройства, позволяющее ему существовать в условиях непрерывных, постоянных атак. При построении встраиваемых систем осуществляется разработка электрической схемы, связывающей микропроцессор или микроконтроллер с интерфейсными микросхемами, микросхемами обеспечения питания и микросхемами системной логики. Центральным вычислительным компонентом на печатной плате устройства является микропроцессор или микроконтроллер, он осуществляет решение основной задачи управления, проводит обработку результатов измерений, выполняет возложенные на устройство алгоритмы. Атаки, направленные на устройства, приходятся на главный вычислительный узел этого устройства. Поэтому, для повышения киберустойчивости устройства, необходимо обеспечить безопасную среду для выполнения исполнительного кода вычислительного узла.

Для обеспечения безопасности низкоресурсных узлов можно использовать как программные подходы [2, 3], так и аппаратные [4]. Эти подходы выполняют функции защиты информации криптографическими методами, а также выявляют аномальное поведение устройства и принимают меры, которые минимизируют возможный ущерб.

В рамках работы предлагается метод повышения киберустойчивости встраиваемых систем на основе поведенческого анализа аппаратных компонентов с помощью интегральной микросхемы.

Список литературы:

1 Björck, F., Henkel, M., Stirna, J., Zdravkovic, J. (2015). Cyber Resilience – Fundamentals for a Definition. In: Rocha, A., Correia, A., Costanzo, S., Reis, L. (eds) New Contributions in Information

Systems and Technologies. Advances in Intelligent Systems and Computing, vol 353. Springer, Cham.

2 Ovasapyan T. D., Ivanov D. V. Security provision in wireless sensor networks on the basis of the trust model //Automatic Control and Computer Sciences. – 2018. – Т. 52. – №. 8. – С. 1042-1048.

3 Шкоркина Е.Н., Александрова Е.Б. Принципы реализации симметричных криптографических алгоритмов на малоресурсных устройствах. // Методы и технические средства защиты информации. -2020. № 29. -114-115 с.

4 Макаров А.С. Архитектура защиты микроконтроллера. //Проблемы информационной безопасности. Компьютерные системы. -2019. № 2. -94-99 с.

Шулепов А.А.[(1)], Новикова Е.С.[(2)]
[(1)]Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», [(2)]СПб ФИЦ РАН

ВЫЯВЛЕНИЕ АНОМАЛИЙ В ПОТОКАХ ДАННЫХ ОТ СЕНСОРНЫХ СЕТЕЙ МЕТОДАМИ ВИЗУАЛЬНОГО АНАЛИЗА

Обнаружение аномалий является распространенной аналитической задачей, цель которой - выявление редких случаев, отличающихся от типичных, составляющих большую часть набора данных. Сенсорные сети, являясь подмножеством кибер-физических систем, могут испытывать множество внешних воздействий, как естественных, так и злонамеренных, негативно влияющих на их функционирование, из-за чего инфраструктура и люди могут подвергнуться серьезным опасностям. Несмотря на то, что проблема обнаружения аномалий в кибер-физических объектах (КФО) является довольно хорошо изученной темой, до сих пор в этой области остается ряд практических и научных проблем. Одной из таких задач является выявление и исследование аномалий, особенно в ручном режиме, т.к. любая кибер-физическая система описывается сотнями разнородных параметров. В действительности, анализ аномалий в функционировании КФО во многих случаях проводится вручную, аналитики, эксперты и технические специалисты используют простые методы визуализации, такие как линейные и точечные графики в сочетании с заданием пороговых значений. Применение методов обнаружения аномалий на основе обучения с учителем ограничено из-за недостатка аннотированных наборов данных реальных КФО, одна из причин которого является тот факт, что такие наборы данных во многих случаях представляют собой данные с ограниченным доступом. Также такие подходы требуют для работы существенные вычислительные ресурсы. Существует еще одна проблема, связанная с частотой возникновения аномалий, поскольку аномалии случаются достаточно редко, а наборы данных от реальных КФО часто имеют большие объемы, формирование меток для дальнейшего анализа является трудоемкой задачей.

В настоящей работе предлагается подход, который позволяет значительно упросить процесс исследования многомерных временных рядов с целью выявления аномалий. Он отличается простотой по сравнению с методами обнаружения аномалий с учителем, не требует для работы заранее размеченных данных и больших вычислительных ресурсов. В его основе лежит предложенный авторами показатель Change_Measure, который характеризует, как изменяется поведение исследуемого объекта во времени. Он рассчитывается на множестве параметров, которые определяются аналитиком, и является, таким образом, интегральным показателем, который можно легко визуализировать с помощью стандартного линейного графика с временной осью.

Метод вычисления показателя Change_Measure основан на определении суммарной площади последовательности треугольников, сгенерированных алгоритмом триангуляции Делоне на множестве точек в двумерном пространстве, полученном посредством применения методов сокращения размерности, таких как PCA или t-SNE. Краеугольным камнем данного подхода является предположение о том, что нормальному поведению системы в проекции данных на двумерное пространство соответствует линейная и последовательная траектория расположению точек, в то же время аномальному поведение соответствует случайная траектория с сильным разбросом точек. Таким образом, при нормальном поведении системы площади треугольников, и, следовательно, значение метрики Change_Measure, будут стремиться к нулю. Построение графика метрики Change_Measure позволяет визуально оценить, в какие моменты времени система находилась в аномальном состоянии, либо же испытывала существенный переход из одного состояния в другое. В этим моменты времени на графике можно увидеть хорошо заметные пиковые значения, и именно они должны представлять интерес аналитика для дальнейшего, более детального анализа причин возникновения подобных отклонений.

Для проведения экспериментов и проверки предложенного подхода было разработано интерактивное программное средство для исследования и визуализации многомерных данных с возможностью автоматически выявлять аномалии. Программа позволяет пользователю выделить участок линейного графика показателя Change_Measure, который, по его мнению, является аномалией, и на основе выделенного участка графика построить паттерн аномалии и провести поиск подобного паттерна на остальной части графика. Предлагаемая визуализация показателя Change_Measure дает представление о функционировании системы, но не детализует ее, поэтому для полного понимания состояния предусмотрена возможность визуализации выбранного множества параметров. Перед проведением анализа пользователь может ограничить временной интервал и задать множество параметров для анализа, тем самым исключив из обработки ненужные, ошибочные или не представляющие интереса данные. В программе реализовано несколько методов предварительной обработки данных: минимаксная нормализация, нормализация средним, z-score нормализация, и два метода сокращения размерности: PCA и t-SNE, с последующим построением диаграмм рассеяния точек в двумерном пространстве. При расчете показателя пользователь может задать размер подмножества точек, на основе которого будет производиться триангуляция Делоне, тем самым изменяя уровень детализации анализа. В качестве постобработки полученных данных реализованы механизмы фильтрации данных на основе среднего или медианы с возможностью задания окна и шага фильтрации.

С помощью данного инструмента была проведена серия экспериментов на двух хорошо изученных наборах данных с известными аномалиями, один синтетический, другой реальный, которые описывали функционирование двух разных КФО. Эксперименты показали, что выдвинутое предположение о том, что нормальному поведению системы соответствует линейная и последовательная траектория расположению точек в двумерном пространстве после применения методов сокращения размерности, подтвердилось. Во многих случаях аномальные паттерны на графиках показателя Change_Measure были хорошо различимы визуально, в некоторых случаях потребовалась дополнительная постобработка методами фильтрации.

Таким образом, полученные результаты показывают, что предложенный авторами показатель Change_Measure и разработанный программный инструмент позволяют выполнять анализ многомерных данных, выявлять аномалии в данных, и таким образом могут быть полезны аналитикам данных, специалистам по безопасности объектов инфраструктуры и сенсорных сетей. Дальнейшая работа связана с развитием инструмента путем добавления возможности полуавтоматического аннотирования данных с последующим сохранением меток, подходящим для использования в алгоритмах машинного обучения с учителем.

Андрианов П.С.
Институт Системного Программирования РАН им. Иванникова, Москва

ПОИСК СОСТОЯНИЙ ГОНКИ В СИСТЕМНОМ ПРОГРАММНОМ ОБЕСПЕЧЕНИИ ПРИ ПОМОЩИ СТАТИЧЕСКОЙ ВЕРИФИКАЦИИ

Статическая верификация (англ. software model checking) занимает промежуточное положение между статическим анализом и формальной дедуктивной верификацией. Методы статической верификации еще могут применяться к программам в 10 и даже 100 тысяч строк кода, но при этом такие методы позволяют обнаруживать более сложные ошибки, чем обычные статические анализаторы. В определенных предположениях возможно даже доказательство корректности программы.

Ошибки, связанные с многопоточным выполнением программы, являются одними из наиболее сложных ошибок, как для обнаружения, так и для исправления. Для воспроизведения такой ошибки требуется специальная последовательность чередований инструкций в нескольких потоках, что затрудняет поиск таких ошибок динамическими методами анализа, например, ThreadSanitizer [1]. Простые статические анализаторы, например, на основе clang [2], проверяют корректность используемой синхронизации на основе аннотаций.

Системное программное обеспечение традиционно предъявляет повышенные требования к корректности программного кода, и затраты на верификацию кода могут в несколько раз превышать затраты на его создание. Это обусловлено тем, что последствия ошибок в системном программном обеспечении могут быть очень серьезными. В частности, состояния гонки могут приводить к небезопасному использованию памяти, например, двойному освобождению. Такие ошибки являются эксплуатируемыми, то есть, могут позволить злоумышленнику выполнить вредоносный код.

Платформа Klever [3] позволяет проводить автоматическую верификацию программ на языке Си на соответствие различным требованиям. Klever автоматизирует подготовку верификационных задач: выделяет фрагменты для отдельного анализа, проводит инструментацию, готовит модель окружения. Затем подготовленная задача передается инструменту статической верификации CPAchecker [4] для решения. Результаты верификации загружаются в Klever для последующего анализа экспертом. В процессе такого анализа может быть создана оценка с описанием предупреждения, которая будет сохранена. Существующие оценки применяются ко всем похожим предупреждениям на основе некоторого критерия сравнения, что позволяет сократить трудозатраты при проведении регулярной верификации.

Метод поиска состояний гонки реализован в инструменте CPALockator [5], который интегрирован в платформу CPAchecker. Этот метод основан на подходе с раздельным рассмотрением потоков (англ. thread-modular approach), что позволяет избежать комбинаторного взрыва числа состояний при анализе программ со множеством потоков. Кроме этого, такой подход успешно комбинируется с другими техниками анализа программ: предикатной абстракцией и уточнением абстракции по контрпримерам (англ. CEGAR [6]). Это позволяет эффективно применять данный метод для верификации даже кода ядра операционных систем, который содержит сотни тысяч строк кода и десятки потоков.

По исходному коду программы строится абстрактный граф достижимости (англ. Abstract Reachability Graph) инструментом CPALockator. Далее на этом графе можно решать

различные задачи, например, задачу достижимости или задачу поиска гонок. Чтобы обнаружить состояние гонки в программе, необходимо найти пару операций в различных потоках, которые производят доступ к одной разделяемой памяти и не используют никакую синхронизацию, при этом один из доступов должен быть записью. В статической верификации традиционной сложностью является определение равенства областей памяти и определение возможности одновременного доступа. Для решения первой проблемы CPALockator использует модель памяти на регионах. Для определения одновременности используется понятие совместности, которое задается каждым используемым анализом, что позволяет повысить точность и эффективность.

Основной сложностью при применении инструментов статической верификации к коду системного ПО являются подготовка модели потоков. В ядре ОС потоки не всегда создаются напрямую. Обычно модули ядра регистрируют свои обработчики, которые затем могут вызываться в этом или в другом потоке. Кроме того, существуют неявные зависимости между различными активностями. Например, если обработчик прерываний уже зарегистрирован, то он уже доступен для ядра ОС, и оно может вызывать этот обработчик. Но, в то же время, если устройство еще не активно, то оно не может генерировать прерывания, а значит, обработчик прерываний пока не может работать. Такие зависимости определяются экспертом вручную и находят отражение в модели потоков (модели окружения).

Платформа Klever и инструмент верификации CPALockator успешно применялись к модулям ядра ОС Linux и ядрам ОС РВ. С их помощью было найдено более 50 ошибок, связанных с состояниями гонки. Далеко не все признанные ошибки легко исправить, и зачастую ошибки в устаревшем коде не были исправлены. Тем не менее, более 20 патчей к ядру ОС Linux были приняты разработчиками. Инструмент может быть полезен разработчикам системного ПО, в том числе, в процессе сертификации кода.

Список литературы:

1. Serebryany, Kostya and Timur Iskhodzhanov. “ThreadSanitizer: data race detection in practice.” WBIA '09 (2009).

2. Hutchins, DeLesley S. et al. “C/C++ Thread Safety Analysis.” 2014 IEEE 14th International Working Conference on Source Code Analysis and Manipulation (2014): 41-46.

3. E. Novikov, I. Zakharov. Towards automated static verification of GNU C programs. In: Petrenko A., Voronkov A. (eds) Proceedings of the 11th International Andrei Ershov Memorial Conference on Perspectives of System Informatics (PSI’17), LNCS, volume 10742, pp. 402–416. Cham, Springer, 2018. https://doi.org/10.1007/978-3-319-74313-4_30.

4. Beyer, D., Keremoglu, M.E. (2011). CPACHECKER: A Tool for Configurable Software Verification. In: Gopalakrishnan, G., Qadeer, S. (eds) Computer Aided Verification. CAV 2011. Lecture Notes in Computer Science, vol 6806. Springer, Berlin, Heidelberg.

5. Andrianov, Pavel. “Analysis of Correct Synchronization of Operating System Components.” Programming and Computer Software 46 (2020): 712-730.

6. Clarke, E., Grumberg, O., Jha, S., Lu, Y., Veith, H.: Counterexample-guided abstraction

Лаврова Д.С.
Санкт-Петербургский политехнический университет Петра Великого

ПОДХОД К БЕЗОПАСНОЙ ИНТЕГРАЦИИ КОНЦЕПЦИИ INTERNET OF THINGS С МЕДИЦИНСКИМИ АППАРАТАМИ ИСКУССТВЕННОЙ ВЕНТИЛЯЦИИ ЛЕГКИХ

Пандемия COVID-19 навсегда изменила жизнь современного общества и вывела на новый уровень облачные технологии, концепцию Internet of Things (IoT) и техники искусственного интеллекта. Данная работа посвящена интеграции технологии IoT с современными медицинскими аппаратами искусственной вентиляции легких (ИВЛ).

Согласно проведенным исследованиям [1-3], ведущие мировые производители аппаратов ИВЛ уже признали важность и удобство технологии IoT для медицины. На сегодняшний день многие компании используют IoT для удаленного мониторинга состояний пациентов, а ряд компаний планируют продвигать свои продукты в сторону использования IoT и искусственного интеллекта. Однако внедрение новых технологий может быть замедлено в связи с рядом нерешенных проблем безопасности.

Актуальность создания подхода к безопасной интеграции концепции Internet of Things с медицинскими аппаратами ИВЛ также подчеркивается тем, что на данный момент на рынке выявлено только одно решение, которое позиционируется создателями как ИВЛ на базе концепции IoT – BiWaze ION от компании ABM Respiratory Care [36]. Согласно производителю, BiWaze ION использует новую технологию, которая в 25 раз быстрее и генерирует в 50 раз меньше трафика данных, чем традиционные веб-технологии. Это обеспечивает быструю, последовательную телеметрию и широкомасштабный безопасный доступ к аппаратам ИВЛ.

В работе рассмотрены следующие аспекты безопасной интеграции IoT с аппаратами ИВЛ:

1. Подключение аппарата ИВЛ к IoT. Выделено 3 ключевых подхода:

- подключение с помощью модуля Wi-Fi, встроенного в аппарат ИВЛ.
- использование внешнего адаптера, который позволяет переключиться с последовательного интерфейса на беспроводной (RS-232 → Wi-Fi).
- применение внешнего 3G/4G/5G модема, подключенного к последовательному интерфейсу.

2. Взаимодействие между аппаратом ИВЛ и устройствами IoT. Использовать сетевые IoT-протоколы способен только вышеуказанный BiWaze ION, остальные производители аппаратов ИВЛ не поддерживают такую возможность. Поэтому предлагается осуществлять косвенное взаимодействие между аппаратом ИВЛ и вспомогательными устройствами. IoT-устройство передает информацию через свой шлюз на Ops сервер, затем конечный пользователь (врач, "оператор") получает доступ к переданным данным через веб-сервер приложения и реагирует на них командой, отправленной на целевой аппарат ИВЛ.

3. Связь с Ops сервером и обработка медицинских данных. Подсистема анализа данных должна включать: Ops-сервер, хранилище данных и сервер веб-приложений, сочетающий в себе функциональность веб-сервера и сервера приложений. Веб-сервер необходим для организации удобного доступа врачей к медицинским данным через веб-браузер, работающий на стационарном компьютере, смартфоне и других цифровых устройствах. Сервер приложений предоставляет набор физических и программных средств, которые могут обеспечить клиентам доступ к программам, запущенным непосредственно на серверном оборудовании.

В работе систематизированы киберугрозы, характерные для систем медицинского IoT, к которым и относится описываемая концепция интеграции аппаратов ИВЛ с IoT. Предложен ряд мер по нейтрализации этих киберугроз:

1. Технические меры:

– выделение сервера Ops в DMZ и использование систем обнаружения и предотвращения атак для защиты сервера Ops;

– добавление резервного хранилище данных для репликации медицнских данных из основного хранилища;

– добавление модуля обнаружения аномалий;

– добавление связи с центром сигнализации для немедленного оповещения врача (по электронной почте, SMS или PUSH-уведомлением)

– использование надежного шифрования для защиты всех каналов передачи данных в инфраструктуре;

– использование двухфакторной аутентификации конечного пользователя (врача);

– для предотвращения ситуаций злонамеренного использования приложений дистанционного управления аппаратами ИВЛ возможно применение схемы с общим секретом, при которой решение может быть принято только при согласии нескольких участников схемы (например, выключение аппарата ИВЛ должно быть подтверждено главным врачом).

2. Организационные меры: крайне важно использовать шифрование, интегрировать только доверенное IoT-оборудование и проводить регулярные аудиты безопасности для выявления проблем безопасности и их устранения.

Список литературы:

1. How IoT revolutionized medical care during the pandemic. URL: https://thenextweb.com/science/2020/08/19/how-iot-revolutionized-medical-care-during-the-pandemic/.

2. IoT Medical Devices Market by Product (Blood Pressure Monitor, Glucometer, Cardiac Monitor, Pulse Oximeter, Infusion Pump), Type (Wearable, Implantable Device), Connectivity Technology (Bluetooth, Wifi), End User (Hospital) - Global Forecast to 2023. URL: https://www.marketsandmarkets.com/Market-Reports/iot-medical-device-market-15629287.html.

3. Ventilators Market by Mobility (Intensive Care, Portable), Type, Mode (Volume, Pressure, Combined), Interface (Invasive, Non-invasive), End-User (Hospital, Home Care, ACC, Emergency Medical Services) Covid-19 Impact - Global Forecast to 2025. URL: https://www.marketsandmarkets.com/Market-Reports/ventilators-market-11018337.html?gclid=CjwKCAjw8-78BRA0EiwAFUw8LPfX6d29pbW91m_su15s8gTSLnbx3ZrKVD72qNG0fPq6r4W0GwiiHRoCztwQAvD_BwE

4. ABM Respiratory Care creates world's first IOT-enabled Tele-Ventilator for COVID-19 Pandemic. URL: https://www.hhmglobal.com/industry-updates/press-releases/abm-respiratory-care-creates-worlds-first-iot-enabled-tele-ventilator-for-covid-19-pandemic.

Самарин Н.Н.
Федеральное государственное унитарное предприятие
Научно-исследовательский институт «Квант», г. Москва

ОБНАРУЖЕНИЕ ПРОГРАММНЫХ ДЕФЕКТОВ НА ОСНОВЕ ОБРАТНОГО СИМВОЛЬНОГО ВЫПОЛНЕНИЯ

Процесс разработки современного программного обеспечения (ПО) стал технически сложнее за счет задействования большого числа различных технология и функционирования в разнородной среде. Это привело к существенному увеличению объема исходных текстов – как на языках высокого уровня, так и, соответственно, в бинарном коде.

В связи с этим, увеличивается актуальность решения задачи эффективного анализа безопасности бинарных файлов такого объема, как вручную, так и с использованием специализированных программ-анализаторов. В условиях необходимости анализа таких больших и сложных программ становится невозможным рассмотрение всех возможных путей выполнения программы. Данный доклад посвящен теме разработки специальных эвристических методов, сокращающих перебор путей выполнения ПО таким образом, чтобы решение этой задачи стало реализуемо [1].

В работе предлагается модель обратного символьного выполнения (ОСВ), способная сократить количество ветвлений, возникающих при прямом символьном выполнении. Ее особенностью является использование символьных переменных при выполнении кода от результата его выполнения.

При использовании ОСВ объектом анализа выступает отдельно взятый путь из графа управления, цель – проверка достижимости. Первоначально методом статического анализа производится поиск потенциально уязвимых участков бинарного кода. Далее фиксируется путь выполнения p, после чего производится символьное выполнение заданного пути p в обратном порядке.

Изначально все параметры представляются как некоторые символьные переменные. Так как производится анализ бинарного файла, в качестве переменных выступают регистры и отдельно взятые участки памяти. Изначально на переменные накладываются некоторые ограничения, то есть условия при которых может сработать исключение в коде. Далее на каждом шаге анализа множество ограничений дополняется новыми в точках ветвления (команды условного перехода), а значения переменных преобразуются в соответствие с выполняемой командой, такую процедуру назовем уточнением.

Важной особенностью предложенного подхода является тот факт, что при обработке прямых ветвлений программы, система не дублируется с различными условиями, как происходит при прямом выполнении. В данном случае изначально существует две копии системы, пришедшие к этому условию с различных веток. То есть, данные копии системы изначально могли быть запущены параллельно с различными начальными условиями.

Как только система доходит до начала функции, производится синхронизация переменных из набора ограничений с локальными переменными на стеке. После чего производится финальное решение системы ограничений. В результате получается множество ограничений на входные параметры для конкретной функции.

Результаты экспериментальных исследований реализованной модели показали, что, наряду с достоинствами (сокращение числа ветвлений внутри кода, возможность дополнения своими наборами ограничений), есть ряд недостатков. Они связаны с обработкой необратимых операций (сложность обращения большого числа функций), сильным ветвлением при сортировке и неоднозначностью переменных (проблема возникает при обращении к одному участку памяти с использованием разных регистров).

Для устранения вышеописанных недостатков предложено использовать такой вариант оптимизации метода классического символьного выполнения, как зеркальное объединение. Оно позволяет сократить промежуточное множество ограничений за счет его объединения с результатами работы метода обратного символьного выполнения (на отдельно взятых участках кода_.

Оценка эффективности предложенной модели проводилась на реализованном экспериментальном макете с использованием набора исходных данных Software Assurance Reference Dataset [2] от NIST (National Institute of Standards and Technology). В качестве результатов тестирования оценивалось время работы классического алгоритма символьного выполнения и алгоритма оптимизированного с помощью метода обратного символьного выполнения по логике зеркального объединения на определенных ветвях.

По результатам тестирования, время работы алгоритма с предложенным методом оптимизации при обнаружении программных ошибок и уязвимостей, порожденных вычислительными ошибками или отсутствием инициализации переменных, в среднем сократилось на 2.1%, что говорит о практической применимости разработанной модели.

Список литературы:

1. G. Balakrishnan. T. Reps. "Fast library identification and recognition technology". URL: http://www.datarescue.com/idabase/flirt.htm.
2. NIST’s Software Assurance Reference Dataset. URL: https://samate.nist.gov/SRD/view.php.

5. Конференция «Неделя науки Института кибербезопасности и защиты информации СПбПУ».

Вопросы информационной безопасности: взгляд молодых учёных

Сабиров Э.Р.[(1)], Маршалко Г.Б.[(2)]
[(1)]ВМиК МГУим. М.В. Ломоносова, г. Москва
[(2)]Академия криптографии Российской Федерации, г. Москва

ИСПОЛЬЗОВАНИЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ СЕТЕЙ ДЛЯ ЗАЩИТЫ ИЗОБРАЖЕНИЙ ОТ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ

При публикации в интернет изображения могут классифицироваться и индексироваться поисковыми службами или специальными программами, что в ряде случаев может нести негативные последствия для безопасности частной жизни граждан.

Существует множество способов ограничить доступ к этой информации и обеспечить конфиденциальность этих данных, одним из которых является построение и публикация состязательных изображений [2]. Такие изображения визуально схожи с оригинальными, но некорректно классифицируются алгоритмами распознавания. Они могут быть получены с помощью генеративно-состязательных сетей (GAN) [1].

Генеративно-состязательная сеть — архитектура, состоящая из двух нейронных сетей -- генератора и дискриминатора, настроенных на работу друг против друга. Дискриминационные алгоритмы классифицируют входные данные, в то время как генеративные алгоритмы формируют образы таким образом, чтобы затруднить классификацию.

Функция потерь - функция, которая в теории статистических решений характеризует потери при неправильном принятии решений на основе наблюдаемых данных.

Цель генератора G заключается в генерации таких изображений, чтобы дискриминатор как можно хуже их различил (принял за настоящие) и может быть описана следующим выражением: $\min_G \log(1 - D(G(z, x)))$.

Цель дискриминатора D, математически описанная с помощью следующего выражения, противоположна — он должен отличить подделку с как можно более высокой вероятностью: $\max_D [logD\ (x) + log\ \left(1 - D\big(G(z, x)\big)\right)]$, где D — функция дискриминатора; G — функция генератора; z — случайный гауссовый шум; x-оригинальное изображение.

В данной работе был предложен следующий алгоритм обучения генеративно-состязательной сети:

1. На каждой итерации обучения выбираются k-реальных изображений $x_1, \dots, x_k$, выбираются случайным образом k-случайных гауссовских векторов $z_1, \dots, z_k$, где k - размер мини-батча (выборка k изображений из всей выборки).

2. Происходит построение модифицированных изображений $img_i = x_i + G(z_i, x_i)$.
3. Происходит вычисление $D(x_i)$ и $D(img_i), i = 1, \dots, k$.
4. Происходит обновление весов дискриминатора D_ω, через вычисление значения градиента согласно функционалу из [1]:

$$D_\omega \leftarrow \nabla_{\theta_D} \frac{1}{k} \sum_{i=1}^{k} [logD(x_i) + \log(1 - D(img_i))].$$

5. Происходит обновление весов генератора G_ω по вновь сгенерированным значениям шумов согласно функционалу из [1]:

$$G_\omega \leftarrow \nabla_{\theta_G} \frac{1}{k} \sum_{i=1}^{k} \log(1 - D(img_i)).$$

Экспериментальное исследование проводилось на наборе данных CelebA, он содержит более 200 тыс. изображений знаменитостей.

Рис. 2а

Рис. 1б

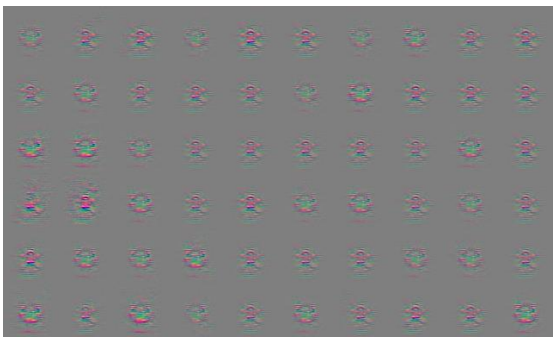

Рис. 1в

На рис.1 представлены оригинальные и модифицированные изображения, сгенерированный генератором шум. Можно отметить визуальное сходство оригинальных и модифицированных изображений.

Результаты работы дискриминатора до и после применения алгоритма GAN (в таблице указана доля корректно классифицированных изображений):

	Тренировочная выборка	Тестовая выборка
Оригинальный дискриминатор	-	0.969
Реализация GAN	0.046	0.051

Результаты работы показывают возможность использования генеративно-состязательных сетей для модификации изображений с целью их защиты от автоматической классификации. В дальнейшем представляется целесообразным модифицировать алгоритм за

счет добавления большего количества дискриминаторов с целью защиты от различных алгоритмов распознавания [3].

Список литературы:

1. Goodfellow I. J. [и др.]. Generative Adversarial Networks // arXiv e-prints. — 2014.

2. Goodfellow I. J., Shlens J., Szegedy C. Explaining and Harnessing Adversarial Examples // CoRR. — 2015.

3. Sabirov E., Marshalko G. Protecting the privacy of images being published // Mitsobi. — 2021.

Кобрин И.А., Вишняков А.В., Федотов А.Н.
ИСП РАН, Москва

ГИБРИДНЫЙ ФАЗЗИНГ ФРЕЙМВОРКА МАШИННОГО ОБУЧЕНИЯ TENSORFLOW

Системы искусственного интеллекта повсеместно внедряются в обыденную жизнь. В связи с этим повышается актуальность вопроса безопасности фреймворков машинного обучения, применяемых для проектирования таких систем. Данная работа раскрывает особенности применения фаззинга для поиска ошибок в фреймворке машинного обучения TensorFlow.

Одной из наиболее распространенных технологий поиска ошибок остается фаззинг с обратной связью по покрытию [1-2]. Гибридный фаззинг извлекает выгоду из динамической символьной интерпретации [3-5], открывая состояния программы, трудно досягаемые с помощью простого фаззинга. Фаззинг представляет собой метод автоматического тестирования программ, при котором программе на вход подаются входные данные, после чего анализируется реакция программы и генерируются новые входные данные. Целью фаззинга является поиск входных данных, которые приведут выполнение программы к аварийному завершению. Символьная интерпретация представляет собой метод автоматического тестирования, при котором происходит интерпретация программы, где конкретным значениям переменных, зависящих от входных данных, сопоставляются символьные переменные, принимающие произвольные значения. Анализируемый путь выполнения программы может быть описан предикатом пути – системой уравнений, зависящих от символьных переменных, решение которой обеспечивает прохождение потока управления по тому же пути. Предикатом безопасности [6] мы называем дополнительные условия на предикат пути, которые позволяют проявить ошибочную ситуацию в программе.

В репозитории фаззинга проектов с открытым исходным кодом Google OSS-Fuzz были добавлены очень примитивные цели для фаззинга, покрывающие маленькие единичные функции. В репозитории TensorFlow были найдены другие фаззинг-цели, которые вызывали более сложный и интересный код, но сборочные файлы для них в OSS-Fuzz отсутствовали. В директории с новыми целями лежали нерабочие файлы для их сборки, поэтому пришлось искать по истории изменений в OSS-Fuzz какие-то намеки на то, как собрать эти цели из core/kernels/fuzzing. Были найдены соответствующие файлы для сборки, однако они также не приводили к успешной сборке целей для фаззинга. В результате пришлось написать сборочные скрипты самостоятельно, которые совмещали в себе сборку целей из текущей версии OSS-Fuzz и сборку, найденную на старом коммите OSS-Fuzz. В изначальной версии репозитория TensorFlow эти цели собираются как динамическая библиотека, но это приводит к ошибкам линковки. Чтобы собрать эти цели, мы отключили динамическое связывание,

поменяв его на статическое, и для целей сделали сборку исполняемых файлов, а не только библиотек.

Нами была налажена сборка фаззинг-целей в TensorFlow. Цели для фаззинга были собраны в трех вариантах: с санитайзерами для libFuzzer [1], без инструментации для инструмента символьной интерпретации Sydr [7] и для просмотра покрытия llvm-cov. После чего был запущен гибридный фаззинг каждой цели. Далее полученный в результате фаззинга корпус был минимизирован, чтобы уменьшить число файлов для проверки на предикатах безопасности. Более того, минимизация корпуса позволила быстрее собирать покрытие по исходному коду TensorFlow, имеющего обширную собственную кодовую базу, а также большое число зависимостей. На минимизированном корпусе была запущена проверка предикатов безопасности (деление на нуль, выход за границу массива, целочисленное переполнение и др.). Далее срабатывания предикатов безопасности проверялись на версии исполняемого файла, собранного с санитайзерами. Подтвержденные на санитайзерах ошибки анализировались вручную.

Применение предикатов безопасности позволило обнаружить истинную ошибку целочисленного переполнения в коде TensorFlow. Однако ручной анализ данной ошибки показал, что найденное переполнение на практике не может привести к некорректному поведению программы. Разработанная система сборки фаззинг целей для TensorFlow была принята в проект Google OSS-Fuzz. Теперь CI фаззинг на облачном кластере Google покрывает больше кода из TensorFlow, который становится в результате более безопасным.

Список литературы:

1. K. Serebryany. Continuous fuzzing with libFuzzer and AddressSanitizer. 2016 IEEE Cybersecurity Development (SecDev), page 157. IEEE, 2016.

2. A. Fioraldi, D. Maier, H. Eißfeldt, and M. Heuse. AFL++: combining incremental steps of fuzzing research. 14th USENIX Workshop on Offensive Technologies (WOOT 20), 2020.

3. S. Poeplau and A. Francillon. Symbolic execution with SymCC: don’t interpret, compile! 29th USENIX Security Symposium (USENIX Security 20), pages 181–198, 2020.

4. S. Poeplau and A. Francillon. SymQEMU: compilation-based symbolic execution for binaries. Proceedings of the 2021 Network and Distributed System Security Symposium, 2021.

5. L. Borzacchiello, E. Coppa, and C. Demetrescu. FUZZOLIC: mixing fuzzing and concolic execution. Computers & Security, 108:102368, 2021.

6. A. Vishnyakov, V. Logunova, E. Kobrin, D. Kuts, D. Parygina, A. Fedotov. Symbolic security predicates: hunt program weaknesses. 2021 Ivannikov ISPRAS Open Conference (ISPRAS), pages 76–85. IEEE, 2021.

7. A. Vishnyakov, A. Fedotov, D. Kuts, A. Novikov, D. Parygina, E. Kobrin, V. Logunova, P. Belecky, and S. Kurmangaleev. Sydr: cutting edge dynamic symbolic execution. 2020 Ivannikov ISPRAS Open Conference (ISPRAS), pages 46–54. IEEE, 2020.

Рудницкая Е.А. Полтавцева М.А.

Санкт-Петербургский политехнический университет Петра Великого

МЕТОДЫ ЗАЩИТЫ ОТ АТАК НА СИСТЕМЫ МАШИННОГО ОБУЧЕНИЯ

Использование систем машинного обучения в последние годы стало практически повсеместным, это обусловлено необходимостью обработки большого количества данных.

Вместе с тем увеличилось количество атак, производимых на системы машинного обучения, что делает важным подход к защите данных систем.

Целью данной работы является повышение безопасности систем машинного обучения. В ходе работы были решены следующие задачи: рассмотрение существующих атак на системы машинного обучения и их систематизация, анализ методов защиты от атак на системы машинного обучения, экспериментальная апробация защиты системы на примере атак уклонения, а также определение дальнейших направлений работы по защите систем машинного обучения.

Существующие атаки на системы машинного обучения относят к одному из трёх видов атак: атаки уклонения, атаки отравления и исследовательские атаки [1-4].

Систему машинного обучения можно рассмотреть как систему из составляющих в виде входных данных, модели и выходных данных (результата). В зависимости от типа атаки происходит воздействие на одну из составляющих частей данной системы и нарушение безопасности такой системы. В таблице 1 приведена систематизация атак на системы машинного обучения (РО – режим обучения, РР – режим работы).

Таблица 1 - Систематизация атак на системы машинного обучения

Нарушение безопасности	**Объект атаки**	**Атаки уклонения**	**Атаки отравления**	**Исследовательские атаки**
Нарушение доступности	Результат	-	-	-
Нарушение целостности	Обучающая выборка	-	+ РО	-
	Входные данные	+ РР	-	-
Нарушение конфиденциальности	Обучающая выборка	-	-	+ РР
	Модель	-	-	+ РР

Все методы защиты от атак на системы машинного обучения можно разделить на активные (оборонные) и проактивные. Проактивные методы направлены на усиление модели: подготовку обучающих данных и/или модели, активные же методы применяются в режиме «реального времени» и позволяют отследить так называемые аномалии во входных данных (обучающих или тестовых).

Проактивные методы защиты, как правило, требуют либо более сложной реализации модели, либо «усиление» обучающих данных, что снижает итоговую точность классификации. В то время как оборонные методы работают с входными данными и не требуют изменения исходной модели, поэтому потери в точности классификации не происходит.

Можно заметить, что единого и универсального подхода к защите систем машинного обучения нет. Требуется поиск новых подходов и методов, в том числе «заимствование» методов из смежных областей, которые смогут производить эффективную защиту от атак на системы машинного обучения. Например, применение поведенческого анализа для ряда атак, направленных на раскрытие конфиденциальности данных (исследовательские атаки).

Таким образом, перспективным направлением работы в области защиты от атак на системы машинного обучения могут быть поиск и реализация таких методов защиты, которые используют поведенческий анализ.

Список литературы:

1. Pitropakis N., Panaousis E., Giannetsos T., Anastasiadis E., Loukas G. A taxonomy and survey of attacks against machine learning // Computer Science Review - 2019.

2. Chakraborty A., Alam M., Dey V., Chattopadhyay A. Mukhopadhyay D. Adversarial attacks and defences: A survey // ACM Computer Survey. - 2018.

3. Barreno M., Nelson B., Sears R., Joseph A. D., Tygar J. D. Can machine learning be secure? // In Proceedings of the 2006 ACM Symposium on Information, computer and communications security. – 2006. – p. 16–25.

4. Liu Q., Li P., Zhao W., Cai W., Yu S., Leung V. C. M. A survey on security threats and defensive techniques of machine learning: A data driven view. // IEEE Access. – 2018. – vol. 6. – p. 12103-12117.

Осипова Л.М. Полтавцева М.А.
Санкт-Петербургский политехнический университет Петра Великого

ФОРМАЛИЗАЦИЯ ДАННЫХ ИЗ ОТКРЫТЫХ ИСТОЧНИКОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ. ИСПОЛЬЗОВАНИЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ СЕТЕЙ ДЛЯ ЗАЩИТЫ ИЗОБРАЖЕНИЙ ОТ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ

Открытое информационное пространство интернета является наиболее популярным источником данных и представляет собой среду, образованную данными и сведениями, циркулирующими в свободно доступных цифровых медиаканалах. Данные из широкодоступных, "открытых" источников используются для решения целого ряда задач информационной безопасности [1,2]. К ним, в частности, относятся:

- борьба с "фейками";
- OSINT;
- деанонимизация субъектов и объектов различной природы.

Весь процесс аналитических задач над этим классом информации можно разделить на основные этапы, в соответствии теорией и с практикой построения аналитических систем: сбор данных, представление данных в некотором формализованном виде, проведение анализа, визуализация [2].

На сегодняшний день список инструментов для работы с открытыми источниками чрезвычайно широк и включает самые разные решения. Однако, если сопоставить их с приведенными этапами, то можно сделать вывод, что практически все средства автоматизации относятся к этапу сбора данных. В свою очередь задачу визуализации результата также можно решить универсальными средствами (например, Kibana [3]), чего нельзя сказать о задачах представления данных и анализа. Сложность их решения обусловлена:

- разнородностью. Данные чаще всего добываются из неоднородных источников и имеют разную степень структурированности. Это же осложняет подготовку наборов данных для машинного обучения;

– недоказанной достоверностью. Информация в источниках может быть полуправдивой или "фейковой", иметь низкий уровень достоверности, что в следствии негативно скажется на результате анализа, исказив его;

– объёмом. Данные имеют большой объём, что чаще всего делает трудновыполнимым или даже невозможным выполнение анализа в ручном или частично автоматизированном режиме.

Из вышеперечисленных факторов проистекает необходимость новых методов и средства автоматизации в этой области. Задача анализа данных не может быть решена без формализации или структурированного представления данных разнородных открытых источников для анализа.

Сложность представляет собой не только разнородность данных, но и разнородность источников, когда в одной схеме данных должны быть представлены социальная информация, результаты текстового анализа и анализа изображений, вывод технических средств, связанные между собой разнородными ассоциациями различной степени достоверности.

Для решения этой проблемы предлагается использовать формализацию на основе подхода "bag of objects" [4]. Данные и связи между ними предлагается формализовать в виде взвешенного метаграфа $G(V, E)$. На данный момент предполагается, что граф G - двумерный. Притом множество вершин V может включаться в себя два типа объектов:

– атомарный объект (факт). Под атомарным объектом подразумевается атомарный информационный элемент графовой структуры, в обобщённом представлении заданный кортежем <тип, значение>;

– комплексный объект. Комплексный объект включает в себя множество вершин, заданных атомарными объектами, и является метавершиной. В качестве комплексного объекта могут выступать человек, организационная структура, источник.

Связи между фактами могут иметь различную природу, в том числе, продиктованную источником данных (объекты одного фото, данные одной страницы из социальной сети и т.д.). Таким образом, факты связаны между собой через объекты опосредовано через другие факты (например, данные аккаунтов, зарегистрированных на один телефон или электронную почту). Такой граф является неориентированным и взвешенным. Веса рёбер определяют категорию доверия на основе источника данных, из которого был извлечён факт.

Конструктивный путь исследования данных в графовом представлении открывает возможности рассмотрения связей между фактами как свойств графа, и применения математического аппарата теории графов для дальнейшего анализа (в том числе прогнозирования поведения и изменяющихся связей). Такой подход является гибким и подходит для дальнейшего использования как в модулях интеллектуальных систем, отвечающих за верификацию данных, так и в модулях построения выводов.

Представляется, что такой подход позволит не только привлечь в данную область исследований математический аппарат, но и поможет свести сложную задачу анализа данных из открытых источников к набору более простых, открывая дополнительные возможности для автоматизации построения рассуждений на основе полученного графового представления.

Список литературы:

1. Кибербезопасность цифровой индустрии : теория и практика функциональной устойчивости к кибератакам : монография / Д.П. Зегжда, Е.Б. Александрова, М.О. Калинин [и др.] ; под редакцией Д. П. Зегжды. Москва : Горячая линия - Телеком, 2020. – 556 с.

2. Poltavtseva M.A., Pechenkin A.I. Intelligent data analysis in decision support systems for penetration tests // Automatic Control and Computer Sciences. – 2017. – Т. 51. – № 8. – С. 985-991. doi: 10.3103/S014641161708017X

3. Активное выявление угроз с Elastic Stack: Построение надежного стека безопасности: предотвращение, обнаружение и оповещение / пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2022. – 326 с.

4. Poltavtseva M.A., Semyanov P.V., Zaitzeva E.A. Heterogeneous semi-structured data analysis in information security // В сборнике: 2020 International Conference Engineering and Telecommunication, En and T – 2020. – 2020. – 1-5 С. doi: 10.1109/EnT50437.2020.9431309.

Хабибуллин А.В., Величко Д.В., Компаниец Р.И.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

АЛГОРИТМ ОБНАРУЖЕНИЯ ВРЕДОНОСНЫХ ПРОГРАММ НА ОСНОВЕ ДВУХУРОВНЕВОЙ ВИЗУАЛИЗАЦИИ И САМООРГАНИЗУЮЩЕЙСЯ ИНКРЕМЕНТНОЙ НЕЙРОННОЙ СЕТИ

Непрерывная эволюция и разнообразие вредоносных программ представляет серьезную угрозу для современных информационных систем. Ежедневно необходимо анализировать огромные объемы данных в поисках потенциальных угроз, но текущих методов просто недостаточно, чтобы удовлетворить высокий спрос [1]. Создатели вредоносного ПО используют различные методы защиты, такие как полиморфизм и обфускация, которые затрудняют обнаружение полезной нагрузки. Типичные подходы часто зависят от обнаружения на основе сигнатур, но несмотря на минимальные требования к вычислениям, они не восприимчивы к простым методам скрытия и эксплойтам нулевого дня. Статический и динамический анализ также не всегда эффективен и требует значительного времени для достижения результата. Для преодоления вышеуказанных проблем исследуется и оценивается эффективный алгоритм обнаружения, основанный на визуальном представлении вредоносных программ и самоорганизующейся инкрементной нейронной сети (SOINN).

Первым шагом двухуровневая визуализация преобразует двоичное содержимое файла в другую область, которая может быть представлена визуально. Предложенный алгоритм визуального представления основан на invis.io [2], онлайн инструмент, который использует цветовые схемы для представления различных двоичных значений или значений ASCII. Примеры визуализаций двух вредоносных файлов изображены на рисунке 2.

а) Trojan-Dropper.Win32.HeliosDinder.p б) Backdor.Win32.shodabot.b

Рисунок 1 – Визуализация вредоносного ПО

На этапе предварительной обработки из изображения извлекаются векторы признаков, способные указать на наличие вредоносной полезной нагрузки, чтобы их можно было использовать в процессе классификации.

Векторы признаков передаются SOINN [3], которая имеет два уровня. Первый уровень направлен изучение топологической структуры нейронной сети (NN), тогда как второй уровень определяет количество кластеров на основе входных данных. Удаление шума – один из важных аспектов SOINN, который отделяет его от других алгоритмов и позволяет NNсохранять только ценные фрагменты информации. Общий алгоритм изображен на рисунке 2.

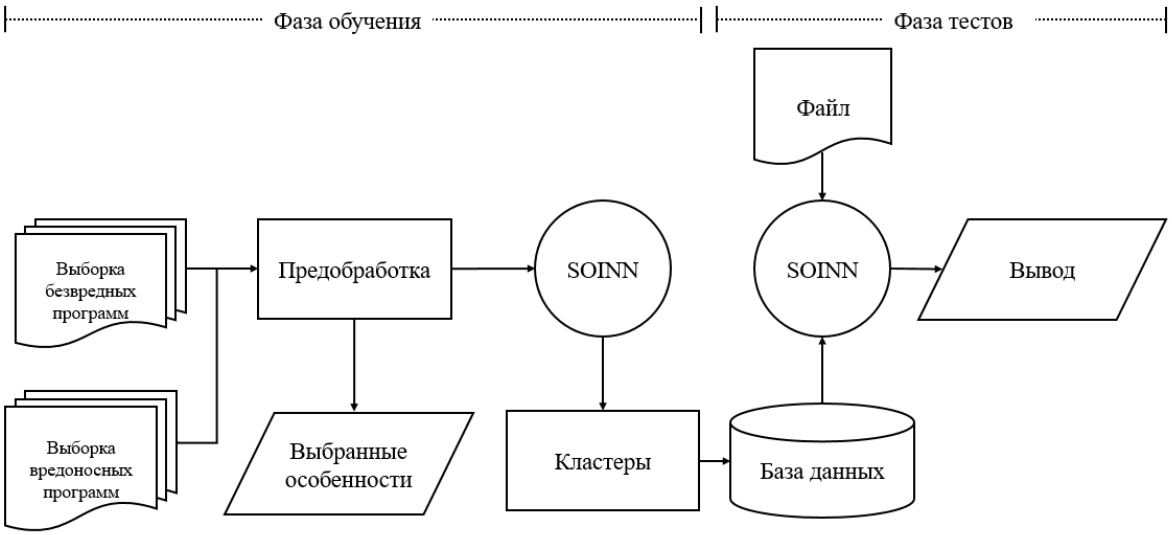

Рисунок 2 – Алгоритм обнаружения вредоносного ПО на основе двухуровневой визуализации и SOINN

Для тестирования алгоритма был собран набор данных, который содержал в общей сложности 4 тыс. безвредных файлов, полученных из надежных источников, и 2 тыс. вредоносных файлов, собранных с VirusShare. Все файлы были различных форматов. В результате работы алгоритма удалось добиться следующих показателей, изображенных на рисунке 3.

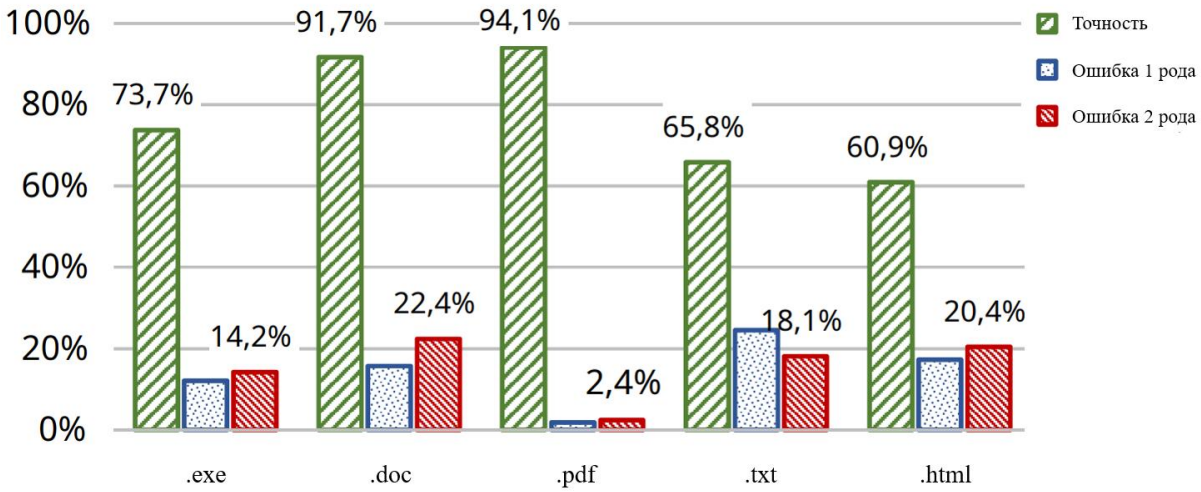

Рисунок 3 – Результаты работы алгоритма

Список литературы

1. GrammatikakisK.-P.,IoannouA.,ShiaelesS.,KolokotronisN.Are cracked applications really free? An empirical analysis on Android devices / 16th IEEE Int'l Conf. Dependable, Autonomic and Secure Computing (DASC).C. 730–735, 2018.
2. CortesiA.binvis.io: Visual Analysis of Binary Files. 2019.
3. SunY., LiuH., SunQ.Online learning on incremental distance metric for person re-identification/IEEE Int'l Conf. Robotics and Biomimetics (ROBIO). 2018.C. 1421–1426.

Хабибуллин А.В., Гомон А.В., Андрушкевич С.С.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

СПОСОБЫ РЕАЛИЗАЦИИ УГРОЗ БЕЗОПАСНОСТИ ИНФОРМАЦИИ НАРУШИТЕЛЯМИ С ПРИМЕНЕНИЯМИ ИНТЕЛЛЕКТУАЛЬНЫХ СРЕДСТВ ПРОВЕДЕНИЯ КОМПЬЮТЕРНЫХ АТАК

В настоящее время продолжается процесс расширения векторов атак, используемых злоумышленниками, что в свою очередь обуславливает необходимость внедрения новых технологий анализа в элементы защиты для отслеживания широкого спектра угроз. Эту нишу продолжают занимать технологии искусственного интеллекта [1], широко применяемые в задачах обнаружения аномалий в сетевом поведении, выявления подозрительной активности потенциально вредоносных программ, а также в задачах обнаружения угроз внутреннего контура на основе поведенческого анализа. Преимущество таких систем состоит в возможности анализа большого объема поступающих данных в режиме реального времени и реагирования на инциденты без вмешательства человека. Но наряду с успешным использованием в технологиях защиты, искусственный интеллект находит своё применение в задачах, реализуемых злоумышленниками.

Одним из способов реализации атак с применением средств искусственного интеллекта являются системы для эксплуатации известных уязвимостей. В основе таких систем лежат алгоритмы с использованием обучения с подкреплением, что облегчает работу настройки модели по причине ненадобности подготовки данных. Примером такого алгоритма является Asynchronous Advantage Actor Critic (A3C), впервые упомянутый в статье Asynchronous methods for deepre in for cement learning в 2016 году [2]. Данный алгоритм в отличие от классических алгоритмов обучения с подкреплением использует несколько агентов, каждый из которых имеет собственную среду и параметры, и, соответственно, взаимодействуют со своими средами асинхронно, что в свою очередь улучшает процесс обучения.

Описываемый тип систем эксплуатации уязвимостей в процессе своей работы может реализовывать несколько функций:

- сбор данных;
- моделирование угроз;
- анализ уязвимостей;
- эксплуатация уязвимостей;
- пост-эксплуатация.

Таким образом, для облегчения ряда процессов, необходимо внедрение в такие системы готовых фреймворков для задач, связанных с поиском, анализом и эксплуатацией

уязвимостей. Например, для реализации функции сбора данных целесообразно использовать nmap (поиск открытых портов и сервисов) и Scrapy (анализ содержимого HTTP-пакетов), а для задач анализа и эксплуатации уязвимостей – фреймворк Metasploit.

Работу такой системы можно описать следующим образом – на первом этапе модель обучается на тестовом сервере, эксплуатируя известные уязвимости в соответствии с данными, полученными в результате разведки, подбирая такие эксплойты, чтобы поддержка при обучении была выше: то есть выбранный эксплойт действительно соответствовал уязвимости сервера. Второй этап – непосредственно эксплуатация уязвимостей на целевом сервере в соответствии с накопленными знаниями на этапе обучения.

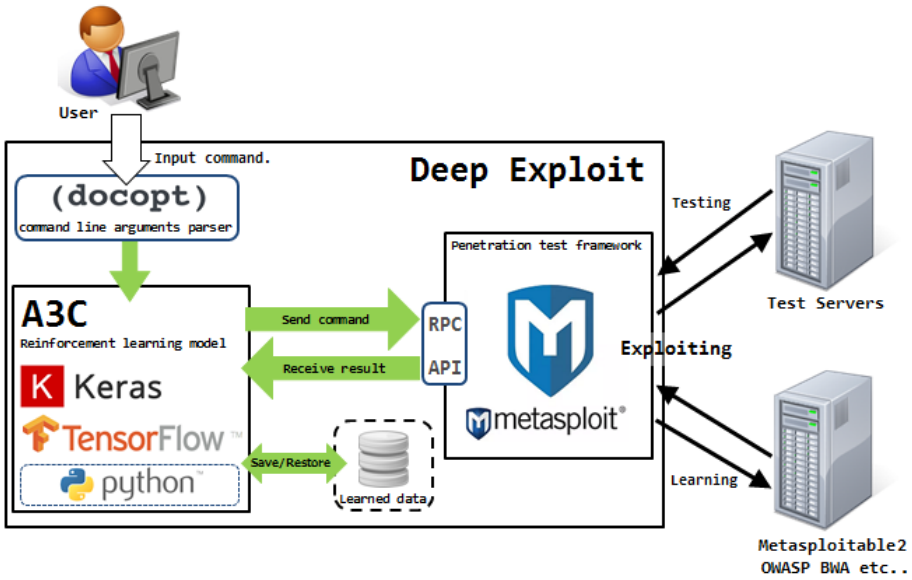

Рисунок 1 – Схема работы системы эксплуатации уязвимостей с использованием интеллектуальных средств

Таким образом, описанный способ реализации угроз безопасности с применением интеллектуальных средств является удобным для злоумышленника средством эксплуатации уязвимостей, что в свою очередь обуславливает необходимость применения дополнительных средств защиты.

Список литературы:

1. Петренко С.А., Бирюков Д.Н., Петренко А.С.В сборнике: The 2019 Symposium on Cybersecurity of the Digital Economy - CDE'19. третья международная научно-техническая конференция. 2019. С. 160-172.
2. Mnih V. Asynchronous methods for deep reinforcement learning //International conference on machine learning. – PMLR, 2016. – С. 1928-1937.
3. Carlini N., Wagner D. Adversarial examples are not easily detected: Bypassing ten detection methods //Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security. – 2017.
4. Goodfellow I. Nips 2016 tutorial: Generative adversarial networks //arXiv preprint arXiv:1701.00160. – 2016.
5. Arjovsky M., Chintala S., Bottou L. Wasserstein generative adversarial networks //International conference on machine learning. – PMLR, 2017.

6. University of New Brunswick. NSL-KDD dataset. [Электронный ресурс]. URL: https://www.unb.ca/cic/datasets/nsl.html(дата обращения 13.05.2022г.).

7. Lin Z., Shi Y., Xue Z. Idsgan: Generative adversarial networks for attack generation against intrusion detection //arXiv preprint arXiv:1809.02077. – 2018.

Асадуллин А.Я., Менисов А.Б.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

МЕТОДИКА ОПРЕДЕЛЕНИЯ АНОМАЛИЙ ФУНКЦИОНИРОВАНИЯ ОБЪЕКТОВ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ НА ОСНОВЕ СВЕРТОЧНЫХ АВТОЭНКОДЕРОВ

** Исследование выполнено в рамках гранта Президента РФ для государственной поддержки молодых российских ученых – кандидатов наук(МК-2485.2022.4).*

Критическая информационная инфраструктура (КИИ) – это сложная система, элементы которой используют различные программно-аппаратные компоненты при функционировании. Включение Интернета вещей (IoT) в КИИ открывает новые возможности для злоумышленников использовать уязвимости системы для проведения кибератак [1].

Среди мер по обеспечению безопасного функционирования объектов КИИ выделяют регламентацию правил и процедур реагирования на компьютерные инциденты, выявление и анализ компьютерных инцидентов, защиту информации и информирование о компьютерных инцидентах, устранение их последствий и принятие мер по недопущению их повторного возникновения.

Процесс обеспечения безопасности может состоять из различного числа циклов – от двух до многих десятков. Поэтому в каждом цикле проводиться оценка и коррекция управляющего воздействия. Сущность обеспечения защиты объектов КИИ состоит в следующем. После того, как установлены специфические особенности функционирования, можно переходить к построению прогноза и исхода защиты объекта КИИ.

Цель разработанной методики заключается в том, чтобы классифицировать новые наблюдения (временные ряды) как можно раньше, предпочтительно до того, как будет доступен полный временной ряд. Этот подход известен как ранняя классификация временных рядов. Автоэнкодеры используются для неконтролируемых методов обучения, они имеют возможность сжимать информацию до минимального количества узлов во внутренних слоях, что позволяет, с проверенным кодировщиком, хранить ту же информацию в уменьшенном количестве.

В большинстве задач, направленных на определение аномального состояния функционирования объектов КИИ, данные функционирования представляют собой временные ряды и делятся на две категории: нормальные (N) и аномальные (A). Кроме того, характерной чертой является несбалансированное распределение классов. Таким образом, выявление аномального функционирования может быть формализована как задача классификации несбалансированных временных рядов.

Высокие значения метрик полноты(recall) и точности (precision) предлагаемой методики позволяют минимизировать пропуск аномалии. Именно поэтому свёрточный автоэнкодер имеет преимущества по сравнению с другими архитектурами нейронных сетей.

Таким образом, представленная методика на основе сверточных автоэнкодеров увеличивает качество определения аномалий функционирования объектов критической информационной инфраструктуры

Симаков А.А. Шалькин Д.О. Дудкин А.С.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

МЕТОДИКА ИССЛЕДОВАНИЯ ЗАЩИЩЕННЫХ JAVA ПРИЛОЖЕНИЙ

Основной целью исследований является разработка программного средства обратного проектирования Java-приложений. Разработанное программное средство предназначено для упрощения исследования и понимания хода работы Java-приложений, в том числе мобильных (Android), позволяет ускорить процесс анализа приложений, путем деобфускации кода, встраивание трассировки, а также отображением управляющего графа программы как для классов, так и для методов (рисунок 1).

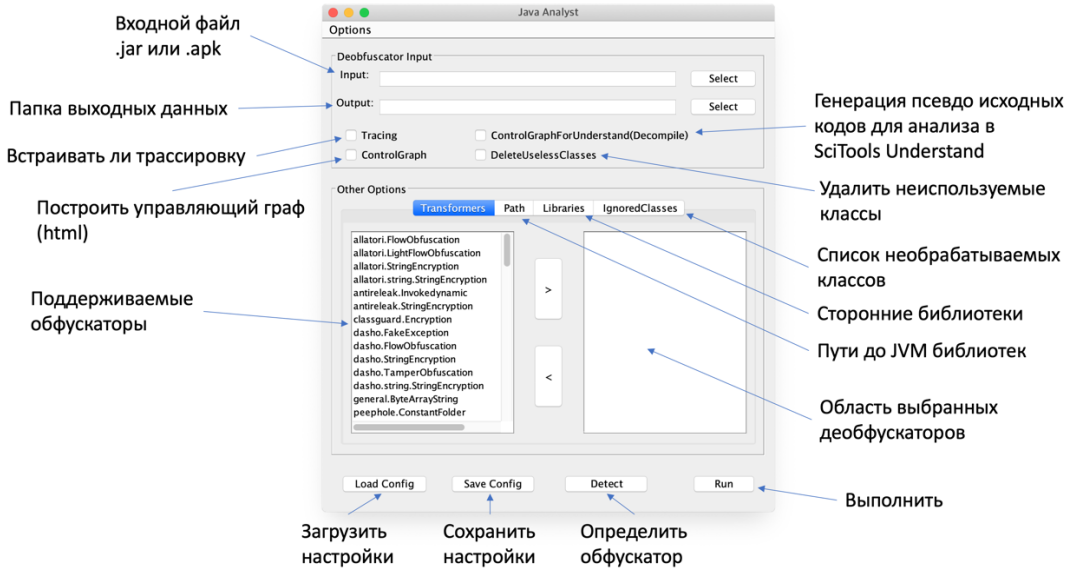

Рисунок 1 – Интерфейс программного комплекса

Основные функции:

- детектирование обфускатора;
- деобфускации;
- встраивание трассировки;
- выявление обращений на удаленные сервера;
- построение управляющего графа.

За основу программного комплекса был взят opensource проект java-deobfuscator [1]. Разработанная программа с помощью заданных алгоритмов определяет тип обфускатора, а также предполагаемую последовательность инструкций для деобфускации.

Поддерживаемые обфускаторы:

- ZelixKlassmaster;

- Stringer;
- Allatori;
- DashO;
- DexGuard;
- ClassGuard;
- Smoke.

Для встраивания трассировки программа дизассемблирует Java-приложение и в полученные файлы, представляющие собой байт-код, вставляет системный вызов вывода в консоль строки с названием метода и именем класса:

GETSTATIC java/lang/System.out Ljava/io/PrintStream;

LDC "<Название класса и метода>"

INVOKEVIRTUAL java/io/PrintStream.println(Ljava/lang/String;)V

Параллельно с этим для построения управляющего графа программа анализирует дизассемблированные файлы и рекурсивно ищет в них методы, в которых другие методы вызываются из тела первого и так далее. Аналогично и для классов (рисунок 2):

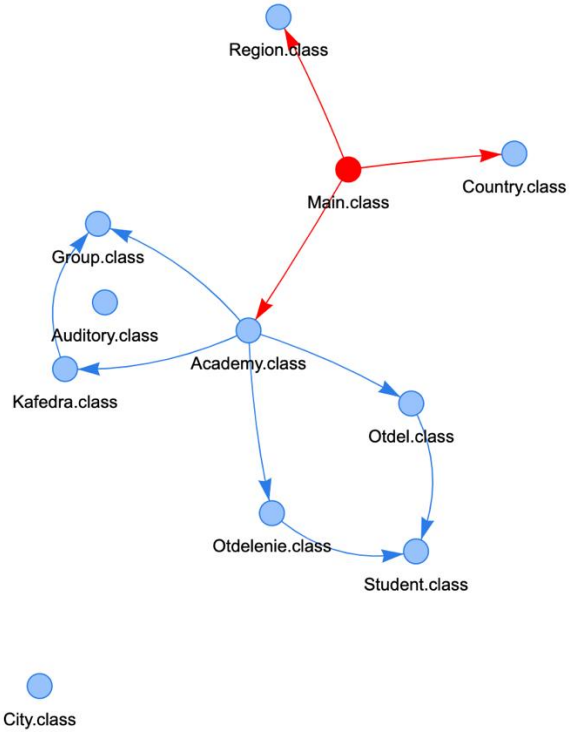

Рисунок 2 – Пример графа классов

Список литературы

1. Java-Deobfuscator[Электронный ресурс] – https://github.com/java-deobfuscator/deobfuscator;
2. Внутренности JVM, Часть 2 — Структура class-файлов [Электронный ресурс] – https://habr.com/ru/company/otus/blog/478584/;

3. Введение в байт-код Java [Электронный ресурс] – https://medium.com/nuances-of-programmingвведение-в-байт-код-java-которое-вам-пригодится-даже-если-вы-считали-иначе-69b8de0bcf3b;
4. Основы JavaBytecode [Электронный ресурс] – https://habr.com/ru/post/568402/;
5. Структура программ на языке Java [Электронный ресурс] – http://zonakoda.ru/struktura-programm-na-yazyke-java.html.

Чичалов М.Р., Крюков Р.О.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

ПОДХОД К ВЫЯВЛЕНИЮ В СИСТЕМЕ ВРЕДОНОСНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ, СОДЕРЖАЩЕГО SHELLCODE

В условиях сложившейся геополитической обстановки в мире наблюдается тенденция роста количестваAPT-атак, которые представляют огромную угрозу для всего цифрового пространства. При осуществлении APT-атак в соответствии с матрицей MITREATT&CK можновыделить следующие основные фазы от получение первоначального доступа (Initial Access) до воздействий (Impact) на скомпрометированную систему.

В настоящее время значительно увеличилось количество успешных APT-атак, в результате которых злоумышленник получает полный или частичный доступ к системе. После успешного получения первоначального доступа нарушитель осуществляет попытки «закрепления» в системе путем несанкционированного создания учетных записей или кражи существующих учетных данных, скрытой установки и запуск средств удаленного доступа или внесения в конфигурацию атакуемой системы изменений, с помощью которых становится возможен многочисленный запуск вредоносного кода. Вместе с тем присутсвие злоумышленника в сети остается незамеченным для большинства средств защиты информации. Поэтому для обнаружения и предотвращения APT-атак необходимы специальные решения.

В проведенном исследовании использовались индикаторы компрометации. Индикатор компрометации (Indicator of Compromise, IoC) - это объект/артефакт, обнаруженный в ИТ-инфраструктуре компании, наличие которого с высокой долей вероятности может свидетельствовать о готовящейся, совершающейся или уже осуществленной компьютерной атаке. В качестве индикаторов компрометации (IoCs) могут быть использованы такие статические объекты, как хэш-суммы файлов и их имена и расположение, IP-адреса, DNS-имена серверов в сети Интернет или конкретные URL (например, ссылки на фишинговые страницы), названия веток и ключей реестра Windows, названия мьютексов (mutex, специализированный механизм синхронизации исполняемого программного кода). Кроме этого, могут быть использованы и динамические объекты, такие как определенная последовательность несанкционированных действий на атакуемой системе, которые также называют индикатором атаки, Indicator of Attack (IoA), необычное поведение учетных записей в системе (это может быть выявлено системами класса UEBA - User & Entity Behavior Analysis, системы анализа поведения пользователей и сущностей), несанкционированное появление новых учетных записей, особенно высокопривилегированных, а также рост числа

подозрительных DNS-запросов, ICMP-трафика, иных видов ранее нехарактерной сетевой активности.

В основе реализации предлагаемого плагина к антивирусной системе будут лежать проверки на наличие ВПО на сервере, путем обнаружения следующих индикаторов атак (Рисунок 1):

- периоды аномально высокой нагрузки на сервер;

- наличие файлов с подозрительной временной меткой (например, более поздней, чем время последнего обновления ПО);

- наличие подозрительных авторизаций из внутренней сети;

- наличие файлов, генерирующих несвойственный им трафик.

Таким образом предложенный плагин для антивирусной системы защиты позволит дополнительно осуществлять поиск сигнатур известных шелл-кодов в файлах, которые могут быть использованы злоумышленником после получения первоначального доступа для закрепления в системе.

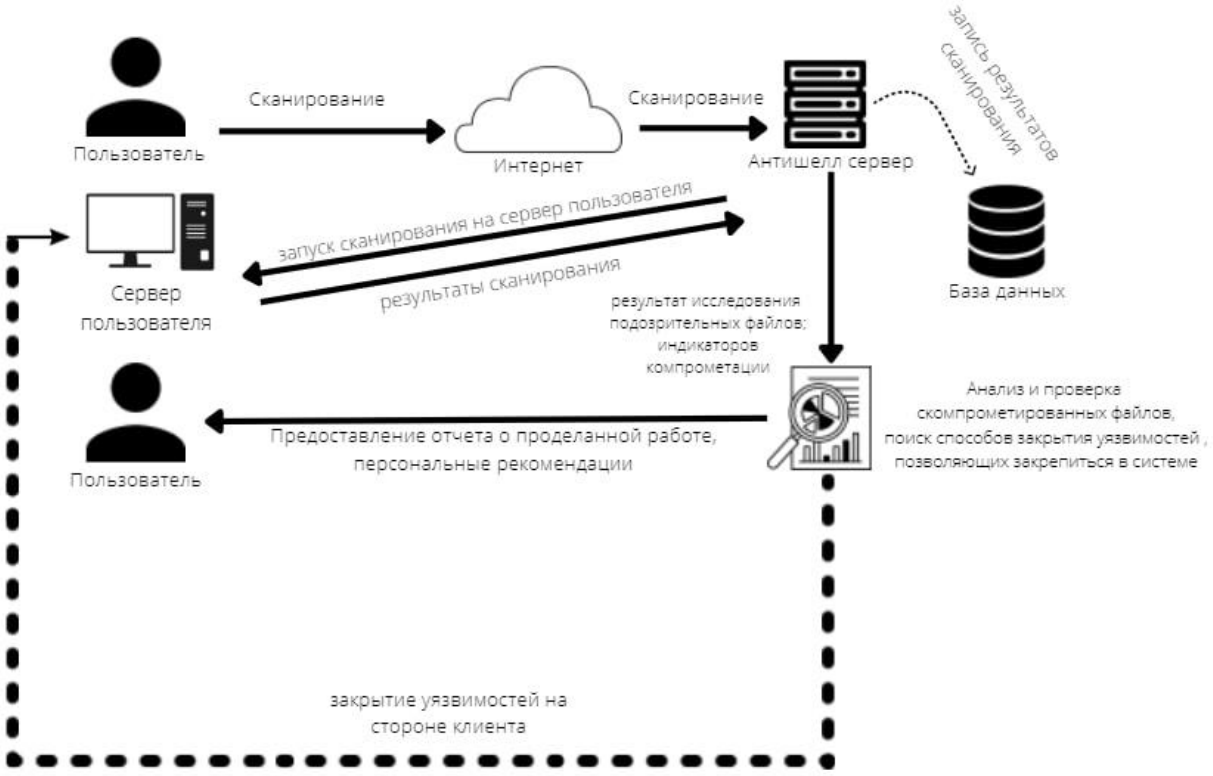

Рисунок 1 – Порядок обнаружения файлов, содержащих шелл-код

Список литературы

1. Реверсинг малвари для начинающих. Внедрение shellcode и шифрование malware-кода [Электронный ресурс]. Режим доступа: https://xakep.ru/2017/05/11/reversing-malware-tutorial-part5/ (Дата обращения: 25.05.22).

2. MITRE ATT&CK [Электронный ресурс]. Режим доступа: https://attack.mitre.org/ (Дата обращения: 25.05.22).

3. Белоус А.И., Солодуха В.А. Основы кибербезопасности. Стандарты, концепции, методы и средства обеспечения Москва: ТЕХНОСФЕРА, 2021. – 482 с.

Иванов Д.С., Миннигалин Д.Р., Крюков Р.О.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

СИСТЕМА ОБНАРУЖЕНИЯ СЕТЕВЫХ АТАК НАПРАВЛЕННЫХ НА МОБИЛЬНЫЕ УСТРОЙСТВА, ПУТЕМ АНАЛИЗА ИХ ИНДИКАТОРОВ

В настоящее время функциональные возможности современных мобильных устройств практически не уступают возможностям персональных компьютеров. Для мобильных устройств существует большое количество разнообразного программного обеспечения, которое доступно для установки в магазинах приложений. Программное обеспечение для мобильных устройств имеет широкий спектр применения, включая доступ веб-ресурсам, социальные сети, банковские и прочие услуги. Поэтому мобильные устройства стали неотъемлемой частью повседневной жизни.

Однако тенденция активного использования мобильных устройств и рост ценности информации, обрабатываемой в них способствовало увеличению количества целевых атак (APT-атак).Кроме того, в официальных магазинах приложений (например, App Store и Google Play Store), несмотря на реализованную защиту от угроз, возможно нахождение вредоносного программного обеспечения, которое может маскироваться под легитимные мобильные приложения. Поэтому задача обеспечения информационной безопасности мобильных устройств является весьма актуальной.

На основе анализа реальных APT-атак, в базе знаний MITRE ATT&CK (Adversarial Tactics, Techniques & Common Knowledge) описаны тактики и техники атак на мобильные устройства. Знание того, как действуют реальные APT-группировки, позволяет сформировать индикаторы атак (Indicator of Attack, IoA) для обнаружения вредоносных действий при функционировании недоверенного программного обеспечения на мобильном устройстве.

Рисунок 1 – Алгоритм работы системы анти-APT для мобильных устройств

Предлагается подход к обнаружению известных APT-атак, на основе анализа сетевого трафика мобильного устройства с использованием индикаторов атак. Суть работы предложенного подхода заключается в следующем (*Рис. 1*):

– весь трафик мобильного устройства попадает в систему анти-APT, где осуществляется сравнение UID программного обеспечения со списком доверенных приложений. В случае отсутствия UID программного обеспечения в списке доверенных приложений трафик проходит проверку на наличие угроз;

– процесс анализа трафика от недоверенного программного обеспечения осуществляется на основе обнаружения индикаторов атак IoA;

– при выявлении вредоносных действий недоверенному программному обеспечению присваивается уровень опасности в зависимости от того каким этапамAPT-атаки соответствуют обнаруженные индикаторы атак;

– в зависимости от уровня опасности недоверенного программного обеспечения система будет принимать одно из следующих действий: уведомление пользователя об аномальном сетевом трафике, перенаправление потенциально опасного трафика на дополнительную проверку или блокировка вредоносного трафика.

Таким образом, предложенная система анти-APT позволит повысить результативность проактивного [5] обнаружения угроз информационной безопасности для мобильных устройств.

Список литературы

1. MITRE ATT&CK [Электронный ресурс]. Режим доступа: https://attack.mitre.org/ (Дата обращения: 08.05.22).
2. Коллинз М. Защита сетей. Подход на основе анализа данных / перевод с анг. А.В. Добровольская. – М.: ДМК Пресс, 2020. – 308 с.
3. Диогенес Ю., Озкайя Э. Кибербезопасность: стратегии атак и обороны / пер. с анг. Д. А. Беликова. – М.: ДМК Пресс, 2020. – 326 с.
4. Белоус А.И., Солодуха В.А. Основы кибербезопасности. Стандарты, концепции, методы и средства обеспечения Москва: ТЕХНОСФЕРА, 2021. – 482 с.
5. Бирюков Д.Н., Ломако А.Г., Петренко С.А., Ступин Д.Д. В сборнике: РТИ Системы ВКО – 2016. Труды IV Всероссийской научно-технической конференции. 2017. С. 758-767.

Захаров О.О., Бирюков Д.Н., Тимашов П.В., Дудкин А.С.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург

ПОДХОД К СОЗДАНИЮ ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ БЛОКИРОВАНИЯ ПОТЕНЦИАЛЬНО ВРЕДОНОСНЫХ ОФИСНЫХ ДОКУМЕНТОВ С МАКРОСАМИ

Документы MicrosoftOffice являются наиболее распространенным типом файлов для доставки вредоносного программного обеспечения [1], поскольку сотрудники в организациях постоянно обмениваются ими – подобные файлы с большей вероятностью загрузят на рабочую станцию даже от неизвестного отправителя, чем файлы исполняемого или иного

формата. Вредоносные документы используют подсистему макросов для загрузки и выполнения полезной нагрузки. Несмотря на то, что защитные решения учитывают данный вектор атаки, они зачастую неэффективно противодействуют ему. Авторами был предложен подход к созданию программного комплекса блокирования потенциально вредоносных документов с макросами, полученных из Интернета, а также разработан прототип, основанный на нем.

Создание подобного программного комплекса обусловлено необходимостью наличия дополнения к существующим защитным решениям, а также отсутствием групповой политики «Блокирование запуска макросов в файлах Microsoft Office, полученных через Интернет» на тех системах, где установлен MicrosoftOffice ниже версии 2016. Кроме того, данная политика отсутствует для офисных пакетов LibreOffice и OpenOffice.

Авторами предложен следующий алгоритм работы программного комплекса:

1. Перехват операций с файловой системой;
2. Отбор файлов по расширениям, соответствующим документам MicrosoftOffice, LibreOffice, OpenOffice;
3. Анализ файла
    a. Если расширение *«*.docm»* или *«*.xlsm»* - переход к шагу 4;
    b. Если расширение «*.*doc»* или «*.*xls»* - производится поиск и чтение потоков _VBA_PROJECT или Macros при помощи библиотеки от Microsoft для формата CFBF (Compound File Binary File)[2];
    c. Если расширение *«*.odt»и др.* - производится чтение zip-архива и проверяется наличие директории Basic, в которой обычно находится макрос.
4. Запрет на загрузку документов с макросами из сети Интернет – проверяется альтернативный поток данных (ADS) Zone.Identifier для уточнения источника файла (ZoneId=3 – Интернет) [3];
5. Оповещение системного администратора (офицера безопасности) об обнаруженной угрозе.

На основе алгоритма был разработан прототип программного комплекса, состоящий из 3 компонентов:

1. Драйвера минифильтра файловой системы – для получения событий, связанных с файловой системой;
2. Приложения пользовательского режима – для взаимодействия с драйвером, обработки запросов на предоставление доступа к файлу;
3. Серверного приложения – для принятия сообщений от клиентов, визуализации полученной информации.

В результате анализа существующих методов получения первоначального доступа, алгоритмов работы защитных решений, а также политик безопасности операционных систем семейства Windows, авторами предложен подход, учитывающий ряд проблем и недостатков, а разработанный на его основе программный комплекс реализует следующие механизмы:

- мониторинга файловой системы;

- обнаружения офисных документов различных форматов с макросами, полученных из Интернета;
- оповещения системного администратора (офицера безопасности) об обнаруженных потенциально опасных офисных документов.

Список литературы

1. Annual State of Phishing Report 2021 // cofense.com [Электронный ресурс]. URL:https://cofense.com/wp-content/uploads/2021/02/cofense-annual-report-2021.pdf (дата обращения 23.05.2022)

2. A simple header file to read Microsoft compound file with minimal efforts.// github.com [Электронный ресурс]. URL:https://github.com/microsoft/compoundfilereader (дата обращения 23.05.2022)

3. About URL Security Zones// docs.microsoft.com [Электронный ресурс]. URL:https://docs.microsoft.com/en-us/previous-versions/windows/internet-explorer/ie-developer/platform-apis/ms537183(v=vs.85) (дата обращения 23.05.2022)

Швец Н.П., Андрушкевич Д.В.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

ИССЛЕДОВАНИЕ РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ

На сегодняшний момент в рамках межгосударственных конфликтов все больше внимания и финансирования выделяется на ведение деструктивных кампаний на различных информационных ресурсах. В частности, все больше и больше внимания уделяется развитиюсредств распространения информации (СМИ) в сети Интернет. Некоторыми способами ведения деструктивной деятельности в СМИ принято считать [1]:

- манипулирование с истинной информацией;
- фильтрация актуальных тем и материалов;
- выбор выгодного момента для информирования населения (например, для выброса компрометирующих материалов);
- эмоциональное комментирование представления происходящего;
- двухступенчатый коммуникационный поток: использование в комментариях мнений компетентных персон, уважаемых обществом, и другие.

Для выявления такого рода деятельности прежде всего необходимо иметь возможность определять направленность публикуемых материалов и сравнивать их популярность, а также тематику в различные промежутки времени. Это позволит выявить отклонения от нормальной деятельности и определить тему, на которую ведется смещение внимания.

Большинство пользователей предпочитает получать информацию в специализированных приложениях.Данные приложения для максимального упрощения и ускорения процесса получения информации содержат функции обмена сообщениями и звонками. Яркими примерами таких приложений являются мессенджер Telegram и социальная сеть Вконтакте.

В рамках исследования распространения информациив социальных сетях был разработан ряд программных модулей, выполняющих функции по сбору данных с анализируемых Telegram каналов, их структуризации, предобработке, тональной оценке выделенных сообщений. Программный модуль, отвечающий за парсинг информации, основан на библиотекеTelethon.Модуль сбора данных представляет собой telegramклиент, подключающийся через apitelegram и выгружающий всю историю постов посредством запроса GetHistoryRequest. Программа сохраняет собранную информацию в файл формата ".json". Определяется значения следующих ключевыхпараметров:

- media–наличие медиафайлов;
- pinned –закреплено ли сообщение;
- views –количество просмотров сообщения;
- forwards –количество репостов сообщения;
- replies–количество комментариев сообщения;
- message –сам текст сообщения;
- date–дата отправки сообщения;
- id–идентификатор сообщения и пр. параметры (в дальнейшем не используются).

В результате получается следующая выборка данных:

	_	id	date	message	out	pinned	media	views	forwards	replies
0	Message	14898	2022-05-01 10:46:11+00:00	⚡В Северной Осетии появилось изображение с ба...	False	0.0	{'_': 'MessageMediaPhoto', 'photo': {'_': 'Pho...	43908.0	140.0	None
1	Message	14897	2022-05-01 10:39:08+00:00	⚡По данным издания Financial Times, Германия ...	False	0.0	{'_': 'MessageMediaPhoto', 'photo': {'_': 'Pho...	86782.0	420.0	None

	_	id	date	message	out	pinned	media	views	forwards	replies	token
0	Message	14898	2022-05-01 10:46:11+00:00	в северной осетии появилось изображение с бабу...	False	0.0	{'_': 'MessageMediaPhoto', 'photo': {'_': 'Pho...	43908.0	140.0	None	[северной, осетии, появилось, изображение, баб...
1	Message	14897	2022-05-01 10:39:08+00:00	по данным издания financial times германия за ...	False	0.0	{'_': 'MessageMediaPhoto', 'photo': {'_': 'Pho...	86782.0	420.0	None	[данным, издания, financial, times, германия, ...
				путин подписал							[путин, подписал

Рисунок 1 – Проверка полученной выборки данных и их препроцессинг

Для обработки текстовых данных были применены методы по очистке от избыточности слов, пунктуации, понижения регистра посредством использования библиотеки spacy. Далее была применена токенизация, леммантизация текстовых данных. В результате получена выборка обработанныхтекстовых данных. Обработка осуществлялась при помощи библиотеки Natasha – русскоязычной NERмодели для выделения сущностейв новостных лентах[2]. С ее помощью может быть выявлена в текстовых постах канала/сообщениях чата информация об организациях, персонах, событиях и ряд другой информации, задаваемой пользователем. В рамках проведенного исследования также осуществлялся анализ тональности всех постов в выбранном телеграмм канале с выделением сущностей различного характера. Тональность является одним из ключевых факторов, позволяющим определить наличие или отсутствие в нем деструктивной составляющей. В результате выполнения указанного анализа формируетсяграфическое представление о корреляции между эмоциональной окраской,

даваемой автором, и выделенными сущностями, а также предупреждение о наличии/отсутствии деструктивной составляющей в сообщениях канала.

Для описания процесса исследования распространения информации предлагается применить нотациюIDEF0 [3].

Данное моделирование позволяет отобразить используемые ресурсы, средства, которыеприменяются привыполнении исследования, а также потоки информации, преобразуемые в ходе исследования. Одной из наиболее важных особенностей нотации является постепенное введение все больших уровней детализации процесса по мере создания диаграмм, отображающих модель [4]. На рисунке 2 представлена диаграмма декомпозиции первого уровня.

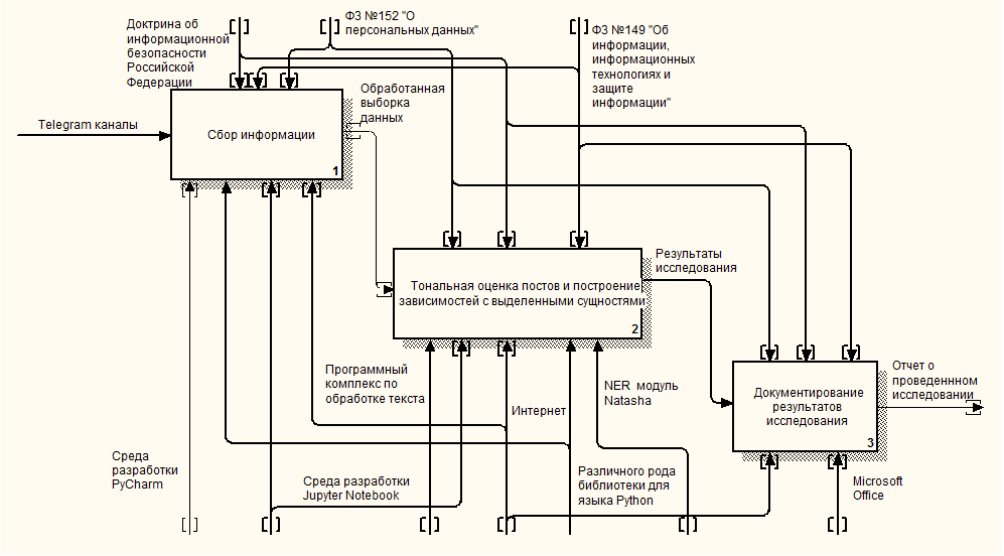

Рисунок 2 – Диаграмма исследования распространения информации

Такое моделирование позволяет упростить восприятие проводимого исследования распространения информации аудиторией в социальных сетях. Построенная модель закладывается в основу программного комплекса по предсказанию распространения негативного контента в социальных сетях.

Список литературы

1. Манойло А.В. Государственная информационная политика в особых условиях: Монография. М.: МИФИ, 2003. – С.115-122
2. Natasha – качественное компактное решение для извлечения именованных сущностей из новостных статей на русском языке[Электронный ресурс] URL: https://natasha.github.io/ner/(дата обращения: 28.03.2022)
3. Как перезагрузить ИБ с помощью процессного моделирования [Электронный ресурс] URL: https://www.anti-malware.ru/practice/methods/process-modeling-in-information-security(дата обращения: 22.03.2022)
4. Моделирование бизнеса. Основные подходы [Электронный ресурс] URL: https://habr.com/ru/company/trinion/blog/332772/(дата обращения: 15.03.2022)

Нигматуллин Р.А. , Андрушкевич Д.В., Бирюков Д.Н.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт Петербург

ПОДХОД К УНИФИЦИРОВАННОЙ КЛАССИФИКАЦИИ ТЕКСТОВЫХ СООБЩЕНИЙ

Рост количества информации, которую необходимо анализировать,способствует увеличению времени, требуемого на ее обработку. Использование удобной, понятной и общепринятой системы классификации позволит упростить этот процесс.

Классификатор – систематизированный перечень наименованных объектов, каждому из которых в соответствие дан уникальный код[1].Для создания унифицированного классификатора разумнее всего использовать уже имеющиеся, которые достаточно распространены, постоянно дополняются и понятны без предварительного глубокого изучения.

Для разработки более обширного и универсального классификатора было проведено сравнение следующих классификаторов: УДК, ББК, МПК.

УДК – универсальная десятичная классификация – система классификации информации, широко используется во всём мире для систематизации произведений науки, литературы и искусства, периодической печати, различных видов документов и организации картотек [2].

ББК – библиотечно-библиографическая классификация – национальная классификационная система России, которая используетсядля систематизации произведений науки, литературы и искусства, периодической печати, различных видов документов и организации картотек [3].

МПК – международная патентная классификация – иерархическая система патентной классификации. МПК является средством для классификации патентных документов единообразной в международном масштабе [4].

Сравнения производились по следующим параметрам:

- наличие иерархичной структуры;
- возможность комбинации с другими классификаторами;
- возможность обозначения территориальной местности, рассматриваемой в сообщении;
- возможность обозначения времени, рассматриваемого в сообщении;
- возможность обозначения информации о людях, рассматриваемых в сообщении;
- наличие отсылок.

Учитывая важность времени поступления информации и ее источника, разработанный классификатор был дополнен соответствующими специальными параметрами, которые позволяют учесть место откуда поступила информация и время ее поступления.

Однако недостаточно просто предоставить оператору классификатор. При большом объеме обрабатываемых данных нецелесообразно вручную распределять их по имеющимся классам. Поэтому необходимо создать удобную среду, которая позволит оператору правильно и быстро распределять требуемую информацию.

Предлагается с помощью семантического анализа [5] создать систему ввода текста, которая будет автоматизировано анализировать сообщение и после предлагать оператору возможные варианты его отнесения к соответствующей категории.

В качестве примера рассмотрим классификацию следующего сообщения:«Как сообщилаMicrosoft, в последние полгода началось активное распространение вируса XorDDoS, предназначенного для Linux-систем» (табл. 1).

Таблица 1. Пример классификации сообщения

Наимено-ваниеклассифи-цируемого объекта	Место нахож-дение объекта	Период времени	Место полу-чениянфор-мации	Время поступ-лениянфор-мации	Кто объект (кто воздейст-вует)	Кто субъект (на что воздейст-вуют)	Инфор-мация о людях (если требует-ся)
004.056.57	-	21.11.21-21.05.22	*73	*21.05.22 16:01	004.49-051 (XorDDoS)	004.451-052 (Linux)	-

Согласно УДК:

004.056.57 – защита от компьютерных вирусов;

*73 – США;

004.49-051 (XorDDoS) – компьютерный вирус XorDDoS;

004.451-052 (Linux) – операционная система Linux.

Рассмотренный подход реализован в программном комплексе и может использоваться при решении задач регистрации событий для их обработки и последующего анализа.

Список литературы:

1. Классификатор-Википедия [Электронный ресурс]: URL: https://ru.wikipedia.org/wiki/%D0%9A%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%82%D0%BE%D1%80 (дата обращения: 20.02.2022)
2. УДК: структура, свойства и принципы[Электронный ресурс]: URL: http://www.spsl.nsc.ru/win/nelbib/UDK/strukt.html (дата обращения: 01.03.2022)
3. Библиотечно-библиографическая классификация [Электронный ресурс]: URL: https://dic.academic.ru/dic.nsf/ruwiki/84376 (дата обращения: 13.03.2022)
4. Руководство к МПК-ФИПС [Электронный ресурс]: URL: https://fips.ru/elektronnye-servisy/klassifikatory/mezhdunarodnaya-patentnaya-klassifikatsiya/rukovodstvo-k-mpk-.php(дата обращения: 01.03.2022)
5. Пескова О. В. Алгоритмы классификации полнотекстовых документов // Автоматическая обработка текстов на естественном языке и компьютерная лингвистика. – М.: МИЭМ (Московский государственный институт электроники и математики), 2011. – С. 170 – 212.

Привалов А.Ю., Исаков С.А., Сабиров Т.Р.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург

ПОДХОД К ОБНАРУЖЕНИЮ И УСТРАНЕНИЮ УЯЗВИМОСТИ "DIRTY PIPE" В ОС ASTRA LINUX

Одним из основных направлений повышения уровня защищенности критической информационной инфраструктуры Российской Федерации является применение отечественных информационных технологий. Данный подход является основным, даже можно сказать единственным посылом, о котором заявили 20 мая 2022 г. члены Совета безопасности России, одобрившие проект основ государственной политики в сфере информационной безопасности [1].

Флагманом отечественной операционной системы без сомнения является AstraLinux, которая вытесняет традиционные системы во всех государственных структурах и корпорациях, а также организациях, которые непосредственно взаимодействуют с ними(МО РФ, ФСБ, Роскосмос и т.д.). Следовательно, существует реальная необходимость поддерживать высокий уровень безопасности таких систем, например, осуществлять контроль за наличием уязвимости функционирующих систем.

Так, в феврале 2022 года была обнаружена уязвимость CVE-2022-0847, которая получила название «Dirtypipe». Она была отнесена к классу «Уязвимостей кода» и позволяла перезаписывать данные в произвольных файлах.На сайте ФСТЭК России уровень опасности был оценён в 6.8 (средний) и 7.8 (высокий) баллов по стандартам CVSS 2.0 и CVSS 3.0 соответственно.

Уязвимость возникает из-за использования частично неинициализированной памяти в структуре буфера конвейера во время его построения. Отсутствие нулевой инициализации нового члена структуры приводит к использованию устаревшего значения флагов. С помощью уязвимости можно получить доступ на запись к страницам памяти в кэше, даже если эти страницы были изначально помечены атрибутом «Только для чтения».Существует множество способов получения привилегии суперпользователя с помощью «DirtyPipe», например, через подмену исполняемых файлов с правами SUID, модификации /etc/passwd и т. п.[2].

Для противодействия описанной уязвимости существует несколько подходов:

— применение всех необходимых обновлений безопасности системы по мере их поступления;

— ограничение работы с интерпретируемыми языками программирования;

— процесс безопасной разработки по ГОСТ Р 56939-2016, собственные методики и подходы к разработке безопасной архитектуры, использование автоматизированных проверок (статический и динамический анализ) кода[3].

На данный момент предложен единственный подход к обнаружению данной уязвимости – проверка версии ядра. Для этого необходимо запустить скрипт, изображенный на рисунке 1, позволяющий установить точную версию ядра дистрибутива ОСAstraLinux «Орёл» и возможность эксплуатации уязвимости. Устранение уязвимости заключается в обновлении версии ядра операционной системы от 5.15.11, 5.15.25, 5.10.102 и выше.

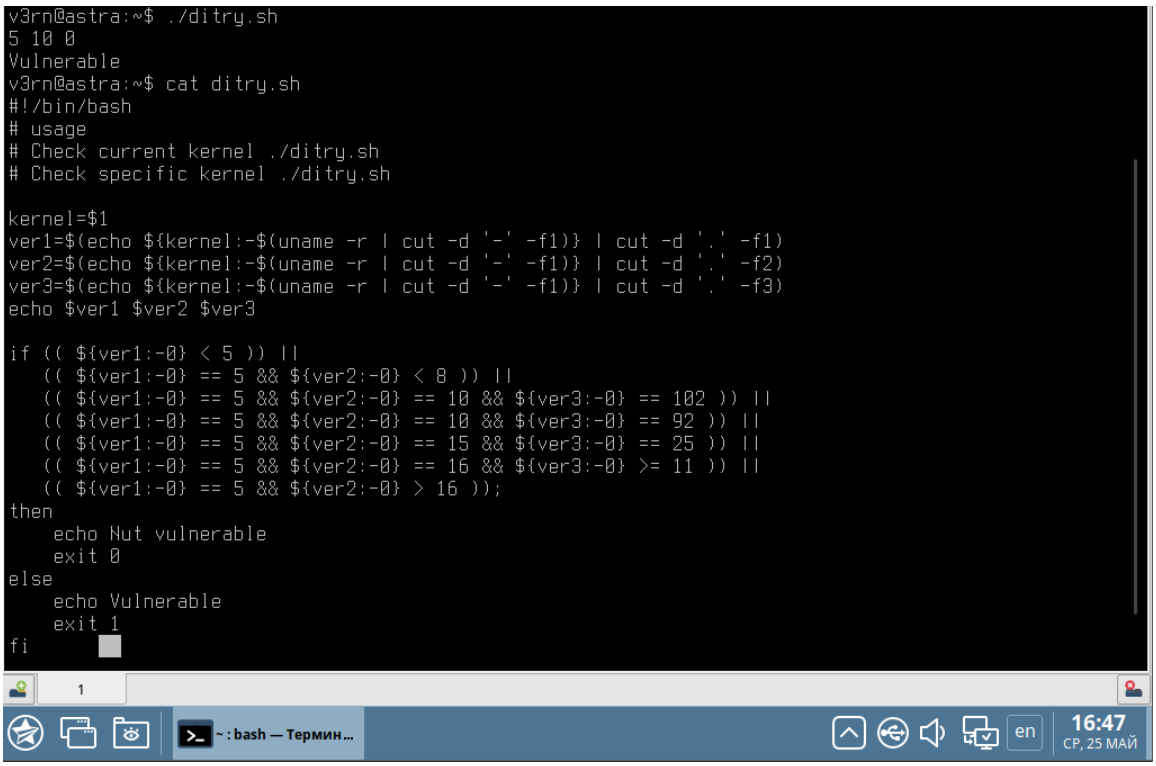

Рисунок 1 – Проверка версии ядра и попытка эксплуатации уязвимости.

Анализ возможных путей эксплуатации уязвимости позволил сформулировать следующие способы повышения безопасности информационной инфраструктуры:

— автоматическая проверка электронной цифровой подписи любого файла в системе для защиты от несанкционированного изменения;

— режим киоска, разрешающий запуск строго определенного набора приложений на уровне ядра;

— ограничение доступа пользователей к консоли;

— блокировка подключения «незнакомых» внешних устройств, препятствующая занесению в систему вредоносного программного обеспечения.

Список литературы:

1. Совбез одобрил проект основ госполитики в сфере инфобезопасности [Электронный ресурс]. – Режим доступа: https://ria-ru.turbopages.org/ria.ru/s/20220520/bezopasnost-1789775131.html.

2. Шива Парасрам, Алекс Замм. KaliLinux. Тестирование на проникновение и безопасность, 2016. -448 с.

3. Подходы по противодействию уязвимостям [Электронный ресурс]. – Режим доступа: https://habr.com/ru/company/astralinux/blog/461691/.

Захаров И.А., Зима В.М., Русанов Д.А.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

АВТОМАТИЗАЦИЯ ФОРМИРОВАНИЯ ЗАЩИЩЕННОЙ ТЕХНОЛОГИЧЕСКОЙ ПЛАТФОРМЫ ДЛЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ НА ОСНОВЕ СИСТЕМЫ КОНТЕЙНЕРИЗАЦИИ DOCKER

С увеличением масштабов работы и ростом сотрудников в организации, актуальным становится решение проблем, связанных с созданием собственных информационных систем (ИС), их масштабированием и обеспечением бесперебойной работы. При этом хотелось бы, чтобы разрабатываемые системы могли давать возможность гибкого использования ресурсов физических серверов, включая перераспределение нагрузки, независимость от аппаратной платформы, надежную и отказоустойчивую работу. В связи с этим становится неизбежным вопрос о внедрении и развертывании технологической платформы, на основе которой будет эффективно работать данная информационная система

В качестве решения предлагается использовать контейнерную виртуализацию, которая является технологией создания множества экземпляров изолированной программной среды со специфическим набором компонентов эмулируемой операционной системы (ОС). В отличие от аппаратной виртуализации с использованием гипервизора, при которой эмулируется аппаратное окружение и может быть запущен широкий спектр гостевых ОС, при контейнерной виртуализации в каждом контейнере может быть запущен экземпляр ОС только с тем же ядром, что и у хостовой ОС (все контейнеры узла используют общее ядро), и отсутствуют дополнительные ресурсные накладные расходы на эмуляцию виртуального оборудования и запуск полноценного экземпляра ОС, характерные для аппаратной виртуализации.

У контейнерной виртуализации по причине интенсивного развития этой технологии помимо более высокой производительности появились другие преимущества перед аппаратной (классической) виртуализацией, связанные с более высокой степенью автоматизации следующих процессов:

- разработка, включая тестирование, приложений-контейнеров;
- развертывание контейнеров;
- подержание безотказной работы приложений (сервисов) за счет кластеризации контейнеров;
- масштабирование – при превышении нагрузки на систему, платформа контейнеризации может увеличить число контейнеров на количество, указанное в настройках.

К наиболее апробированным, масштабируемым и автоматизируемым программным платформам контейнерной виртуализации относятся платформы, построенные на основе кодовой базы Open Source системы контейнеризации Docker. В отличие от Open Source платформ контейнеризации Kubernetes, OpenVZ и других, система контейнеризации Docker входит в состав последних версий сертифицированной по требованиям безопасности отечественной ОС Astra Linux Special Edition (версии 1.6, 1.7). В системе контейнеризации Dockerиспользуется декларативный подход на основе модели IaC – в файлах конфигурации (yaml-файлах) на близком для понимания языке YAML (Yet Another Markup Language) описывается не как нужно сделать, а что требуется достичь при разработке, развертывании, кластеризации и масштабировании контейнеров. Платформа контейнеризации на базе этого описания и доступных физических машин разворачивает контейнеры и делает все возможное для поддержания требуемой конфигурации. Работа системы становится самоорганизованной и прогнозируемой. Если все развернулось и работает в тестовом окружении, то будет работать и в рабочей среде.

Несмотря на все свои достоинства платформа контейнерной виртуализации Docker является достаточно сложной программной системой, для которой необходимо обеспечить автоматизацию процесса развертывания, а также конфигурирования и аудита для достижения высокого уровня защищенности технологической платформы.

Автоматизация процесса развертывания платформы Dockercостоит в том, что необходимо разработать bash-скрипты, которые автоматизируют следующие процессы:

1) разворачивание необходимого количества контейнеров на выделенных серверах;

2) объединение серверов с контейнерамив единый отказоустойчивый кластер;

3) назначение узла управления (manager node) и автоматизация процесса проведения тестов отказоустойчивой работы и распределения нагрузки между хостами.

Для достижения высокого уровня защищенности технологической платформы Docker необходима автоматизация ее конфигурирования, связанная с установкой требуемых по защищенности параметров настройки, а также автоматизация последующего аудита. Главная проблема автоматизации конфигурирования и аудита – высокоуровневоенеформализованное описание требований. Предлагаемое решение основано на использовании набора спецификаций протокола автоматизации управления данными безопасности SCAP (Security Content Automation Protocol) для автоматизации следующих процессов:

1) поиск и исправление уязвимостей;

2) автоматическая настройка конфигураций;

3) оценка уровня безопасности.

Используя язык XCCDF (eXtensible Configurtion Checklist Description Format) необходимо обеспечить контроль правильности конфигураций системы и ее приведение к эталонному состоянию. Правила XCCDF непосредственно не описывает, как именно должна проводиться проверка. Вместо этого в документе XCCDF содержатся ссылки на другие XML-документы (определения OVAL), которые, в свою очередь, должны содержать актуальные инструкции по выполнению этой проверки.

Развернутая в итоге проделанной работы защищенная технологическая платформа Docker будет способна обеспечить эффективную, отказоустойчивую, масштабируемую и безопасную работу тех систем, которые будут запущены на ее основе.

Списоклитературы

1. Bhat, Sathyajith. Practical Docker with Python. – Bangalore, Karnataka: Apress, 2018.

2. Hutten Dennis. Docker: Build. Ship. Run. –Amazon, 2021.

3. Ian Miell, Aidan Hobson Sayers. Docker in Practice second edition. –Manning, 2019.

4. Nickoloff Jeff. Docker in Action. – Manning, 2016.

5. МоуэтЭ. ИспользованиеDocker. –Москва: ДМК, 2017.

6. Ной Гифт, Кеннеди Берман, Альфредо Деза, Григ Георгиу. Python и DevOps: Ключ к автоматизации Linux. –СПб.:Питер, 2022.

Хабибуллин А.В., Гомон А.В., Компаниец Р.И.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

АЛГОРИТМ ИДЕНТИФИКАЦИИ ФУНКЦИЙ В БИНАРНЫХ ФАЙЛАХ НА ОСНОВЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Двоичный анализ позволяет использовать множество полезных приложений в области компьютерной безопасности, учитывая множество возможных ситуаций, в которых исходный код высокого уровня недоступен, утерян или неудобен в использовании. Например, обнаружение вредоносных программ, защита программного обеспечения от распространенных уязвимостей и реверс-инжиниринг протоколов наиболее полезны, когда задействованные процедуры могут напрямую работать с двоичными файлами.

Основная проблема бинарного анализа заключается в отсутствии семантической структуры высокого уровня в двоичных файлах, поскольку компиляторы удаляют ее из исходного кода в процессе компиляции. Авторы вредоносных программ часто идут еще дальше и запутывают свои выходные данные, пытаясь помешать исследователям провести любой возможный анализ.

Функции -базовая, но фундаментальная часть структуры всех программ, но большинство двоичных файлов представляют собой недифференцированную последовательность инструкций машинного языка без какой-либо информации о том, как части объединяются в функции. Следовательно, многие методы бинарного анализа, которые полагаются на информацию о границах функций, должны сначала попытаться восстановить ее с помощью идентификации функций. Например, идентификация функций может помочь в добавлении контроля целостности потока управления к двоичным файлам для надлежащего ограничения переходов. Точно так же декомпиляторы и отладчики должны знать расположение функций, чтобы предоставить пользователю полезный вывод [1]. Популярными инструментами, такими как IDAPro, Ghidra, для идентификации функций создаются сигнатуры запуска функций в виде взвешенных деревьев префиксов, но данный метод имеет относительно низкую точность. Для повышения точности идентификации функций предлагается новый подход, основанный на сверточной нейронной сети (CNN) [2].

Чтобы идентифицировать функции, для каждого байта в двоичном коде просматривается 10 байтов до него и 10 байтов после него, чтобы определить, является ли он началом или остановкой функции. При этом существуют глобальные функции, которые могут помочь определить границы функции. Например, опкоды функции call позволяют определить начало функции. Входными данными модели будет вектор, где каждое значение находится в диапазоне от 0 до 257. Результатом модели является матрица, в которой каждая строка содержит два значения — вероятность того, что байт будет началом функции, и вероятность того, что байт не будет началом функции. Архитектура модели изображена на рисунке.

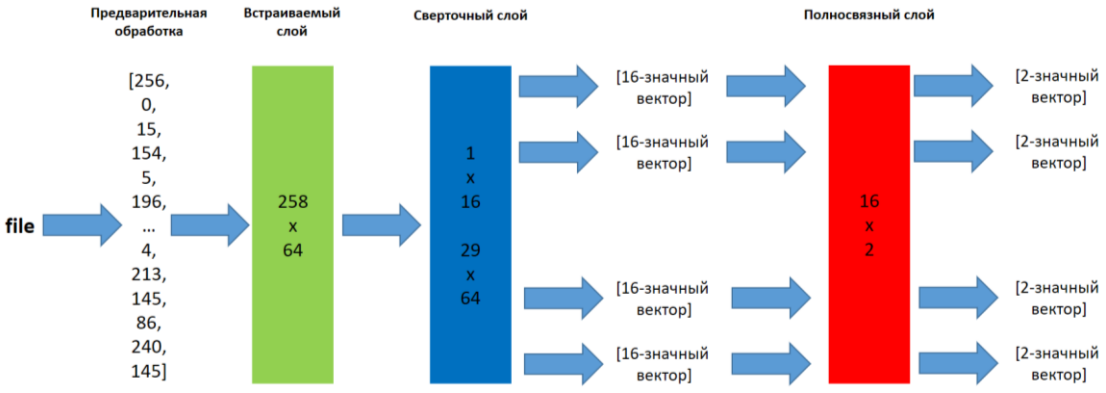

Рисунок 1 – Архитектура модели

Набор данных состоял из 2200 отдельных двоичных файлов, 2064 для linux, 136 для Windows. Для реализации модели использовался Pythonc библиотекой Theano [3]. Экспериментальные результаты приведены в таблице.

Таблица 1.Экспериментальные результаты

ELF x86			**ELF x86-64**		
P	R	F1	P	R	F1
97,75%	95,34%	96,53%	94,85%	89,91%	92,32%
PE x86			**PE x86-64**		
P	R	F1	P	R	F1
97,53	95,27%	96,39%	98,43%	97,33%	97,88%

Списоклитературы:

4. BaoT., BurketJ., WooM., TurnerR., Brumley D.Byteweight: Learning to Recognize Functions in Binary Code. In 23rd USENIX Conference on Security Symposium (SEC), C. 845–860.
5. SutskeverI.,VinyalsO. Sequence to sequence learning with neural networks. In Advances in Neural Information Processing Systems, C. 3104–3112.
6. GitHub – официальный сайт [Электронный ресурс]. URL: https://github.com/Theano/Theano (дата обращения 13.05.2022г.).

Житихин А.Е. , Андрушкевич С.С.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург

ОПРЕДЕЛЕНИЕ СТЕПЕНИ ВАЖНОСТИ ПРИЗНАКОВ ПРИ ВЫЯВЛЕНИИ КОМПЬЮТЕРНЫХ АТАК НА ОСНОВЕ АЛГОРИТМА FP-GROWTH

С ростом количества компонентов в информационных инфраструктурах и их сложностьюувеличиваются риски прекращения функционирования, а проблемы информационной безопасности становятся все более серьезными. Использование только традиционных методов защиты информации может быть неэффективным и может нанести ущерб интересам отдельных лиц и организаций.

Поскольку существует разрыв между существующими мерами информационной безопасности и текущими потребностями, крайне важно разрабатывать новые меры защиты. Хотя многие исследователи и специалисты предлагают решения, основанные на технологиях искусственного интеллекта, количество таких решений в организациях крайне мало,так как есть довольно много проблем в их разработке и управлении. В докладе представлен подход к выявлению компьютерных атак на основе ассоциативных правил.

Под получением информации о ситуации понимается процесс извлечения важной информации о состоянии информационной безопасностииз крупномасштабных источников данных, которая служит основой для оценивания и прогнозирования ситуации. Использование ассоциативных правил заключается в извлечении важных признаков компьютерных атак, описывающих взаимосвязь между данными. Есть два основных этапа выявления компьютерных атак на основе ассоциативных правил:

Этап 1. Поиск часто встречающихся признаков с минимальным уровнем поддержки.

Этап 2. Генерация ассоциативных правил с определенным уровнем достоверности.

В процессе генерации ассоциативных правилвысокий уровень поддержки выдвигает более высокие требования к производительности алгоритмов, а при низком уровне происходит генерация большого количества тривиальных правил. Более производительный алгоритм FP-Growth использует дерево поиска и сжатую структуру хранения данных. Используя этот подход, алгоритмFP-Growth снижает затраты на поиск, рекурсивно отыскивая и объединяя короткие паттерны.

Для проверки применимости алгоритма FP-Growthдля выявления компьютерных атак взят за основу набор данных KDD 99[1].Набор данных KDD 99состоит примерно из 4 900 000 отдельных соединений, каждое из которых содержит 41 признак и помечено как нормальный или как компьютерная атака. По результатам эксперимента алгоритм FP-Growth обеспечивает точность от 69% до 78% для разных соединений. Кроме того, он показывает очень низкий уровень ложных срабатываний около 3%.

Такимобразом, применение алгоритма FP-Growth может ускорить процесс выявления компьютерных атак.

Список литературы:

1. Olusola A. A., Oladele A. S., Abosede D. O. Analysis of KDD'99 intrusion detection dataset for selection of relevance features //Proceedings of the world congress on engineering and computer science. – WCECS, 2010. – Т. 1. – С. 20-22.

Шевченко С.В., Бурнаев О.Р., Ткаченко С.Ф.
Военно-космическая академия имени А.Ф. Можайского, г.Санкт-Петербург

РАЗРАБОТКА СЕРВИСА УСТАНОВЛЕНИЯ ИСТОЧНИКОВ И ПУТЕЙ РАСПРОСТРАНЕНИЯ ДЕСТРУКТИВНОЙ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ

Непрерывно в глобальной сети Интернет публикуется негативный контент, повсеместное распространение которого ведет к дестабилизации общественно-политической ситуации в Российской Федерации, популяризации материалов террористических и экстремистских организаций, призывам к массовым беспорядкам, осуществлению экстремистской деятельности, участию в массовых мероприятиях, проводимых с нарушением установленного порядка, совершению самоубийства, осуществлению пропаганды криминального образа жизни, потреблению наркотических средств и психотропных веществ, размещению иной противоправной информации [1]. Основным объектом такого деструктивного информационного воздействия является молодежь. По этой причине в октябре 2018 года по указу Президента был создан Центр изучения и сетевого мониторинга молодёжной среды (ЦИСМ), решающий эту проблему. Более того, последние события доказывают, что деструктивное информационное воздействие в социальных сетях может являться указателем на место и время планируемых преступлений, совершаемых в учебных заведениях. За последние 3 года зафиксировано 3 вооруженных нападения на учащихся ВУЗов и школ: в Керчи, Казани, Перми. Но и другие города не застрахованы от этого. Уже погибло 36 людей, более 100 пострадали.

В процессе исследования был разработан сервис, осуществляющий следующие этапы:

Этап 1. Сбор данных.

Этот этап проводится с помощью разработанных методик и технических средств сбора данных с помощью API социальных сетей, а также парсеровHTML-страниц. Результатом

сбора является большой объем взаимосвязанных данных о сетевой активности пользователя (частоте публикаций в зависимости от времени и дня), взаимодействиях с другими пользователями и Интернет-ресурсами. Также осуществляется сбор информационных постов в режиме реального времени по существующим трендам (хэштегам) и указанных на географической карте точки и радиуса мониторинга.

Этап 2. Обработка данных.

Этап предварительной обработки важен для последующего выявления деструктивного информационного воздействия в информационных постах пользователей. Он состоит из удаления стоп-слов, знаков препинания, спам-контента, не несущих какой-либо смысловой нагрузки. Затем оставшиеся слова необходимо привести к их исходным формам для следующего этапа.

Этап 3. Извлечение признаков негативного контента.

В рамках разработанного сервиса предлагается использовать не только текстовые признаки, но и признаки пользователей с учетом их взаимоотношений. Извлекаются следующие признаки:

1. признаки профиля - сведения о возрасте учетной записи, факта ее верификации, количестве статусов (т.е. числе твитов, включая ретвиты и цитаты, опубликованные пользователем и друзьями по переписке), число сообщений в списке (т.е. число общедоступных списков, участником которых является пользователь), а также число избранных сообщений (твиты, которые пользователь «лайкнул» за время существования своей учетной записи);

2. признаки, которые измеряют коммуникабельность пользователя на соответствующих платформах, например, число подписчиков и подписок, а также популярность на основе соотношения подписок и подписчиков, а также активности последних в обсуждении информационных постов;

3. признаки выявления активных распространителей негативного контента на основе извлечения перекрестных подписчиков популярных лидеров мнений;

4. признаки информационных сообщений пользователей по тематикам.

Этап 4. Выявление деструктивного информационного воздействия.

Пусть множество пар «информационное сообщение, класс негативного контента» $X \times Y$ является вероятностным пространством с неизвестной зависимостью X от Y. Имеется конечная обучающая выборка наблюдений $X^m = \{(x_1, y_1), \cdots, (x_m, y_{,m})\}$. Требуется построить алгоритм $a: X \to Y$, способный классифицировать произвольный объект $x \in X$, чтобы ошибка классификации $L \to \min$. Наиболее результативным при решении данной задачи оказался наивный байесовский классификатор (F1=0,92).

Таким образом, был разработан сервисустановления источников и путей распространения деструктивной информации в социальной сети Twitter. На данный момент разрабатывается поддержка других социальных сетей и мессенджеров, а также методика по нейтрализации негативного контента группой сгенерированных ботов. Серверная часть сервиса написана с использованием DjangoRestFramework, с помощью которого был разработан API, легко интегрируемый в сторонние системы. В качестве средства визуализации работы методов был использован прогрессивный JavaScript-фреймворк Vue.js, так как он позволяет создавать одностраничные приложения (SinglePageApplication, SPA), отвечающие принципам реактивности. Для хранения данных используется СУБД PostgreSQL, так как она является конкурентоспособной российской разработкой, что уменьшает риски по выводу из

строя программного комплекса в будущем. Для мониторинга социальных сетей используются следующие программные решения: Celery, Redis, Selenium, Prometheus и Graphana.

Список литературы:

1. Указ Президента Российской Федерации от 02.07.2021 № 400 «О Стратегии национальной безопасности Российской Федерации».

Бурнаев О.Р., Шевченко С.В., Ткаченко С.Ф.
Военно-космическая академия имени А.Ф. Можайского, г.Санкт-Петербург

МЕТОДИКА РАЗРАБОТКИ ЛИНГВИСТИЧЕСКИХ РЕСУРСОВ ДЛЯ ВЫЯВЛЕНИЯ ДЕСТРУКТИВНОГО КОНТЕНТА В СОЦИАЛЬНЫХ СЕТЯХ

Современные средства коммуникации между людьми (электронная почта, блоги, мессенджеры и социальные сети) расширили возможности по причинению умышленного вреда человеку или группе людей [1-3]. Данное деструктивное воздействие можно разделить на следующие типы:

- (виртуальное) преследование, домогательство – то есть неоднократные угрозы кому-либо;
- очернение/клевета – то есть разглашение ложной информации о ком-либо;
- оскорбления – то есть краткие оскорбительные онлайн-взаимодействия;
- распространение компромата.

В рамках проведенных исследований, под термином «деструктивное информационное воздействие» (далее ДИВ) будем понимать умышленное и (или) повторяющееся действие с целью причинить кому-либо вред или унизить его с помощью информационных и коммуникационных технологий, таких как мессенджеров, электронной почты и социальных сетей. Такие действия носят деструктивный характер и варьируются от эмоционального (гнев, страх, снижение морально-психологического состояния и т.д.) [4] и психологического вреда (заниженная самооценка, депрессия, суицидальные наклонности и т. д.) [5] до физического вреда (потеря сна, головная боль, расстройство пищевого поведения и т. д.) [6].

Несмотря на осуществления различных мер по обнаружению, предотвращению и ликвидации последствий деструктивных информационных воздействий на индивидуальное и групповое сознание [7-9], количество данных случаев увеличивается с каждым годом [11]. Учитывая сложность обнаружения деструктивных информационных воздействий по сравнению с более простыми типами негативного контента, такими как расизм (специально нацеленный на национальность или этническую принадлежность), женоненавистничество (например, ругательства, относящиеся к женщинам, или явное объявление женоненавистничества в профиле/биографии) или спам, особенно актуально достижение следующей цели: повышение эффективности обнаружения такого типа воздействий и их классификация по степени угрозы для дальнейшего предотвращения (отражения) и ликвидации их последствий. Угроза настолько критична, что для профилактики и ликвидации последствий ДИВ есть ряд российских [7-10] и международных инициатив [12-14].

Все подходы по профилактике и ликвидации последствий ДИВ должны:

1) повышать осведомленность о потенциальных угрозах ДИВ, основанной на потребностях пользователей ИТКС Интернет;

2) иметь адекватную систему оценивания угроз ДИВ и уровня защищенности от них;

3) повышать эффективность как мер реагирования (например, удаление, блокирование и опровержение сообщений ДИВ), так и превентивных мер (например, повышение прогноза радикализации информационной обстановки);

4) предоставить практические стратегии и инструментарии, которые позволят нейтрализовать ДИВ на групповое сознание.

Лингвистические ресурсы для выявления деструктивного контента - векторные модели, которые основываются на том, что так или иначе "считают" слова и их соседей, и на основе этого строят вектора для слов.

В процессе исследования социальной сети Twitterбыл собран набор из 135 хэштегов по 5 темам. В ходе изучения информационных постов по этим хэштегам было обнаружено, что вероятность обнаружить деструктивный контент варьируется от 0,03 до 0,7. Однако для эффективного автоматизированного обнаружения такого контента необходимо разработать лингвистические ресурсы, решающие следующие вопросы:

1. идентификация стиля речи;
2. идентификация вида предложения;
3. идентификация литературных приемов;
4. обработка риторических вопросов;
5. анализ графических материалов (смайлов);
6. скоринг информационных постов с учетом комментариев.

Часть информационных постов, которые можно отнести к классу деструктивного контента, содержат завуалированные призывы к несанкционированным митингам, подводу подростков к самоубийству и т.д. Именно поэтому необходимо учитывать такие признаки, как стиль речи (научный, официально-деловой, публицистический, художественный, разговорный), вид предложения (восклицательные, побудительные, вопросительные), используемые литературные приемы (эпитеты, метафоры, ирония, сарказм и т.д.). Также в особо эмоциональных диалогах в качестве аргументов, подтверждающих истинность позиции автора, часто используются риторические вопросы, которые также могут содержать завуалированный деструктивный контент. Также для общения в социальных сетях используются графические материалы (смайлы) и другие символы для выражения эмоций автора, однако помимо базового анализа тональности таких материалов необходимо высчитывать корреляцию между текстом пользователя и его эмоциями, так как возможные противоречия могут позволить выявлять субъектов взаимодействия в рамках треугольника Карпмана (психологическая модель деструктивных отношений, где каждый субъект имеет преимущество над одним, но уступает другому), что позволит в дальнейшем перейти к нейтрализации деструктивного контента.

Таким образом, необходимо разработать лингвистические ресурсы на базе разработанного программного комплекса по обнаружению первоисточников и путей распространения деструктивного контента в социальных сетях (программа для ЭВМ №2022618104) и проведенного исследования информационных постов в социальной сети Twitter.

Список литературы:

1. Шевко Н.Р., Исхаков И.И. Особенности проявления кибербуллинга в социальных сетях [Электронный ресурс] // Ученые записки Казанского юридического института МВД России. – 2017. – №3. – Режим доступа: https://cyberleninka.ru/article/n/osobennosti-proyavleniya-kiberbullinga-v-sotsialnyh-setyah (дата обращения: 19.02.2021).
2. Кирюхина Д.В. Кибербуллинг среди молодых пользователей социальных сетей // Современная зарубежная психология. – 2019. – Т. 8. – № 3. – С. 53-59.
3. Balakrishnan V., Khan S., Arabnia H. R. Improving cyberbullying detection using Twitter users' psychological features and machine learning //Computers & Security. – 2020. – Т. 90. – С. 101710.
4. Kowalski R. M. et al. Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth //Psychological bulletin. – 2014. – Т. 140. – №. 4. – С. 1073.
5. Hinduja S., Patchin J. W. Bullying, cyberbullying, and suicide //Archives of suicide research. – 2010. – Т. 14. – №. 3. – С. 206-221.
6. Bottino S. M. B. et al. Cyberbullying and adolescent mental health: systematic review //Cadernos de saudepublica. – 2015. – Т. 31. – С. 463-475.
7. О безопасности [Текст]: Федеральный закон от 28 декабря 2010 г. N 390-ФЗ // "Российская газета". – 2010. – № 295(5374) от 29 декабря 2010 г.
8. Доктрина информационной безопасности Российской Федерации [Текст]: Указом Президента Российской Федерации от 5 декабря 2016 г. №646 // "Российская газета". – 2016. – от 5 декабря 2016 г.
9. Основные направления исследований в области обеспечения информационной безопасности Российской Федерации [Электронный ресурс] / Совет Безопасности Российской Федерации // Режим доступа: http://www.scrf.gov.ru/security/information/document155/ (дата обращения: 19.02.2021).
10. Mail.ru Group объявила 11 ноября Днём борьбы с кибербуллингом в России [Электронный ресурс] – Режим доступа: https://corp.mail.ru/ru/press/releases/10510/ (дата обращения: 19.02.2021).
11. Чебакова, Дарья. Российские геймеры пожаловались на кибербуллинг [Электронный ресурс] / Д. Чебакова // РБК. 2021. – Режим доступа: https://www.rbc.ru/technology_and_media/18/02/2021/602dfc2b9a7947603febb92e, свободный.
12. Harris J. An evaluation of the use and effectiveness of the Protection from Harassment Act 1997. – London : Research, Development and Statistics Directorate, Home Office, 2000.
13. Vaillancourt T., Faris R., Mishna F. Cyberbullying in children and youth: implications for health and clinical practice //The Canadian journal of psychiatry. – 2017. – Т. 62. – №. 6. – С. 368-373.
14. Görzig A., Ólafsson K. What makes a bully a cyberbully? Unravelling the characteristics of cyberbullies across twenty-five European countries //Journal of Children and Media. – 2013. – Т. 7. – №. 1. – С. 9-27.

Менисов А.Б., Житихин А.Е., Казаков М.В.
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург
ИСП РАН, г. Москва

ПРОГНОЗИРОВАНИЕ СОСТОЯНИЯ ЗАЩИЩЕННОСТИ MLAAS ПРИ ТРАНСФЕРНОМ ОБУЧЕНИИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

** Исследование выполнено в рамках гранта Президента РФ для государственной поддержки молодых российских ученых – кандидатов наук (МК-2485.2022.4).*

Благодаря достижениям в области методов машинного обучения (МО) и глубокого обучения (ГО), а также потенциалу облачных вычислений стали популярными платформы Machine Learning as a Service (MLaaS)[1]. Кроме того, все чаще используются сторонние облачные сервисы для аутсорсинга обучения моделей ГО, что требует значительных дорогостоящих вычислительных ресурсов (например, высокопроизводительных графических процессоров (GPU)). Столь широкое использование облачных MLaaS открывает перед злоумышленниками широкий спектр векторов атак.

В практике МО и ГО часто используется трансферное обучение для дообучения уже существующих моделей МО и ГО на собственных данных. Хотя такиесистемы имеют явные преимущества,они могут быть легкоскомпрометированы [2].

Оценивание защищенности MLaaS является сложной задачей. Несмотря на большой объем исследований и практических методик защиты МО и ГО от различных атак, сложно комплексно оценить их надежность и сравнить их. В докладе представлено описание состояний системы: уязвимые (S); зараженные (I); иммунные (CI); частично иммунные (CP), и определены возможные переходы из одного состояния к другому.

Переходы определены угрозами безопасности информации, специфичных для MLaaS. Эти угрозы безопасностиинформации возникают на всех этапах таких систем:

— на этапе сбора данных для обучения моделей машинного обучения. Существуют примеры «отравления» неразмеченных данных, поступающих экспертам в прикладных областях для разметки, при этом «отравление» может осуществляться незаметной для эксперта модификацией данных. В результате обучения на «отравленных» данных модель может быть легко атакована злоумышленником, производившим «отравление», при помощи аналогичных модификаций;

— на этапе обучения и формирования обученных параметров модели машинного обучения, если злоумышленник может влиять на этот этап. Так, известны методы встраивания зловредного кода в обученные параметры нейронной сети, при этом сеть показывает хорошую точность на тестовой выборке. Зловредный код при этом может быть извлечен с помощью «вспомогательных» программ, созданных злоумышленником как дополнение к поставляемой модели;

— на этапе эксплаутации моделей, в том числе их обновления. В последние годы разрабатываются атаки черного ящика на нейросетевые модели машинного обучения с помощью состязательных примеров, сопоставимые по эффективности с атаками белого ящика.

Прогнозирование состояния MLaaS во время функционирования обеспечивает раннее обнаружение неисправностей и их устранение в процессе обслуживания. При решении такой задачи могут быть использованы методы машинного обучения, предназначенные для классификации. В данной статье в качестве исходных данных рассматриваются известные результаты (прецеденты) оценки состояния системы. Используется несколько различных подходов к классификации: классические статистические модели, методы, специально

ориентированные на машинное обучение, композиционные методы и другие. Для повышения качества прогнозирования может быть использован агрегированный подход – комбинация нескольких методов классификации.

Список литературы:

1. Wang C. et al. MIASec: enabling data in distinguish ability against membership inference attacks in MLaaS //IEEE Transactions on Sustainable Computing. – 2019. – Т. 5. – №. 3. – С. 365-376.
2. Qayyum A. etal. Securing machine learning in the cloud: A system a tic review of cloud machine learning security //Frontiers in big Data. – 2020. – С. 43.

Кришталь И.В., Давыденко В.С., Первушин А.В., Нагибин Д.В.
Военно-космическая академия имени А. Ф. Можайского, г. Санкт-Петербург

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ЗАЩИЩЁННОГО ОБМЕНА ТЕКСТОВОЙ И МУЛЬТИМЕДИЙНОЙ ИНФОРМАЦИЕЙ

В настоящее время невозможно быть уверенным в защищенности данных, передаваемых с помощью популярных мессенджеров. Осуществление внутренней (корпоративной) переписки и передачи служебных документов – сложные задачи с точки зрения информационной безопасности (ИБ).

Анализ функционирования мессенджеров, действующих на территории России, показал наличие следующих проблем ИБ: подмена данных [1], несовершенство систем авторизации, аутентификации и регистрации пользователей в системе [2], хранение данных на зарубежных серверах [3].

Целью работы являлось обеспечение надежности, безопасности и конфиденциальности информации, передаваемой с помощью мессенджеров.

В результате был разработан программный комплекс обмена текстовой и мультимедийной информацией («Агат»), который позволил разрешить указанные выше по тексту проблемы. В нём реализовано две схемы шифрования:

1. «человек-сервер-человек» – для обычных сообщений;

2. «человек-человек» («сквозное шифрование») – для секретных чатов.

В первом случае информация передается и шифруется по схеме, основанной на протоколе HTTPS, но с использованием отечественного криптографического алгоритма симметричного шифрования «Кузнечик» (ГОСТ 34.12-2018) [4, 5] с хеш-функцией «Стрибог» (ГОСТ 34.11-2018) [6, 7]. Во втором случае обмен основывается на протоколе MTProto (Telegram), также на российских алгоритмах.

Распределение ключей осуществляется по схеме Диффи-Хеллмана (DH) [8]. Для пресечения атаки «человек посередине» (MITM-атака) закрытые ключи генерируются следующим образом [9]:

1) Каждый участник генерирует 2048 бит, пусть это будет r_c.

2) Сервер генерирует 2048 бит (r_s) и рассылает их всем участникам диалога.

3) Участники вычисляютзакрытые ключи:$r_c \oplus r_s$, где $\oplus$ – операция XOR.

Далее собеседникивырабатывают общий секретный ключ.

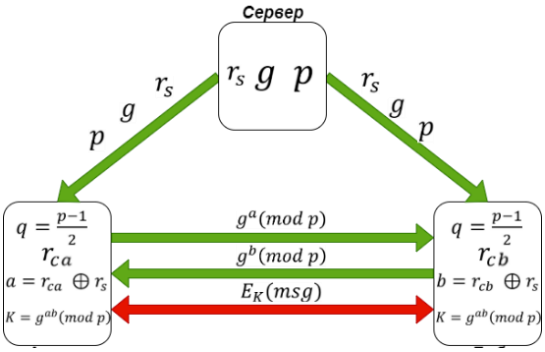

Рисунок 1 – Схема передачи сообщений в секретном чате

Затем сообщения шифруются алгоритмом «Кузнечик». При работе по схеме «человек-человек» «следов» на сервере не остаётся, т.к. он используется только при генерации открытых ключей. Схема создания секретного чата между двумя участниками представлена на рисунке 1.

Архитектура приложения состоит из двух частей: серверной и клиентской. Сервер написан с использованием СУБД Postgres, фреймворка Django и библиотек: REST Framework, TWJ – token, WebSocket, Channels.

У клиента общая схема приложения реализована на React-native, версия для Android – наязыке программирования Kotlin, для IOS–на Swift.

Результаты сравнения разработанного программного комплекса «Агат» с аналогами представлены в таблице 1.

Таблица 1. Результаты сравнительного анализа криптографических протоколов известных мессенджеров

Название	**Страна-разработчик**	**Протокол безопасности**	**Наличие сквозного шифрования**	**Длина ключей шифрования**		**Схема распределения ключей**
				a*	**k****	
Агат	РФ	MTProto (отеч. криптоалг.)	+	2048	256	DH
Telegram	Международная команда	MTProto	+	2048	128	RSA
Signal	США	Signal Protocol	+	256	256	ECDH
Threema	Швейцария	NaCl (шифрXSalma20)	-	256	128	ECDH
VIPole	Великобритания	HTTPS (AES-256 + RSA)	-	3072	256	RSA
Wickr	США	Wickr Messaging	+	256	256	ECDH

* - длина ключа алгоритма распределения ключей в битах.

** - длина ключа алгоритма симметричного шифрования в битах.

В дальнейшем планируется увеличить криптостойкость предложенной схемы шифрования за счет применения эллиптических кривых для вычисления общего секретного ключа.

Список литературы

1. Уязвимость Media File Jacking [Электронный ресурс]. – URL: https://symantec-enterprise-blogs.security.com/blogs/expert-perspectives/symantec-mobile-threat-defense-attackers-can-manipulate-your-whatsapp-and-telegram-media(дата обращения: 15.05.22).
2. Атака на систему аутентификацииWhatsApp [Электронный ресурс]. – URL: https://zen.yandex.ru/media/id/5e459cf16e1cd54e7a5c7b53/whatsapp-vzlomali-kak-prestupniki-vzlamyvaiut-messendjer(дата обращения: 15.05.22).

3. Беликов Ю.В. Система криптографической защиты в мессенджерах // Современные проблемы проектирования, применения и безопасности информационных систем в цифровой экономике. – 2018. – С. 43-46.

4. Марков Е.С., Маро Е.А. Программная реализация криптографического алгоритма «Кузнечик» // Информационное противодействие угрозам терроризма. – 2015. – №. 24. – С. 292-299.

5. ГОСТ 34.12-2018 Межгосударственный стандарт. Информационная технология. Криптографическая защита информации. Блочные шифры. М: Стандартинформ. – 2019. – 12 с.

6. Диченко С.А. Анализ построения функции хэширования семейства «Стрибог» // Информационный бюллетень Омского научно-образовательного центра ОмГТУ и ИМ СО РАН в области математики и информатики. – 2018. – С. 71-73.

7. ГОСТ 34.11-2018 Межгосударственный стандарт. Информационная технология. Криптографическая защита информации. Функция хэширования. М: Стандартинформ. – 2019. – 23 с.

8. Давтян А.В. Система открытого распределения ключей Диффи-Хеллмана // Лучшая исследовательская работа 2021. – 2021. – С. 76-81.

9. Мозолевский А.А. Принцип шифрования сообщений мессенджера Telegram // Перспективы развития науки и образования. – 2018. – С. 133-136.

Пилькевич С.В., Кайзер М.С.
Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

РАСПОЗНАВАНИЕ И БЛОКИРОВАНИЕ НЕГАТИВНОГО КОНТЕНТА СРЕДСТВАМИ ИНТЕРНЕТ-БРАУЗЕРА

В информационном пространстве наблюдается тенденция увеличения негативного контента. По данным ТАСС [1]в 2020 году Роскомнадзор заблокировал порядка 500 тысяч материалов с деструктивной информацией, а общее количество жалоб граждан на противоправный контент выросло в 16 раз по сравнению с 2013 годом. Сложившаяся ситуация в сети создает необходимость в оперативном блокировании такой информации. Использование интернет-браузера пользователями, как средства поиска информации создает возможность разработчику использовать API браузера для точечной и быстрой блокировки негативного контента, а применение различных методов машинного обучения позволит реализовать автоматическое распознавание.

На сегодняшний день негативный контент блокируется следующими способами:

- по IP-адресу и протоколу;
- с помощью технологии DPI;
- по URL-адресу;
- с помощью DNS-сервера [2];

Краткий анализ перечисленных способов показал следующее. Блокировка по IP-адресу и DNSне обеспечивают избирательности блокировки, так как они блокируют весь контент на сервере. Кроме того, эти способы блокировки можно обойти с помощью использования

прокси-серверов и VPN. Фильтр URL может обеспечить избирательность на уровне элементов веб-страниц, однако при применении многоуровневого шифрования можно обойти этот механизм. Технология DPI высоко избирательна, так как она затрагивает только контент, соответствующий правилам блокировки, при условии эффективности этих правил. В противном случае точность распознавания контента уменьшается. Необходимо отметить, что, DPIможно обойти, использовав многоуровневое шифрование. Общей чертой вышеперечисленных методов является использование промежуточного устройства между устройством и сервером, что влечет расходы на аппаратную составляющую этого устройства.

Обобщая недостатки, указанные выше, можно предложить реализацию блокировки в форме расширения для браузера. Анализ структуры просматриваемой веб-страницы, позволит выделить элементы, потенциально содержащие негативный контент, что обеспечит высокую избирательность, и сделает блокировку невосприимчивой к шифрованию трафика. Использование API браузера позволит эффективно реализовать поиск таких элементов. Общая схема системы блокирования негативного контента средствами интернет-браузера представлена на рисунке 1.

Программа, блокирующая негативный контент должна состоять из плагина, контролирующего приложения и сервера. Плагин выполняет задачи распознавания и блокировки. Обращает на себя внимание то, что размещение классификатора на стороне пользователя (в плагине) не только сократит время на обработку элемента, потенциально содержащего негативный контент, но и избавит разработчика от издержек на аппаратную часть сервера, который выполняет задачи классификации. Контролирующее приложение должно иметь доступ к блокируемым категориям на каждом плагине и заблокированном контенте. Связующим звеном между расширением и контролирующим приложением является сервер. На нём ведутся две базы данных, в первой записаны идентификатор контролирующего приложения и закрепленные за ним плагины, а во-второй идентификатор плагина и заблокированный им контент. Передачу данных между контролирующим приложением и расширением следует реализовать с использованием TCP сокета. Хотя сокет и ограничивает количество подключений, но обеспечивает быструю передачу данных.

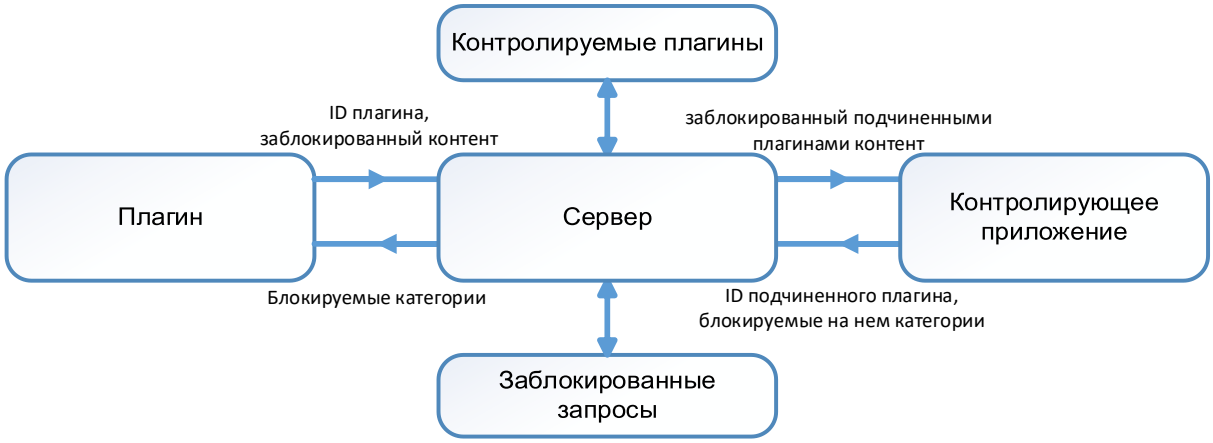

Рисунок 1 – Основные элементы системы блокирования негативного контента на основе интернет-браузера

Для макетирования рассматриваемого подхода используем приложение основным функционалом которого является блокировка негативного графического контента. Контент при этом разделен на три класса: «нейтральные» изображения, «порнография» и «кровавые фотографии». Классификатор реализован нейросетью, энкодер которой – предобученный на ImageNetMobilenet, классификатор – полносвязная нейронная сеть,состоящая из 3х слоёв (по 128, 32 и 3 нейронов соответственно, вероятность исключения нейрона в течении одной эпохи

равна 0,65), обучающая выборка состоит из 1559 изображений. На валидационной выборке получены следующие метрики: доля правильных ответов - 0.8558, точность - 0.8678, доля распознанных обьектов класса из всех обьектов класса- 0.8205. Их можно улучшить, составив более репрезентативную выборку. Среднее время распознавания одного изображения на «слабом» компьютере (CPU - AMD A6-9225 RADEON R4, 4Гб RAM) составило 5 секунд, а на «мощном»(CPU - Ryzen5 4600H, GPU - GeForceGTX1650Ti, 16Гб RAM) - 1,7 секунд. При тестировании приложения на веб-странице (https://2ch.hk/gg/), которая содержала 70 изображений, «мощный» компьютер обработал изображения за 2 минуты 4 секунды, «слабый» за 5 минут 17 секунд. Как видно, скорость обработки изображений на «слабом» компьютере затрудняет комфортное использование браузера. Исходя из результатов тестирования понятно, что приложение устанавливает требования к аппаратной составляющей ПК, что сужает потенциальный круг пользователей.

Подводя итоги, реализация блокирования негативного контента средствами интернет-браузера позволяет достичь высокой избирательности, так как с помощью APIбраузера выделяются конкретные HTML-элементы, потенциально содержащие негативный контент, гибкого контроля блокируемых категорий, но требует производительной аппаратной части ПК пользователя.

Список литературы:

1. Новости в России и мире - ТАСС. URL: https://tass.ru/obschestvo/11482847
2. Internet Society — обзор перспектив блокировки интернет-контента. 2017. URL:https://www.internetsociety.org/wp-content/uploads/2017/03/ISOC-ContentBlocking Overview_ru.pdf
3. Бирюков Д.Н., ЛомакоА.Г., Петренко С.А. Интеллектуальные системы предотвращения кибератак // The 2019 Symposium on Cybersecurity of the digital economy - CDE'19. – С. 184-195.

Пилькевич С.В., Ковальчук В.С.

Военно-космическая академия имени А.Ф. Можайского, г. Санкт-Петербург

К ВОПРОСУ О БЕЗОПАСНОСТИ ТЕХНОЛОГИИ «СМАРТ-КОНТРАКТ»

В современном мире, в результате научно-технического прогресса, появляются технологии, позволяющие решать задачи в различных областях,одной из них является смарт-контракт. Данная технология реализует автоматизацию процессов, для которых характерны высокий уровня доверия и отсутствие человеческого фактора. В частности, реализация системы финансовых взаимоотношений; организация прозрачного и, в тоже время, анонимного голосования или опроса; повышение уровня доверия к товару, посредством достоверного знания о цепочке поставок.

Смарт-контракт – это технология, реализующая протокол, позволяющий проводить сделки и контролировать их исполнение с помощью математических алгоритмов. Он составляется с помощью компьютерной программы.

Использование технологии блокчейн, в смарт-контракте, повышает устойчивость к мошенничеству и распространению искаженной информации. В свою очередь блокчейн – это децентрализованная база данных, которая одновременно хранится на множестве компьютеров, соединённых друг с другом в интернете и состоит из последовательно выстроенной цепочки цифровых блоков, в каждом из которых содержится информация о предыдущем и следующем блоках.

Использование смарт-контракта на блокчейн платформе приводит к некоторым особенностям. Одной из них, является своеобразная плата узлам (звеньям цепочки) комиссии, выплачиваемой во внутренней валюте, за запись данных в блок. Такую плату называют «газ» [1]. Плата за газ обусловлена компенсацией вычислительной энергии (вычислительных ресурсов узла) и производится пользователями, являясь необходимым элементом технологии, потребным для обработки и проверки транзакций.

Также, особенностью является определение достоверных данных, получаемых из внешней сети. Блокчейн позволяет сохранять информацию неизменной, но, при ее получении из вне, требуется, проверка поступающих данных. За решение данной проблемы отвечают так называемые «оракулы» [2]. Они бывают как обычными API сервисами, так и смарт-контрактами.

В последнее время технология смарт-контракт приобрела широкое распространение. С помощью данной технологии, выполняются действия с критическими данными, в следствие чего необходимо обеспечение высокой безопасности данного класса протоколов.

Наличие уязвимостей в реализации технологии смарт-контракт, в случае ее эксплуатации нарушителем, может послужить причиной значительного ущерба пользователям. Для недопущения подобного, разработчикам необходимо знать о существовании подобного рода ошибок и избегать их при написании программ.

Разработка смарт-контракта на блокчейн платформе Etherium состоит из следующих этапов:

а) составление на языке Solidity;

б) компилирование кода программы;

в) исполнение на виртуальной машине Etherium.

Рассмотрение известных уязвимостей[3], позволило провести их классификацию по типам эксплуатации:

а) эксплуатация логики контрактов;

б) эксплуатация архитектуры;

в) эксплуатация доверенных источников.

К первому типу относятся уязвимости: код программы которых содержит логику, приводящую к непредвиденным действиям с внешними контрактами; использование критических переменных, значения которых являются функциями от времени или других, легко прогнозируемых переменных; применение арифметических операций, в результате которых возможно переполнение целочисленного значения.

Ко второму типуотносятся такие ошибки эксплуатация которых обусловлена особенностями технологии блокчейн. Например, плата за газ и его лимит, в пределах одного блока, и видимость операции в течение короткого промежутка времени перед ее выполнением. Примером атаки является манипулирование значением газа для вызова отказа в обслуживании, которое не позволяет смарт-контракту выполнять обязательства перед участниками, а также совершение атаки на опережение, заключающейся в более быстрой записи данных в блок, чем у другого участника, даже если изначально запрос у участника был раньше.

Третий тип возникает при использовании смарт-контрактом некорректных данных от оракулов. Программа, неправильно обрабатывающая информацию или получающая неверный поток данных, автоматически продолжать выполняет действия. Устаревшее или даже

вредоносное содержимое, может иметь катастрофические последствия для всех процессов, связанных с потоком данных.

Существуют различные векторы атаки: от применения метода грубой силы, до захвата оракула, если он является API сервисом. Для некоторых типов, злоумышленник должен потратить свои активы, что делает атаку нецелесообразной если, затраты превышают «награду».

Для обеспечения безопасности смарт-контракта разработчики должны учитывать особенности технологии блокчейн и обеспечивать корректную обработку данных, получаемых из внешней сети. Код программы хранится в децентрализованной базе данных, из-за чего изменение логики, после размещения смарт-контракта, невозможно. Данные требования повышают сложность разработки и реализации протокола, вынуждают тщательно продумывать логику программы, а перед размещением в виртуальной машине проводить независимый аудит.

Список литературы

1. SolidityDocs – официальный сайт [Электронный ресурс]. URL: https://docs.soliditylang.org/en/v0.8.14/introduction-to-smart-contracts.html (Дата обращения 14.05.2022г.)
2. Ethereum – официальный сайт [Электронный ресурс]. URL: https://ethereum.org/en/developers/docs/oracles/ (Дата обращения 14.05.2022г.)
3. EthereumSmartContractSecurity – официальный сайт [Электронный ресурс]. URL: https://consensys.github.io/smart-contract-best-practices/attacks/ (Дата обращения 14.05.2022г.)

Ткачева Е.И. Калинин М.О.

Санкт-Петербургский политехнический университет Петра Великого

ВЫЯВЛЕНИЕ КИБЕРУГРОЗ В СИСТЕМАХ ИНТЕРНЕТА ВЕЩЕЙ С ИСПОЛЬЗОВАНИЕМ МНОГОАГЕНТНОГО ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

** Работа выполнена в рамках Государственного задания на проведение фундаментальных исследований (код темы 0784-2020-0026*

На сегодняшний день технология Интернета вещей становится неотъемлемой частью современной информационной инфраструктуры общества. Из-за быстрого роста и количества подключаемых устройств технология Интернета вещей предоставляет большие возможности нарушителям безопасности.

Системы обнаружения вторжений, используемые для защиты сетевых инфраструктур, зависят от ряда условия: изменения поведения контролируемой системы, необходимости переобучения и настройки детекторов, зависимость от наличия и полноты обучающих выборок [1]. Разновидность машинного обучения – обучение с подкреплением – позволяет решить данную проблему с помощью постоянной подстройки под меняющиеся параметры среды и создать кооперации между устройствами-агентами обучения.

В исследовании рассматривается подход многоагентного обучения с подкреплением для реализации обнаружения нарушений безопасности в системах Интернета вещей. Многоагентная система отличается от обычной тем, что в среде функционирует не один агент, а несколько [2]. На каждом конечном устройстве-жертве находятся агенты, которые и

представляют из себя элементы системы обнаружения вторжений. В качестве алгоритма обучения реализован Q-Learning [3, 4].

Разработано три типа многоагентного обучения с подкреплением: полностью децентрализованное, с передачей информации о прогнозах, с передачей информации о наблюдениях. Полностью децентрализованное обучение агентов представляет собой среду, где агенты работают независимо друг от друга и не обмениваются информацией. Каждый агент принимает решения на основе своих наблюдений без агрегирования данных. Обучение с передачей информации о прогнозах – среда с каналом связи между агентами, где агенты обучаются на своих наблюдениях, рассматривая других агентов как часть окружающей среды. Агент посылает 1, если обнаружил атаку, 0 – в противном случае. Обучение с передачей информации о наблюдениях – среда с каналом связи, через который агенты передают часть своих наблюдений (под наблюдением подразумевается один из выходов нейронной сети, детектирующей признаки атаки).

В практической части исследования реализованы все рассмотренные варианты многоагентного обучения. Параметры обучения подобраны экспериментальным способом. Система Интернета вещей построена на стандарте IEEE 802.15.4, поверх которого накладывается протокол UDP. Реализован модельный сценарий атаки UDP-флуд как пример для обучения агентов. Обученные системы были сопоставлены с известной системой обнаружений вторжений Suricata, в которых используется традиционный сигнатурный метод (таблица 1).

Все три варианта многоагентного обучения с подкреплением уступают сигнатурному методу по точности, что является основным недостатком обнаружения на основе обучения с подкреплением, однако, многоагентно-обученные системы имеют лучшие показатели полноты (F-меры), что показывает адаптивность и полноту обучения таких систем. Наибольшей полнотой (90%) характеризуется архитектура многоагентного обучения с передачей информации о прогнозах.

Таблица 1 – Сопоставление систем, основанных на многоагентном обучении с подкреплением, и традиционного сигнатурного метода

Архитектура системы	*Recall*	*Precision*	*F-мера*
Децентрализованная	0.91	0.82	0.87
С передачей информации о прогнозах	0.90	0.90	0.90
С передачей информации о наблюдениях	0.90	0.73	0.81
Традиционная (СОВ Suricata, сигнатурное обнаружение)	0.93	0.73	0.82

Обучение с подкреплением позволяют системе обнаружений вторжений быть более полной. Новый подход подразумевает сбор и учет при обучении неполных входных данных, что характерно для Интернета вещей. Такая обучаемая система обнаружения вторжений быстрее адаптируется к изменениям контролируемой среды, так как способна доучиваться в процессе работы и с большой вероятностью обнаруживать неизвестные атаки.

Список литературы:

1. Hamdan A. Intrusion Detection System: Overview / Hamdan. O. Alanazi, Rafidah Noor, B.B Zaidan, A.A Zaidan // Journal of Computing. – 2010. – P. 130-133

2. Servin A. Multi-Agent Reinforcement Learning for Intrusion Detection / A. Servin // Adaptive Agents and Multi-Agent Systems III. Adaptation and Multi-Agent Learning. – 2008. – P. 211-223.

3. Thanh Thi N. Deep Reinforcement Learning for Cyber Security / N. Thanh Thi, R. Vijay Janapa. // IEEE Transactions on Neural Networks and Learning Systems. – 2021. – P. 1-17.

4. Alavizadeh H. Deep Q-Learning based Reinforcement Learning Approach for Network Intrusion Detection / H. Alavizadeh, J. Jang-Jaccard. // Cyber Security Research Programme – Artificial Intelligence for Automating Response to Threats. – 2021. – P. 1-12

Богатов Г.В., Александрова Е.Б.
Санкт-Петербургский политехнический университет Петра Великого

КРИПТОГРАФИЧЕСКАЯ ЗАЩИТА В ВЕБ-ПРИЛОЖЕНИИ СПОРТИВНОГО СКАУТИНГА

С развитием новых цифровых технологий в различных сферах деятельности человека возникает задача обеспечения безопасности данных, которые система обрабатывает и предоставляет своему пользователю. При разработке современных веб-приложений, особенно в тех, где одним из главных пользователей является несовершеннолетний ребенок, соблюдение принципов конфиденциальности, целостности и доступности становится сложной задачей не только с технической и организационной, но и с правовой точки зрения.

На сегодняшний день остро стоит проблема взаимодействия спортивных организаций и молодых спортсменов, у огромного количества детей нет даже шанса быть замеченными. Для решения данной проблемы предлагается использовать веб-приложение со следующими участниками:

- спортсмен (несовершеннолетний ребенок до 18 лет);
- родитель (законный представитель ребенка-спортсмена);
- спортивная организация, представленная юридическими лицами;
- скаут (спортивный агент), официальный представитель спортивных организаций;
- администратор.

Пользователи типа «спортсмен» в системе имеют возможность выкладывать видео различного характера, которые, по их мнению, демонстрируют уровень мастерства в конкретном виде спорта. Пользователи «скаут» имеют возможность просматривать видео всех спортсменов. Если пользователь «скаут» оценил возможности пользователя «спортсмен», он приглашает данного пользователя в личный чат, в котором будут присутствовать пользователи «скаут», «родитель»-«спортсмен». В личном чате происходит обсуждение и приглашение несовершеннолетнего ребенка на просмотры в спортивный клуб, решаются организационные вопросы и передача нужных для этого документов. Если сторонам удается договориться о встрече, то все данные будущей встречи сохраняются в блокчейн.

Для обеспечения безопасности и соблюдения законов (ст. 9 закона «О персональных данных» и ч.1 ст.64 Семейного кодекса) все взаимодействие между пользователями происходит только в рамках данной системы. Использование любых сторонних мессенджеров не гарантирует пользователю конфиденциальность и целостность данных и не аутентифицирует получателя сообщения. Процесс аутентификации каждого пользователя производится с привлечением «администратора», который после регистрации принимает необходимые документы с информацией о пользователях и сверяет их с данными, которые заполнил пользователь. Если полученные администратором документы подлинные и администратор аутентифицировал нового пользователя, то в зависимости от типа каждого

пользователя генерируется и присваивается необходимая электронная цифровая подпись [1]. Пользователи «спортсмен» и «родитель» снабжены личными ключами электронной цифровой подписи, которая позволяет системе автоматически аутентифицировать пользователя, а также обеспечивает целостность передаваемых данных. Для пользователей типа «спортивная организация» генерируется групповая электронная цифровая подпись (так как для организации аутентификации от имени группы достаточно одного участника), которая формируется пользователями, входящими в определенную группу (официальные лица клуба) [2]. Для всех групп проверяющим будет являться «администратор», который всегда будет иметь возможность удостовериться, что подпись сформирована данной группой лиц или от имени группы. Таким образом, в каждую группу можно добавлять аутентифицированных пользователей типа «скаут». Наличие процедуры аутентификации позволяет защититься от атаки «человек посередине» при реализации личного диалога между пользователями «скаут» и «родитель»-«спортсмен».

Протокол Диффи-Хеллмана на эллиптических кривых позволяет двум сторонам, имеющим пары открытый/закрытый ключ, получить общий секретный ключ, используя незащищенный от прослушивания канал связи. Перед генерацией необходимых параметров проверяются электронные цифровые подписи каждого участника диалога, в случае подтверждения всех участников происходит обмен открытыми ключами и вычисления общего секретного ключа для каждой из сторон. Защита личных сообщений обеспечивается симметричным алгоритмом блочного шифрования. После согласования всех организационных вопросов пользователь «скаут» заполняет в системе форму, которую обязан подтвердить пользователь «родитель». После успешного подтверждения обеих сторон, данная форма записывается в блокчейн, чтобы никто из сторон не имел возможности изменять данные без обоюдного согласия.

Реализация данного приложения позволит значительно улучшить взаимодействие между спортивными агентами и несовершеннолетними спортсменами. Выбор современных, в том числе постквантовых, криптографических методов защиты информации, позволяет обеспечить высокую степень защиты передаваемых в веб-приложении данных.

Список литературы:

1. Boldyreva A. et al. Ordered multisignatures and identity-based sequential aggregate signatures, with applications to secure routing // 14th ACM Conf. on Computer and Communications Security. – ACM, 2007. – С. 276-285.
2. Кузьмин А.С., Сабанов А.Г. Анализ зарубежной нормативной базы по идентификации и аутентификации. Инженерный журнал: наука и инновации. – 2013. – 11 (23). – С.1-13.

Васильев А.А

Институт Системного Программирования РАН им. Иванникова, Москва

ИСПОЛЬЗОВАНИЕ СТАТИЧЕСКОЙ ВЕРИФИКАЦИИ ДЛЯ ОБНАРУЖЕНИЯ ОШИБОК ДОСТУПА К ПАМЯТИ НА ПРИМЕРЕ ДРАЙВЕРОВ ОПЕРАЦИОННОЙ СИСТЕМЫ LINUX

Статическая верификация - технология для автоматического доказательства соответствия программы заданным свойствам. Побочным эффектом от применения инструментов статической верификации является выдача трасс, нарушающих проверяемое свойство.

Под свойством корректного доступа к памяти для языка Си мы понимаем отсутствие ошибок менеджмента памяти и разыменований: утечки памяти, использование освобожденной памяти, повторное освобождение, чтение/запись за пределами выделенной памяти, разыменование нулевого указателя. Обнаружение всех таких ошибок в системном программном обеспечении является очень сложной задачей, но наличие даже одной неисправленной ошибки может привести к потенциальной уязвимости. Функционал верификации корректного доступа к памяти реализован в анализе SMGCPA [1] инструмента верификации CPAchecker [2].

Для анализа драйверов ядра ОС Linux используется фреймворк Klever [3] для статической верификации. Фреймворк обеспечивает автоматическое построение модели окружения и сценариев использования драйвера. Результатом применения стало исправление более 50 ошибок доступа к памяти в драйверах ядра ОС Linux. При тщательном анализе предупреждений инструмента выясняется, что истинными ошибками является примерно 10 % рассматриваемых трасс в основном из-за неточности автоматически сгенерированного окружения, которое можно при необходимости дополнительно уточнить.

Список литературы:

1. Васильев А.А., Мутилин В.С. Анализ корректности работы с памятью с использованием расширения теории символьных графов предикатами над символьными значениями. Труды Института системного программирования РАН. 2019;31(6):7-20.

2. Beyer D., Henzinger T., Theoduloz G. Program Analysis with Dynamic Precision Adjustment // Automated Software Engineering, 2008. ASE 2008. 23rd IEEE/ACM International Conference on. — Сент. 2008. — С. 29—38.

3. E. Novikov, I. Zakharov. Towards automated static verification of GNU C programs. In: Petrenko A., Voronkov A. (eds) Proceedings of the 11th International Andrei Ershov Memorial Conference on Perspectives of System Informatics (PSI'17), LNCS, volume 10742, pp. 402–416. Cham, Springer, 2018.

Водяной Д.А., Жуковский Е.В.
Санкт-Петербургский политехнический университет Петра Великого

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ ОБНАРУЖЕНИЯ УЯЗВИМОСТЕЙ В ПРОГРАММНОМ ОБЕСПЕЧЕНИИ БЕЗ ИСХОДНЫХ КОДОВ

Методы машинного обучения позволяют решать большой спектр задач, в том числе, в области применения их для обнаружения ошибок и уязвимостей в программном обеспечении. Но, до сих пор не существует высокоэффективного и универсального решения данной задачи, основанного на применении машинного обучения.

Существуют исследования, которые применяют машинное обучение для обнаружения уязвимостей в программном обеспечении на основе исходных кодов [1]. Такие исследования имеют ряд ограничений, в том числе связанных с использованием синтетических датасетов, наиболее популярным из которых является набор уязвимых примеров Software Assurance Reference Dataset (SARD) [2]. Недостатком искусственно сгенерированных примеров является то, что их использование зачастую не позволяет обучить нейронные сети таким образом, чтобы обеспечивалась возможность выявления уязвимостей в исходном коде коммерческих продуктов. Также исследователи сталкиваются с проблемой предобработки кода программ для получения состоятельных результатов.

Исследование [3] показывает, что одним из важнейших этапов при решении поставленной задачи является именно предварительная обработка датасета и получение нужного представления программы для дальнейшей передачи в искусственную нейронную сеть. По результатам исследований авторы также указывают 15 уязвимостей в открытом программном обеспечении, которые были найдены с помощью модели на основе блока GRU.

В текущей работе предлагается осуществлять анализ бинарного кода скомпилированных программ и производить обучение нейронной сети на дизассемблированных листингах. Помимо прочего, решение данной задачи позволит искать уязвимости и ошибки без исходного кода, что позволит повысить безопасность более широкого спектра программных продуктов. Исследование [4] показывает, что данное направление является перспективным, однако, его результаты и оценка ограничены вышеупомянутым использованием датасета SARD.

При выполнении эксперимента в текущей работе были построены рекуррентные модели на основе блоков LSTM и GRU, а также их двунаправленные версии BLSTM и BGRU. Выбор моделей обосновывается тем, что задачи работы с последовательностью операций в программе отчасти схожи с обработкой естественных языков, так как каждое действие в программе имеет зависимость от предыдущих действий и влияет на будущие.

На текущем этапе работы также производилась работа с датасетом SARD и для проведения обучения были извлечены ассемблерные листинги уязвимых и безопасных функций. Таким образом, поиск ошибок в программном обеспечении производился на функциональном уровне без дополнительной обработки. Результаты эксперимента показали существенные недостатки предложенного метода функционального анализа – метрики обучения оказались низкими, точность составила 73%.

Таким образом, в сравнении с результатами других исследований можно сделать предположение, что для более качественного обучения требуется специальная предобработка датасета. В первую очередь, стоит обратить внимание на анализ графов программы – графа потока управления и графа потока данных. Это позволит удалить лишние инструкции и анализировать именно частные пути в программе. Для решения данной задачи применим статический анализ, также может помочь использование символьного исполнения для определения достижимых путей программы. Данный подход также применен и рассмотрен в [4].

Кроме того, для повышения качества обучения предлагается использование различных представлений бинарного кода. В частности, в данном исследовании используется PalmTree[5], являющаяся адаптацией языковой модели BERT[6] для представления ассемблерного кода. Также при проведении дальнейших исследований считается целесообразным произвести сравнение результатов при различных подходах к векторизации ассемблерного кода. Для повышения эффективности разрабатываемых моделей предполагается использование набора данных, сформированного на основе базы уязвимостей реальных программ с их разметкой, компиляцией, и последующими обработкой и анализом.

Список литературы:

1. Guanjun L., Wei X., Jun Z., Yang X., Deep Learning-Based Vulnerable Function Detection: A Benchmark // ICICS 2019: Information and Communications Security. – 2019. – p. 219-232.

2. Software Assurance Reference Dataset Project // National institute of Standards and Technology URL: https://samate.nist.gov/SARD/ (дата обращения: 26.05.2022).

3. Zhen L., Deqing Z., Shouhuai X., Hai J., Yawei Z., Zhaoxuan C., SySeVR: A Framework for Using Deep Learning to Detect Software Vulnerabilities // IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING. – 2018.

4. Junfeng T., Wenjing X., Zhen L., BVDetector: A program slice-based binary code vulnerability intelligent detection system // Information and Software Technology – 2020.

5. Xuezixiang L., Qu Y., Heng Y., PalmTree: Learning an Assembly Language Model for Instruction Embedding // CCS '21: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security. – 2021 – p. 3236–3251

6. Devlin J., Chang Ming-Wei., Lee K., Toutanova K., BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Transformers for Machine Learning – 2022 – p. 28-46.

Иванова О.Д., Калинин М.О.
Санкт-Петербургский политехнический университет Петра Великого

Беленко В.С., Черненко В.Г.
LG Electronics Inc.

РАЗРАБОТКА МЕТОДА ВЫЯВЛЕНИЯ АТАК УКЛОНЕНИЯ В СИСТЕМАХ МАШИННОГО ОБУЧЕНИЯ

***** *Исследование выполнено при финансовой поддержке LG Electronics Inc.*

Проблема выявления атак уклонения на системы машинного обучения является одной из актуальных проблем, возникающих в ходе использования вычислительного интеллекта при распознавании объектов, выявлении компьютерных атак, фильтрации информационных потоков, автоматизации принятия решения. При этом многие интеллектуальные системы подвержены атакам уклонения и их использование может привести к непредсказуемым результатам, так как атаки уклонения могут проводиться не только злоумышленниками, но и могут возникнуть случайно при работе искусственного интеллекта. Цель работы – анализ методов выявления атак уклонения, разработка универсального метода, который будет выявлять различные атаки уклонения.

Известные методы выявления атак уклонения делятся на два класса:

а) методы, использующие исходную систему машинного обучения – обрабатывают результаты, получаемые от системы, и на их основе принимают решения о том, являются ли данные вредоносными. Подразделяются на два класса: не преобразующие входные данные и преобразующие входные данные;

б) методы, изменяющие модель исходной системы машинного обучения – вносят изменения в систему и на основе результатов, получаемых от измененной системы, принимают решения о том, являются ли данные вредоносными. Подразделяются на две группы: добавляющие ветвь оценки достоверности и использующие методы регуляризации.

Методы выявления атак уклонения, которые реализованы в работе, делят входные данные на два класса: данные, относящиеся к атакам уклонения; и чистые данные, которые не относятся к атакам уклонения. В эксперименте использовалось три набора данных: MNIST, Fashion-MNIST, CIFAR10. Также были сгенерированы состязательные примеры, построенные при помощи методов на основе градиента, на основе карт значимости и граничного метода. Для генерации состязательных примеров использовалось библиотека foolbox.

Для экспериментального исследования эффективности методов выявления атак уклонения реализованы методы MSP и ODIN, использующие исходную систему машинного обучения. Также реализовано три метода, которые изменяют исходную систему. Метод Confidence добавляет ветвь оценки достоверности, при этом входные данные не

преобразуются. Метод Pre-confidence также добавляет аналогичную ветвь оценки достоверности, но при этом входные примеры преобразуются по формуле $x' = x - \varepsilon\, sign(-\nabla_x log f(x))$, где $f(x)$ – предсказание, ε – добавленное возмущение, x – исходный пример, x' - обработанный пример. Метод Uncertainty использует отсев и вычисляет неопределенности.

В работе использовались следующие разновидности атак уклонения: данные вне распределения (OOD), состязательные примеры, построенные при помощи методов на основе градиентов (SGM), на основе карт значимости (MS) и граничным методом (BA). Реализованные пять методов выявления атак уклонения протестированы с использованием простой нейросети со слоями Flatten, Dense (activation=relu), Dense (activation=relu), Dense. Модель обучалась 10 эпох. В таблице 1 приведен фрагмент сравнения эффективности методов выявления атак уклонения для набора данных MNIST, указаны параметры методов, если они применимы для данных методов и для которых получен наилучший результат.

Метод Uncertainty при определенных параметрах способен выявлять все типы атак уклонения. Однако при этом его точность является невысокой в сравнении с другими методами, и выбор границы, по которой можно отделить состязательные примеры и чистые данные, довольно сложен, в связи с чем и падает его точность. Альтернативные методы показывают хорошую точность в выявлении SGM, BA, MS атак уклонения, однако, они плохо справляются с OOD атаками. Также при использование этих методов в большинстве своем оценка для состязательных примеров имеет почти одинаковые значения, что значительно упрощает выбор границ, по которым можно отделить атаки уклонения от чистых данных. Однако эти значения при переобучении модели или ее обучении в ходе использования, могут существенно меняться, при этом границы будут также не сильно отличаться, что не позволяет выбрать границы, которые были бы более универсальны и не изменялись в ходе использования модели. Для устранения указанных недостатков на базе ODIN и Uncertainty разработан новый гибридный метод, который позволяет выявлять все атаки уклонения, включая OOD, и достигать более высокой точности в сравнении с методом Uncertainty.

Таблица 1 – Сравнение эффективности методов выявления атак уклонения (измерение метрик на наборе данных MNIST)

Атака	Метод	Точность	F1-мера	Параметры методов (если применимы)	
SGM	MSP	0,98	0,98		
	ODIN	0,99	0,99	T=100	ε =0,05
	Confidence	0,98	0,98		
	Pre-confidence	0,91	0,91	ε =0,05	
	Uncertainty	0,88	0,86	Доля отсева=0,05	Отклонение=0,25
MS	MSP	0,99	0,99		
	ODIN	0,99	0,99	T=100	ε =0,05
	Confidence	0,98	0,98		
	Pre-confidence	0,92	0,91	ε =0,05	
	Uncertainty	0,88	0,86	Доля отсева=0,05	Отклонение=0,25
OOD	MSP	0,49	0,66		
	ODIN	0,50	0,66	T=100	ε =0,05
	Confidence	0,51	0,66		
	Pre-confidence	0,47	0,61	ε =0,05	
	Uncertainty	0,88	0,86	Доля отсева=0,05	Отклонение=0,25
BA	MSP	0,99	0,99		
	ODIN	0,99	0,99	T=100	ε =0,05
	Confidence	0,98	0,98		

	Pre-confidence	0,91	0,90	ε =0,05	
	Uncertainty	0,88	0,86	Доля отсева=0,05	Отклонение=0,25

Измайлов И.В., Крундышев В.М.
Санкт-Петербургский Политехнический университет Петра Великого

ИДЕНТИФИКАЦИЯ ВНУТРЕННЕГО НАРУШИТЕЛЯ НА ОСНОВЕ АНАЛИЗА ЗАШИФРОВАННОГО ТРАФИКА С ИСПОЛЬЗОВАНИЕМ МЕТОДА ФИНГЕРПРИНТА

** Работа выполнена в рамках стипендии Президента РФ для поддержки молодых ученых и аспирантов (СП-2714.2021.5)*

В настоящее время наблюдается процесс цифровой трансформации в различных сферах деятельности человека. Цифровизация бизнес-процессов происходит как на отдельных предприятиях, так и в крупных общественных и социальных сферах: медицине, массовых коммуникациях, образовании и науке [1, 2]. Разрабатываемые цифровые решения позволяют не только поддерживать и расширять существующие методы, но и создавать инновационные продукты благодаря автоматизации процессов. В таких условиях необходимо обеспечить безопасность сетевой инфраструктуры предприятия как от внешних, так и от внутренних угроз. Решениям, нацеленным на повышение информационной безопасности от внешнего нарушителя, уделяется достаточно много внимания [3, 4]. Но в то же время для многих компаний задача защиты от внутреннего нарушителя порой является не столь приоритетной, несмотря на то, что именно внутренний нарушитель обладает наибольшими возможностями по реализации угроз информационной безопасности [5].

В данной работе рассматривается сценарий компьютерной атаки на веб-ресурсы предприятия, реализуемой внутренним нарушителем. Злоумышленник, пользуясь анонимной сетью (например, Tor), зашифровывает свой сетевой трафик с целью скрыть попытку несанкционированного доступа к веб-ресурсу предприятия. Таким образом, с помощью просмотра заголовков IP-пакетов становится невозможным идентифицировать злоумышленника и определить веб-ресурсы, к которым совершено обращение.

Для решения данной проблемы предлагается использовать фингерпринт трафика. Данный метод применяется для перехвата зашифрованного веб-трафика и его последующего анализа. С помощью утилиты capinfos из дамп-файлов извлекаются следующие признаки: количество пакетов, объем данных (байт), средняя скорость передачи данных(бит/сек), средний размер пакетов (байт) и средняя скорость передачи пакетов (пакетов/сек). При сопоставлении отпечатка веб-трафика с отпечатками из имеющейся базы данных в данной работе используется аппарат нейронных сетей. С использованием библиотек TensorFlow и Keras на языке Python построен классификатор на основе многослойной нейронной сети прямого распространения. Модель имеет входной слой с 5 нейронами, один скрытый слой, который содержит 8 нейронов и функцию активации relu, а также 4 нейрона на выходном слое с функцией активации softmax. Результаты экспериментальных исследований показали, что максимальная точность классификации составляет 97% и достигается при 200 эпохах обучения и обучающем наборе данных равном 1000.

График зависимости точности классификации от конфигурации нейронной сети представлен на рисунке 1.

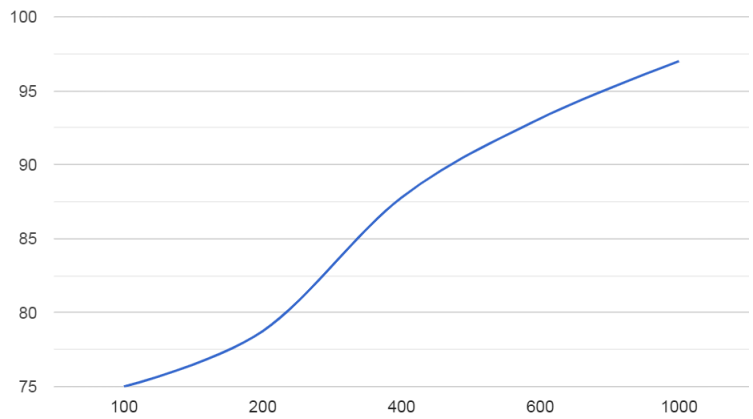

Рис. 1. График зависимости точности классификации от размера обучающей выборки

Список литературы:

1. Полтавцева М. А. Защита данных в системах мониторинга безопасности крупномасштабных объектов / М. А. Полтавцева, М. О. Калинин // Управление развитием крупномасштабных систем (MLSD'2019): Материалы двенадцатой международной конференции, Москва, 01–03 октября 2019 года / Под общей редакцией С.Н. Васильева, А.Д. Цвиркуна. – Москва: Международный научно-исследовательский институт проблем управления РАН, 2019. – С. 1019-1029.

2. Зегжда Д. П. Универсальный метод обнаружения кибератак на глобальные информационные системы поддержки цифровой экономики / Д. П. Зегжда, П. Д. Зегжда, М. О. Калинин // Методы и технические средства обеспечения безопасности информации. – 2019. – № 28. – С. 48-49.

3. Александрова Е. Б. Анализ подходов к обеспечению защищенного взаимодействия в крупномасштабных промышленных системах / Е. Б. Александрова, А. В. Ярмак, М. О. Калинин // Проблемы информационной безопасности. Компьютерные системы. – 2018. – № 4. – С. 140-144.

4. Калинин М. О. Анализ информационной безопасности предприятия на основе мониторинга информационных ресурсов с использованием машинного обучения / М. О. Калинин, С. И. Штеренберг // Интеллектуальные технологии на транспорте. – 2018. – № 3(15). – С. 47-54.

5. Овасапян Т. Д. Применение нейронных сетей для выявления внутренних нарушителей в VANET-сетях / Т. Д. Овасапян, Д. А. Москвин, М. О. Калинин // Проблемы информационной безопасности. Компьютерные системы. – 2018. – № 1. – С. 68-73.

Крашенинников Э.А., Ярмак А.В., Александрова Е.Б.
Санкт-Петербургский политехнический университет Петра Великого

КОНТРОЛЬ ДОСТУПА К ДАННЫМ ОБЛАЧНОГО ХРАНИЛИЩА НА ОСНОВЕ ИЗОГЕНИЙ

Благодаря активному развитию облачных вычислений пользователи и организации всё чаще используют сторонние сервисы для хранения и обмена данными. Облачные провайдеры (такие как Amazon, Microsoft, Apple и т. д.) предоставляют различные пакеты услуг в зависимости от требований клиентов. Однако, с точки зрения безопасности, использование

облачных сервисов может привести к возникновению дополнительных векторов атак, в том числе, со стороны недоверенного провайдера облачных сервисов.

Одним из решений данной проблемы является использование схем криптографического контроля доступа. Данные схемы позволяют обеспечить безопасность в различных окружениях, не требуя значительных изменений фундаментальной архитектуры [1]. Основная идея такого подхода – использование для контроля доступа к данным криптографических алгоритмов и протоколов. Данные, находящиеся в хранилище, зашифрованы, и доступ к ним может получить только пользователь, имеющий необходимые криптографические ключи.

В рамках данной работы предлагается модификация системы Crypt-DAC [2]. В качестве основного математического аппарата используются изогении эллиптических кривых, представляющие собой один из перспективных механизмов для построения схем постквантовой криптографии. В частности, в качестве основных криптографических примитивов были использованы шифрование на основе SIDH [3,4] и схема неоспоримой цифровой подписи [5].

На рисунке 1 проиллюстрирован процесс получения доступа к данным на чтение и запись. В рамках схемы вводятся следующие сущности и объекты: пользователи $U = \{u_1, u_2, \dots, u_n\}$, роли $R = \{r_1, r_2, \dots, r_m\}$, файлы $F = \{f_1, f_2, \dots, f_l\}$. Таблицы RK и FK содержат зашифрованные ключи ролей и ключи файлов. Если роли r_i предоставлен набор разрешений op к файлу f_j, то существует кортеж $FK_{r_i f_j} = \left(r_i, Name(f_j), op, Enc_{pk_{r_i}}\left(k_{f_j}\right)\right)$. Если пользователю u_i была выдана роль r_j, то существует кортеж $RK_{u_i r_j} = \left(u_i, r_j, Enc_{pk_{u_i}}\left(sk_{r_j}\right)\right)$. При получении доступа на чтение пользователь должен последовательно расшифровать данные таблиц и только затем, вычислив ключ шифрования файла, расшифровать необходимый файл. При получении доступа на запись пользователь должен загрузить обновленный зашифрованный файл на сервер и подтвердить внесенные изменения с помощью подписи. Для ускорения операции аутентификации участников роли администратором, вместо схемы подписи можно использовать j-инвариант кривой, полученной путем построения изогении на закрытом показателе роли и значении хэш-функции измененного файла.

В ходе взаимодействия с облачным хранилищем администратору может потребоваться отозвать разрешение роли у какого-либо пользователя или отозвать членство данного пользователя. Удаление пользователя u_i из роли r_j требует следующих действий: генерации новых ключей роли; обновления кортежей RK, связанных с ролью; генерации новых ключей шифрования файлов и обновления кортежей FK, связанных с этими файлами.

Генерация новых ключей роли и обновление кортежей RK предотвращает доступ пользователя u_i к новому ключу роли r_j, а генерация новых ключей шифрования файлов и обновление кортежей FK не позволяет пользователю получить доступ к файлам с помощью кэшированных предыдущих файловых ключей. Лишение прав доступа роли к файлу требует меньшего количества операций, чем генерация новых ключей роли, однако также требует расшифрования и повторного зашифрования файлов.

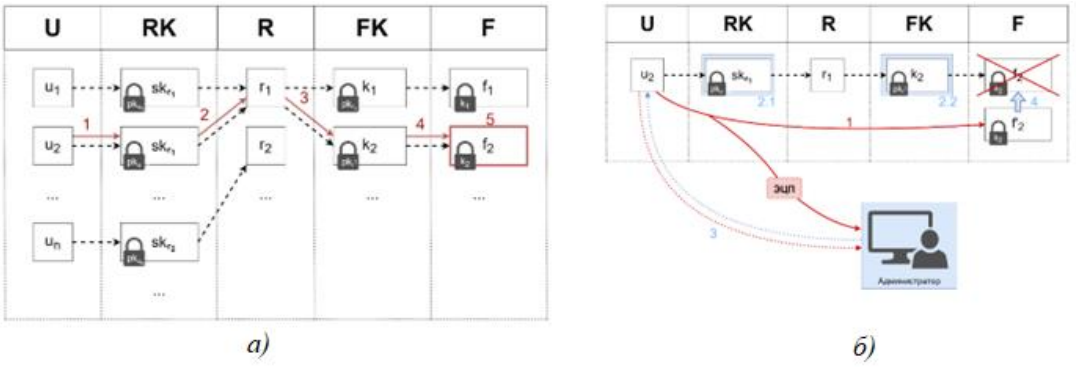

Рис. 1. Получение доступа а) на чтение; б) на запись

Предложенная модель обладает следующим недостатком: если в системе находится некоторая роль, которая имеет доступ к большому количеству файлов, то при попытке отозвать доступ администратору необходимо выгрузить все файлы, доступ к которым имеет данная группа, затем расшифровать и повторно зашифровать каждый файл, а после повторно загрузить их в облачное хранилище. Решением данной проблемы является делегирование облачному провайдеру операции повторного шифрования. Вместо того, чтобы заменять файловый ключ, система будет добавлять дополнительные слои шифрования, используя мощности облачных серверов. Это позволяет увеличить быстродействие выполняемых операций, однако способствует увеличению множества используемых ключей.

Список литературы:

1. Kayem A. V. D. M., Akl S. G., Martin P. Adaptive cryptographic access control. – Springer Science & Business Media, 2010. – Т. 48.

2. Qi S., Zheng Y. Crypt-DAC: cryptographically enforced dynamic access control in the Cloud //IEEE Transactions on Dependable and Secure Computing. – 2019. – Т. 18. – №. 2. – С. 765-779.

3. Jao D., Feo L. D. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies //International Workshop on Post-Quantum Cryptography. – Springer, Berlin, Heidelberg, 2011. – С. 19-34.

4. Costello C. Supersingular isogeny key exchange for beginners //International Conference on Selected Areas in Cryptography. – Springer, Cham, 2019. – С. 21-50.

5. Jao D., Soukharev V. Isogeny-based quantum-resistant undeniable signatures //International Workshop on Post-Quantum Cryptography. – Springer, Cham, 2014. – С. 160-179.

Лазарев К.С., Платонов В.В.
Санкт-Петербургский политехнический университет Петра Великого

МЕТОД МИНИМИЗАЦИИ БУЛЕВЫХ ФУНКЦИЙ БЕЗ ИСПОЛЬЗОВАНИЯ ПОЛНОГО ПЕРЕБОРА

Булевы функции широко применяются при проектировании управляющих и вычислительных устройств. Особую роль играют булевы функции в криптографии, а в криптоанализе актуальна задача минимизации булевых функций [1, 2]. При этом основным способом минимизации является применение метода Квайна, который использует полный перебор [3]. Целью исследования является разработка метода минимизации булевых функций, который лишен данного недостатка.

Минимизация булевой функции состоит в обработке таблицы значений функции и в выявлении переменных, которые не влияют на результат. Возьмем для примера функцию трех переменных F (x, y, z), значение которой записаны в табл. 1.

Табл. 1. Таблица истинности функции F (x, y, z)

	X = 0				X = 1			
x	0	0	0	0	1	1	1	1
y	0	0	1	1	0	0	1	1
z	0	1	0	1	0	1	0	1
F(x,y,z)	1	1	0	0	1	1	0	0

Значения данной функции при X = 0 и X = 1 идентичны – левая и правая части последней строки равны (1100 = 1100). Это означает, что переменная X не влияет на результат функции, а значит можно рассматривать лишь половину всей таблицы истинности. Результат данных действий представлен в табл. 2.

Табл. 2. Таблица истинности функции F (y, z)

	Y = 0		Y = 1	
y	0	0	1	1
z	0	1	0	1
F(y,z)	1	1	0	0

Вновь разделим строку со значениями функции на две равные части, чтобы определить, является ли переменная Y существенной. Значение первой и второй ячейки не равно значению третьей и четвертой (11 ≠ 00). Из этого следует вывод, что переменная Y является существенной, а значит не может быть убрана из функции.

В случае, если минимизация по переменной не удалась, необходимо разделить таблицу истинности пополам и рассматривать их как две отдельные таблицы. При этом, минимизацию можно производить только в том случае, если во всех таблицах истинности это представляется возможным. Поэтому, если хотя бы в одной таблице истинности переменная является существенной, минимизация невозможна.

Этот же алгоритм применяется и в случае, когда рассматривается более одной таблицы истинности. Все имеющиеся таблицы разделяются на две, образуя новое множество. Таким образом, имея k различных таблиц истинности, каждая существенная переменная будет увеличивать значение k в 2 раза. В следствие этого, по окончании работы алгоритма будет верно равенство $k = 2^m$, где m – количество существенных переменных.

Результат разделения таблицы представлен в табл. 3.

Табл. 3. Разделенная таблица истинности функции F (y, z)

	Z = 0	Z = 1
y	0	0
z	0	1
F(y,z)	1	1

	Z = 0	Z = 1
y	1	1
z	0	1
F(y,z)	0	0

Теперь проверяется влияние переменной Z на значение функции. Видно, что в левой и правой частях таблицы 3, значения функции при различных Z одинаковы (1 = 1 и 0 = 0 соответственно). Это означает, что можно минимизировать функцию по переменной Z, поэтому убираем вторую строку и всю правую часть обеих таблиц. Конечный результат представлен в табл. 4.

Табл. 4. Результат минимизации функции F (x, y, z)

y	0	1
F(y)	1	0

Итоговая таблица составлена из k = 2 промежуточных таблиц истинности. Кроме того, в ходе работы алгоритма выявлено m = 1 существенных переменных. Результат удовлетворяет равенству $k = 2^m$ ($2 = 2^1$), а значит алгоритм выполнен верно.

Предложенный подход позволяет минимизировать булевые функции, не используя полный перебор. Данный метод рассматривает каждую переменную по отдельности, при этом обрабатывая несколько таблиц истинности. Так как количество таблиц в худшем случае может быть равно n/2, где n – количество переменных, а количество рассматриваемых переменных – $\log_2(n)$, то метод имеет линейно-логарифмическую сложность $n \cdot \log_2(n)$.

Целью дальнейшего исследования является уменьшение сложности алгоритма, что потребует выбора необходимой структуры данных для представления таблиц истинности в памяти компьютера.

Список литературы:

1. Ростовцев А.Г., Маховенко Е.Б. Введение в теорию итерированных шифров. – СПб.: НПО «Мир и семья», 2003. – С. 66-69.

2. Горбатов В.А. Фундаментальные основы дискретной математики. Информационная математика. // М.: Наука. Физматлит. – 2000. – С. 102-107.

3. Ахметова Н.А., Усманова З.М. Дискретная математика. Функции алгебры логики. Учебное пособие // Уфимский государственный авиационный технический университет. – 2000. – С. 74-76.

Макаров М. В., Штыркина А. А.
Санкт-Петербургский политехнический университет Петра Великого

МЕТОД ОБНАРУЖЕНИЯ ОТРАВЛЕНИЯ СВЕРТОЧНЫХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Существует множество различных архитектур искусственных нейронных сетей (ИНС), каждая из которых может быть предназначена для решения задач классификации и регрессии для определенных типов данных: изображения, видео, временные ряды и т.д. В данной работе рассматриваются сверточные ИНС ввиду их широкого применения в инфраструктурах, где ошибка в прогнозе может привести к катастрофическим последствиям: медицина, автопилотирование транспортных средств, системы распознавания лиц, системы информационной безопасности и т.д.

Ошибка прогноза модели может быть следствием некорректного обучения модели, либо результатов намеренного внесения в модель вредоносной логики на этапе обучения путем «отравления» обучающей выборки [1]. Обнаружить какие-либо изменения в обучающей

выборке невозможно без использования специальных алгоритмов анализа. Помимо этого, довольно часто к обучающей выборке доступ не предоставляется, поэтому провести статический анализ часто не представляется возможным.

На данный момент был найден всего один метод, чья точность превышает 80%, который позволяет динамически обнаруживать вредоносную логику в ИНС [2]. Однако такой метод требует доступа к аналогичной обучающей выборке для обучения нового классификатора, что вызывает проблему: риски того, что новый классификатор также будет отравлен все равно остаются.

Исходя из вышеизложенного, целью данной работы является динамическое обнаружение вредоносной логики в ИНС на основе алгоритмического анализа вывода ИНС.

Для достижения поставленной цели необходимо решить следующие задачи:

1) проанализировать возможные атаки типа отравление на сверточных ИНС;
2) проанализировать способы поиска отравленных данных в обучающей выборке на основе вывода классификатора;
3) разработать подход к анализу вывода ИНС на основе вейвлет-преобразований.

Способов совершить отравление классификатора существует довольно много [1], однако, в данной работе рассмотрен подход, в результате которого на некоторые элементы обучающей выборки наносятся определенные признаки, выбранные атакующим [2]. Именно такой способ отравления позволяет в будущем подать на вход ИНС произвольные данные с желаемым признаком, и они будут классифицированы именно так, как нужно атакующему.

Как показал анализ работ [2, 3], для защиты ИНС не существует четкого алгоритма, который бы проверял классификатор на предмет того, что модель учитывает признаки, добавленные атакующим. Ни одна из статей не предлагает алгоритм, который не оперировал бы обучающей выборкой, а работал только с тестовым набором данных.

Для разработки метода обнаружения вредоносной логики предлагается использовать вейвлет-преобразования – специальные математические преобразования, которые позволяют создать аппроксимирующую функцию из n-мерного потока данных. Это преобразование позволяет получить большое количество избыточных коэффициентов и произвести эффективное сжатие медиафайлов.

Алгоритм работы предлагаемого метода включает в себя следующие шаги:

1. Для каждого объекта из тестового набора данных G выполняется сжатие с помощью вейвлет-преобразования. Результаты сжатия заносят набор G';
2. Объекты из множеств G и G_i' подаются на вход проверяемой модели сверточной ИНС. В результате работы модели получаются выводы Out и Out' для несжатых и сжатых данных, соответственно;
3. Выполняется покомпонентное сравнение векторов Out и Out': если разница между компонентами векторов являются аномальными относительно других компонент, то анализируемый объект добавляется с список кандидатов на проверку на наличие зловредного признака.

Определение того, является ли разница между компонентами векторов аномальной, может сводиться как к выявлению константного порогового значения, так и к поиску разделяющей функции. Решение данного вопроса относится к направлениям дальнейших исследований.

Список литературы:

1. Practical Poisoning Attacks on Neural Networks [электронный ресурс]. – Режим доступа https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123720137.pdf свободный (28.05.2022)

2. Detecting Backdoors in Neural Networks Using Novel Feature-Based Anomaly Detection [электронный ресурс]. – Режим доступа https://arxiv.org/pdf/2011.02526.pdf свободный (28.05.2022).

3. Backdoor Embedding in Convolutional Neural Network Models via Invisible Perturbation [Электронный ресурс]. – Режим доступа https://arxiv.org/pdf/1808.10307.pdf свободный (28.05.2022)

Обидина А.И., Платонов В.В.
Санкт-Петербургский политехнический университет Петра Великого

СОКРАЩЕНИЕ РАЗМЕРНОСТИ БАЗЫ ДАННЫХ СЕТЕВЫХ АТАК UNSW-NB 15 С ПОМОЩЬЮ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

Успешная работа компьютерной системы напрямую связана с работой системы обнаружения вторжений (СОВ), которая в результате анализа различных данных может определить атаку или вторжение [1]. В настоящее время СОВ, основанные на методах машинного обучения, достаточно точны, именно их работа будет рассмотрена в рамках данной статьи. Эффективность таких СОВ зависит от данных, подаваемых им на вход, конкретнее, от размерности данных. При большом количестве параметров, их зависимости друг от друга, производительность может быть значительно снижена. Чтобы не допустить этого, для сокращения размерности больших входных данных может быть использован метод главных компонент, который позволяет уменьшить размерность с сохранением в оставшихся параметрах максимального количества информации.

Сокращение размерности базы данных UNSW-NB 15, исследование метода главных компонент (МГК) было рассмотрено во многих статьях [2, 3]. Реализация МГК может варьироваться, но результаты приблизительно одинаковы – при выборе оптимальных главных компонент их доля от суммарной дисперсии близка к 90 %.

Цель исследования – анализ применимости метода главных компонент для выявления сетевых атак. Метод главных компонент – важный и мощный метод для современного анализа данных [4]. В основе МГК лежат сложные базовые математические принципы для преобразования ряда данных, которые могут быть коррелированы, в меньшее число переменных, называемых главными компонентами. В результате работы МГК получаются главные компоненты (их количество равно количеству исходных параметров) – линейные комбинации исходных, причем главные компоненты не коррелируют между собой, и большая часть информации об исходных данных помещается в первые компоненты (компоненты вычисляют согласно порядку убывания их доли от суммарной дисперсии исходных величин).

Для сокращения размерности выбрана база данных сетевых атак UNSW-NB 15, содержащая 42 параметра обучения. Обученная модель решает задачу классификации - на тестовых параметрах должна определить, имеет тут место атака (1) или норма (0). Для исследования МГК были выбраны четыре метода машинного обучения: линейная регрессия (C = 1) [5], дерево решений [6], метод k ближайших соседей (k = 10) [7], перцептрон с одним скрытым слоем (количество нейронов в нем вычислялось как 2*n – 1) [8, 9]. Модели обучались на всех 42 параметрах, после чего для каждой модели вычислялись метрики precision

(чувствительность), recall (полнота), F1-score (F1-мера), accuracy (точность), ROC-AUC [10] (таблица 1).

Среди полученных главных компонент выбраны 18 из них (их доля от суммарной дисперсии исходных параметров – 90%) (рис. 1). Далее модели обучались с помощью них (таблица 2). Результаты для всех моделей, кроме дерева решений улучшились. Результаты дерева решений незначительно уменьшились.

Таблица 1. Значения метрик моделей после обучения на исходной базе данных

Модель / Метрика, %	Линейная регрессия	Дерево решений	Метод k ближайших соседей	Перцептрон с одним скрытым слоем (83 нейрона)
Precision	95,62	94,31	89,21	98,79
Recall	78,67	94,67	86,98	77,19
F1-score	86,32	94,49	88,08	86,67
Accuracy	80,63	92,97	84,57	80,58
ROC-AUC	83,07	92,33	83,48	86,48

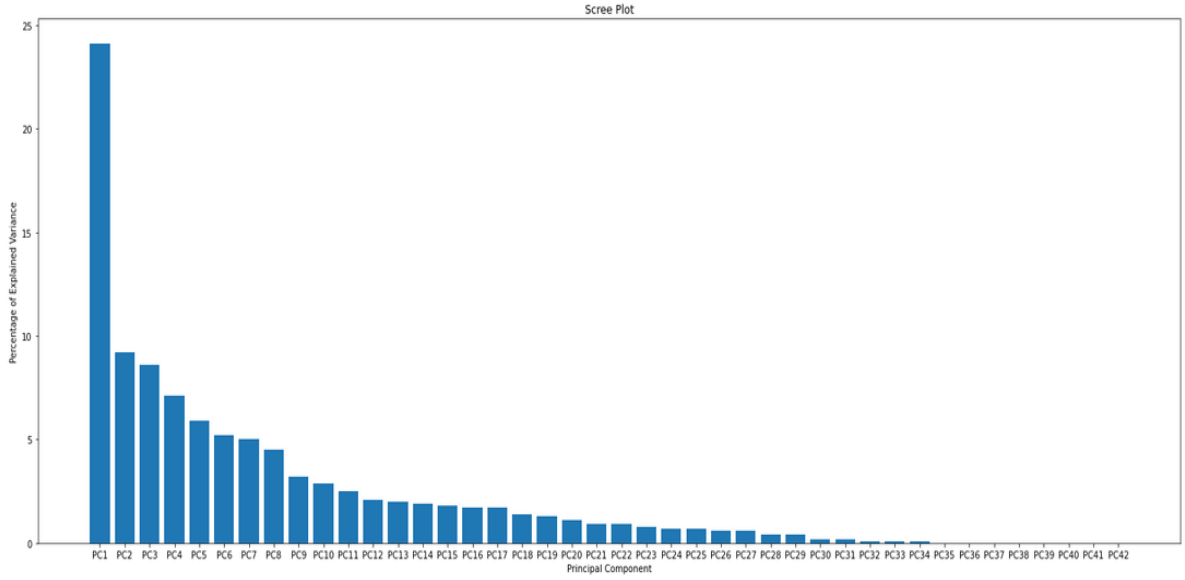

Рис. 1. Значимость главных компонент (%)

Таблица 2. Значения метрик моделей после обучения с помощью главных компонент

Модель / Метрика, %	Линейная регрессия	Дерево решений	Метод k ближайших соседей	Перцептрон с одним скрытым слоем (35 нейронов)
Precision	94,92	92,34	91,35	95,99
Recall	85,20	92,51	94,17	91,31
F1-score	89,80	92,42	92,74	93,59
Accuracy	86,22	90,33	90,86	91,60
ROC-AUC	86,97	89,50	89,82	91,76

Список литературы:

1. Платонов В.В. Программно-аппаратные средства защиты информации // Издательский центр «Академия». – 2013. – С. 177.
2. Sheluhin O.I., Ivannikova V.P. Comparative analysis of informative features quantity and composition selection methods for the computer attacks classification using the UNSW-NB 15 dataset. – 2020. – P. 53-60.
3. Kasongo S.M., Yanxia Sun Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset. – 2020. – P. 8-11, 14.
4. Померанцев А. Метод Главных Компонент (PCA). – 2008.
5. Shichao Zbang, Xuelong Li, Ming Zong Learning k for kNN Classification.– 2017.– P.1-5.
6. Decision tree methods: applications for classification and prediction // Shanghai Archives of Psychiatry. – 2015. – P. 130-134.
7. Jurgen Grob Linear Regression // Springer. – 2003. – P.33-46.
8. Осовский С. Нейронные сети для обработки информации // Горячая линия – Телеком. – 2017. 427 с.
9. Laveen N. Kanal Perceptron // Encyclopedia of Computer Science. – 2003. – P. 1383-1385.
10. Jerome Fan, Saneel Upadhye, Andrew Worster Understanding receiver operating characteristic (ROC) curves // Cambridge University Press. – 2015. – P. 19-20.

Писков А.А., Жуковский Е.В.
Санкт-Петербургский политехнический университет Петра Великого

ИСПОЛЬЗОВАНИЕ УЗЛОВ-ПРИМАНОК ДЛЯ ОБНАРУЖЕНИЯ АТАК НА КОРПОРАТИВНЫЕ ВЕБ-РЕСУРСЫ

В структуру современных компаний зачастую внедряются веб-приложения, которые могут использоваться в различных целях: управление внутренней инфраструктурой, отдельными узлами предприятия, организация видеоконференцсвязи и т. п. Корпоративные веб-ресурсы часто становятся целью для злоумышленников, т. к. многие приложения доступны из внешней сети, следовательно, подвержены сканированию и могут быть использованы для проникновения в сеть организации. Эксплуатация веб-уязвимостей может быть использована как один из этапов целенаправленной атаки (APT). На рис. 1 показаны изменения в OWASP Top10 (10 наиболее распространенных уязвимостей в веб-приложениях на основе анализа инцидентов и тестирования). Веб-приложения подвергаются атакам с использованием новых типов уязвимостей. Для выявления новых типов атак необходимо средство анализа поведения злоумышленников во время проведения атаки.

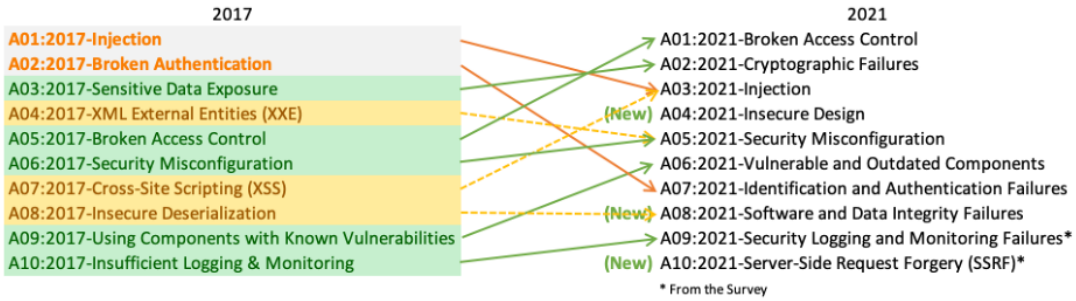

Рис.1 Изменения в списке наиболее распространенных веб-уязвимостей OWASP

Современные решения для выявления атак на веб-ресурсы основаны на анализе трафика и сравнении пакетов с сигнатурами [1-4]. Данный подход не позволяет выявлять новые типы атак и требует своевременного обновления баз данных угроз. Известны решения на базе методов машинного обучения для комплексного анализа поведения системы при проведении атаки [5-6]. Их недостаток – необходимость синтеза качественных данных о нормальном поведении пользователя для обучения модели.

Актуальным вектором развития в области выявления атак на веб-ресурсы является использование узлов-приманок (honeypot-систем). Узлы-приманки, размещаемые в организации, помогают отвлечь злоумышленника от реальных веб-приложений, собрать данные об атаке, вовремя оповестить администраторов безопасности о вторжении. Изученные средства [7-12] для эмуляции поведения уязвимого веб-приложения на основе honeypot-систем имеют ряд ограничений: привязка к языку программирования, на котором реализовано веб-приложение, требование наличия исходного кода для его преобразования, шаблонные пользовательские интерфейсы.

В ходе исследования разработан подход к построению систем узлов-приманок, эмулирующих поведение уязвимых веб-ресурсов, имеющих максимальное сходство с интерфейсом целевого ресурса и не требующих исходного кода веб-приложения. Предлагаемый подход к построению honeypot-системы с низким уровнем взаимодействия включает в себя несколько этапов. На первом этапе происходит копирование веб-интерфейса целевого веб-ресурса для его дальнейшего отображения сервером приманкой. Данный этап необходим для обеспечения правдоподобия эмулируемого приложения. Это может стать дополнительным отвлекающим фактором для злоумышленника при проведении ручной разведки веб-ресурса. На втором этапе модифицируется полученная веб-страница, которая, как правило, представляет собой HTML-разметку: в разметку встраиваются уязвимые поля. Полученная веб-страница отображается веб-сервером узла-приманки.

Помимо модуля копирования страницы система для разворачивания узла-приманки содержит модуль, анализирующий поступающие на поддельный веб-сервер запросы и выявляющий тип проводимой атаки; модуль встраивания в ответный запрос поддельных данных; модуль логирования событий; модуль управления узлом-приманкой. При встраивании данных в запрос необходимо учитывать сигнатуры уязвимостей, которые содержатся в общедоступных базах данных популярных сканерах уязвимостей (OpenVAS, Acunetix), что также повысит вероятность отвлечения злоумышленника от эмулируемого веб-ресурса.

Список литературы:

1. Hoang Dau, Hưng Nguyễn. A survey of tools and techniques for web attack detection // Special Issue CS. – 2022.

2. Kaur Jaspreet, Shakil Samiya, Shakil Sadiya. A brief survey on sandboxing techniques and it's vulnerabilities. // National Conference on Recent Advances in Information and Communication Technologies. – 2017.

3. Sun, F., Xu, L., Su, Z. Static detection of access control vulnerabilities in web applications. – 2017.

4. Le M., Stavrou A., Kang B.B. DoubleGuard: Detecting intrusions in multitier web applications. // IEEE Trans. Dependable Secur. Comput – 2012.

5. Rokia Lamrani Alaoui, El Habib Nfaoui. Deep learning for vulnerability and attack detection on web applications: A systematic literature review. – 2022.

6. Yao Pan, Fangzhou Sun, Zhongwei Teng, Jules White, Douglas C. Schmidt, Jacob Staple, Lee Krause. Detecting web attacks with end-to-end deep learning. // Journal of Internet Services and Applications. – 2019.

7. Mphago, B., Bagwasi, O., Phofuetsile, B., Hlomani, H. Deception in dynamic web application honeypots: Case of glastopf. // Int. Conf. on Security and Management. – 2015.

8. Rahmatullah, D. K., Nasution, S. M., Azmi, F. Implementation of low interaction web server honeypot using cubieboard. – 2016.

9. Muter, M., Freiling, F., Holz, T., Matthews, J. A generic toolkit for converting web applications into high-interaction honeypots. – 2008.

10. Anujot Boparai, Ron Ruhl, Dale Lindskog. The behavioural study of low interaction honeypots: dshield and glastopf in various web attack. – 2020.

11. Amirreza Niakanlahiji, Jafar Haadi Jafarian, Bei-Tseng Chu, Ehab Al-Shaer. HoneyBug: Personalized Cyber Deception for Web Applications. // 53rd Hawaii Int. Conf.on System Sciences. – 2020.

12. Djanali S., Arunanto F., Pratomo B.A., Studiawan H., Nugraha S.G. SQL injection detection and prevention system with Raspberry Pi honeypot cluster for trapping attacker. // Int. Symposium on Technology Management and Emerging Technologies. – 2014.

Саломатин М. В., Штыркина А. А.
Санкт-Петербургский политехнический университет Петра Великого

ИССЛЕДОВАНИЕ МЕХАНИЗМОВ ЗАЩИТЫ СВЁРТОЧНЫХ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ, ИСПОЛЬЗУЕМЫХ ДЛЯ ОБНАРУЖЕНИЯ ВРЕДОНОСНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ, ОТ СОСТЯЗАТЕЛЬНЫХ АТАК

Несмотря на все достижения в сфере искусственного интеллекта (ИИ), не так давно выяснилось [1], что глубокие искусственные нейронные сети (ИНС) — алгоритмы, входящие в состав большинства систем ИИ, — подвержены вредоносным атакам, использующим, на первый взгляд, легитимные входные данные. ИНС можно обмануть, внеся во входные данные незначительные изменения, которые будут незаметны для человека. Такие атаки называют состязательными. Ввиду того, что многие системы ИИ получают свои входные данные из внешних источников, подверженность вредоносным входным данным открывает новую угрозу безопасности. В рамках данной работы исследуется возможность защиты сверточной ИНС от нелегитимных входных данных с помощью метода состязательного обучения.

В качестве исследуемых ИНС были выбраны: модель с архитектурой MalConv [2] и ее улучшенная версия – Enhanced MalConv [3]. Приведенные модели решают задачу бинарной классификации файла на вредоносный и не вредоносный.

Для того чтобы исследовать возможность защиты ИНС, были рассмотрены механизмы генерации состязательных примеров: метод «белого» ящика, который требует знания параметров модели, и метод «черного» ящика. К методам «белого» ящика относятся: метод быстрого градиентного знака (FGSM) [4], метод добавления доброкачественных признаков (BFA) и его улучшенная версия (Enchanced BFA) [4]. В качестве метода «черного» ящика была выбрана атака на основе случайных вставок и ее улучшенная версия [4]. Сравнительные результаты успешности создания состязательных атак представлены в Таблице 1.

Далее были рассмотрены методы защиты от состязательных атак на сверточные ИНС. Среди всех рассмотренных методов был выбран метод на основе состязательного обучения [4], который заключается в добавлении в обучающую выборку правильно размеченных состязательных примеров с последующим дообучением модели. При применении данного метода защиты к полученной модель было замечено, что наличие пяти состязательных примеров в обучающей выборке снижает успех атаки с использованием состязательных примеров до 20 %. Сравнительная статистика успеха атаки на модель в зависимости от числа состязательных примеров в обучающей выборке представлена в таблице 2.

Данный эксперимент показал, что защищенность ИНС зависит от обучающих данных, соответственно, дообучение с достаточным расширением обучающего набора данных может значительно увеличить защищенность сети.

Дальнейшее направление исследований состоит в выявлении наличия закономерностей в подверженности состязательным атакам (и защищенности от них) с точки зрения архитектуры для модели Enhanced MalConv.

Таблица 1. Успешность атак в зависимости от количества добавляемых байт.

Название атаки	Число байтов в состязательном примере	Успешность создания состязательного примера, %
FGSM	1000	1
	2000	2
	5000	3
	10000	3
	20000	4
BFA	1000	6
	2000	15
	5000	31
	10000	67
	20000	93
Улучшенный метод случайных вставок	1000	10
	2000	21
	5000	34
	10000	57
	20000	65

Таблица 2. Успешность атаки после применения метода состязательного обучения

Число состязательных примеров в обучающей выборке	Средняя успешность атак с использованием состязательных примеров, %
1	92
2	84
5	34
10	22
50	12
100	3

Список литературы:

1. Goodfellow I. J., Shlens J., Szegedy C. Explaining and harnessing adversarial examples //arXiv preprint arXiv:1412.6572. – 2014.

2. Raff E. et al. Malware detection by eating a whole exe //Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence. – 2018.

3. E. Raff, W. Fleshman, R. Zak, H.S. Anderson, B. Filar, M. McLean. Classifying Sequences of Extreme Length with Constant Memory Applied to Malware Detection //arXiv:2012.09390v.

4. Suciu O., Coull S. E., Johns J. Exploring adversarial examples in malware detection //2019 IEEE Security and Privacy Workshops (SPW). – IEEE, 2019. – С. 8-14.

Пагуба Г. Ю., Павленко Е. Ю.
Санкт-Петербургский политехнический университет Петра Великого

ГЕНЕРАЦИЯ ВХОДНЫХ ДАННЫХ НА ОСНОВЕ ЦЕПЕЙ МАРКОВА ДЛЯ ФАЗЗИНГА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Поиск ошибок в программном обеспечении становится сложнее с увеличением сложности программного обеспечения, а также сложности механизмов его разработки (компиляторы, IDE и т. д.). Также сложность увеличивается за счет разнообразия входных данных, обрабатываемых программным обеспечением.

Эффективным способом поиска ошибок в программном обеспечении считается фаззинг. Фаззинг – процесс тестирования программного обеспечения, при котором программное обеспечение выполняется с частично или полностью некорректными входными данными [1].

Инструменты, осуществляющие фаззинг программного обеспечения, называются фаззерами и разделяются на следующие виды по принципу генерации входных данных тестируемого программного обеспечения [1]:

1. Мутирующие (Mutation based) – фаззеры, получающие новые наборы входных данных для тестируемого программного обеспечения посредством применения мутаций к существующим наборам.
2. Порождающие (Generation based) – фаззеры, получающие новые наборы входных данных посредством создания на основе специально созданной модели, описывающей свойства корректных входных данных.

Ключевой особенностью порождающих фаззеров является наличие модели, описывающей входные данные тестируемой программы для функционирования. Такая модель может быть получена посредством экспертного анализа тестируемого программного обеспечения или, в случае сетевых протоколов, посредством формального задания спецификации этих протоколов.

Однако, в большинстве случаев, экспертный анализ программного обеспечения – задача очень трудоёмкая. Также, тестируемое программное обеспечение может не иметь исходных кодов и иметь защиту от дизассемблирования. Эти причины существенно усложняют задачу составления модели входных данных тестируемой программы на основе экспертного анализа.

Предложенный в данном исследовании метод генерации входных данных для фаззинга программного обеспечения подразумевает разбиение входных данных тестируемой программы на блоки, где блок – последовательность из одного или более байт и представление модели входных данных тестируемой программы для порождающего фаззера как цепь Маркова, где событие – наличие произвольного блока входных данных тестируемой программы за предыдущим. Цепь Маркова — последовательность случайных событий со счётным числом исходов, характеризующаяся тем свойством, что, при фиксированном настоящем будущее независимо от прошлого [2].

Например, для корректных входных данных тестируемой программы «TEST DATA», представленных в шестнадцатеричном виде как {0x54, 0x45, 0x53, 0x54, 0x20, 0x44, 0x41, 0x54, 0x41} цепь Маркова будет иметь вид, показанный на рисунке 1.

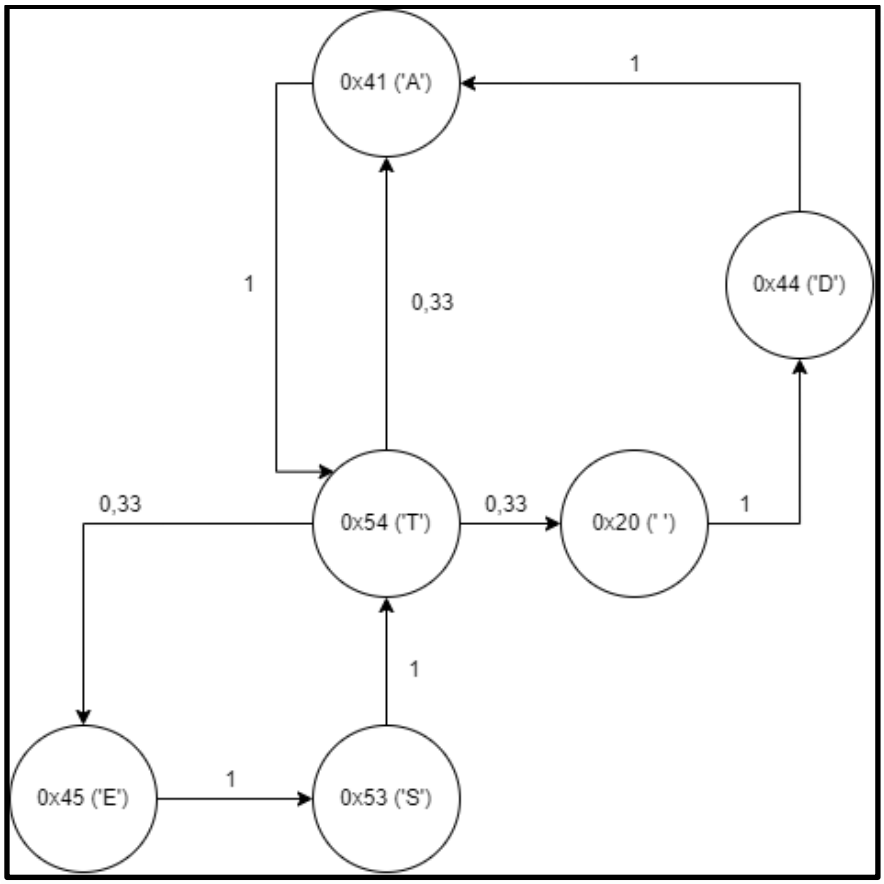

Рисунок 1 – Цепь Маркова для входных данных «TEST DATA»

Матрицей перехода для составленной цепи Маркова входных данных будет являться следующая матрица:

$$P = \begin{matrix} 0 & 0{,}33 & 0 & 0{,}33 & 0 & 0{,}33 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

Затем, по составленной модели можно создать следующие входные данные: «ATESTAT», «TAT DATEST», «TATATATATEST» и т. д.

Предложенный метод позволяет генерировать входные данные тестируемой программы, близкие по структуре к корректным, не имея никаких знаний о тестируемой программе за исключением набора корректных данных.

Список литературы:

1. Valentin J.M. Manes, HyungSeok Han, Choongwoo Han, Sang Kil Cha, Manuel Egele, Edward J. Schwartz, and Maverick Woo. The Art, Science, and Engineering of Fuzzing: A Survey [Электронный ресурс] // https://arxiv.org/pdf/1812.00140.pdf

2. Кельберт М. Я. Вероятность и статистика в примерах и задачах. Т. II: Марковские цепи как отправная точка теории случайных процессов и их приложения. [Текст] / Кельберт М. Я., Сухов Ю. М. – Москва: МЦНМО, 2009. – 587 с.

Павленко Е.Ю., Федоров И.Р.
Санкт-Петербургский политехнический университет Петра Великого

РАЗРАБОТКА МОДЕЛИ ФУНКЦИОНИРОВАНИЯ АДАПТИВНОЙ СЕТЕВОЙ ТОПОЛОГИИ КРУПНОМАСШТАБНЫХ СИСТЕМ НА ОСНОВЕ ДИНАМИЧЕСКОЙ ТЕОРИИ ГРАФОВ

** Исследование выполнено за счет гранта Российского научного фонда № 22-21-20008, https://rscf.ru/project/22-21-20008/*

На сегодняшний день пока что нет оснований говорить об окончательно сложившейся теории динамических сетей (Dynamic network analysis) или сетевой науки (Network science), несмотря на значительный объем эмпирического материала. Для формирования такой отрасли прикладной науки необходима теоретическая основа. Ее ядром может стать динамическая теория графов, основным объектом которой является динамический граф – модель динамической сети [1].

К динамическим сетям относят социальные сети, сети связи, коллективное взаимодействие, структуру фондовых рынков, структуру взаимных обязательств межбанковской системы, а также транспортную систему [2].

В работе предложена модель функционирования сети, построенная на основе динамической теории графов. Динамический граф, как модель динамической сети, представляет собой последовательность “классических” графов, не имеющих параллельных

ребер и петель, переход между которыми описывается различными теоретико-графовыми операциями $\varphi(G_l) = G_{l+1}$.

Для разработки модели функционирования была выбрана матрица смежности, так как в динамическом графе отсутствуют петли, а направление ребра может быть однонаправленным. Данная матрица имеет ряд полезных свойств, в частности, она позволяет быстро получать информацию об смежных вершинах, а ее расширение для использования весов тривиально: вместо логических значений сохраняется значение весов рёбер.

Помимо весов, в модели присутствует матрица реального состояния сети, в которой указывается загруженность ребра в определённый момент времени, что позволяет настроить адаптивное реагирование модели и перераспределение ресурсов системы.

Хранение предыдущих данных сети позволяет использовать их для прогнозирования состояния системы и выявления аномалий, потенциально являющихся угрозой [3] для нормального функционирования системы. Разработанная модель представляет базис для имплементации модели функционирования сетевой топологии крупномасштабных систем на основе динамической теории графов.

Список литературы:

1. Л.И. Сенникова, Р.А. Кочкаров. Некоторые особенности применения динамических графов для конструирования алгоритмов взаимодействия подвижных абонентов [Электронный ресурс] // https://cyberleninka.ru/article/n/nekotorye-osobennosti-primeneniya-dinamicheskih-grafov-dlya-konstruirovaniya-algoritmov-vzaimodeystviya-podvizhnyh-abonentov/viewer.

2. Т.П. Барановская, Д.А. Павлов. Моделирование крупномасштабных транспортных сетей с применением методов многокритериальной оптимизации и учетом структурной динамики [Электронный ресурс] // https://cyberleninka.ru/article/n/modelirovanie-krupnomasshtabnyh-transportnyh-setey-s-primeneniem-metodov-mnogokriterialnoy-optimizatsii-i-uchetom-strukturnoy/viewer.

3. А.В. Уланов, И.В. Котенко. Моделирование адаптивных кооперативных стратегий защиты от компьютерных атак в сети интернет [Электронный ресурс] // http://simulation.su/uploads/files/default/immod-2007-2-211-215.pdf.

Калабишка М.М., Волошина Н. В.
Университет ИТМО, г. Санкт-Петербург

МЕТОД ПОСТАНОВКИ ЦВЗ С ИСПОЛЬЗОВАНИЕМ ВЗВЕШЕННОЙ МОДЕЛИ СТЕГАНОГРАФИЧЕСКОГО КОНТЕЙНЕРА ДЛЯ ЗАЩИТЫ ЦИФРОВЫХ КОПИЙ ДОКУМЕНТОВ

В настоящее время актуальной задачей в области информационных технологий и систем является защита мультимедиа данных, в частности цифровых копий документов. Основанием для данного вывода стало растущее количество судебных разбирательств связанных с нарушением авторских прав и(или) прав правообладателя, а также выросшее число атак, целью которых является перехват персональных данных пользователей различных сервисов, и участившиеся случаи утечки конфиденциальной информации, в том числе персональных данных, различных интернет сервисов.

Одним из самых эффективных методов защиты мультимедиа данных является применение цифровых водяных знаков (далее - ЦВЗ).При этом с целью защиты могут

применяться как видимые водяные знаки, так и не видимые ЦВЗ, для которых используются стеганографические методы встраивания информации. Метки ЦВЗ в основном содержат сведения об авторе и правообладателе, также могут содержать информацию кому предназначалась копия, её назначение, даты и т.п.

В настоящий момент наиболее популярным методом в стеганографии является метод наименьших значащих бит (далее - LSB) [1]. Однако, все большую популярность набирают методы, использующие для встраивания несколько слоев в пикселе MLSB- multilevel least significant bit (далее - MSLB). Известно, что для LSB существуют ограничения связанные с объемом встраивания, например при встраивании максимально заполняется только 12,5% от общего объема файла, для MLSB основной недостаток заключается в том, что если при вставании задействовать высокие битовые плоскости, то за счет вносимых шумов встраивание может стать заметным. Кроме того, существуют различные атаки на скрытые ЦВЗ, одна из самым актуальных атак при защите меток на цифровые копии документов - это атака фальсификации ЦВЗ. Рассматриваемая нацелена на фальсификацию информации об авторе или правообладателе перехваченного цифрового файла, а также информации о целях и сроках его легального использования. Задачей злоумышленника при этом является замена оригинального ЦВЗ на другой, предоставляющий иные права на работу с цифровой копией документа, при сохранении качества результирующей копии документа. Если, злоумышленник предполагает, что встраивание выполнялось с помощью LSB метода, в этом случае атака фальсификации производится простой заменой области LSB на данные фальсифицированной метки. Известно, что такая атака может быть проведена успешно, при сохранении качества результирующего контейнера. Таким образом задача защиты цифровой копии документа не может быть решена с использованием простого LSB встраивания, т.е. с использованием одноуровневой модели контейнера [2].

Для решения задачи защиты от атаки фальсификации возникает необходимость в формировании нового метода встраивания ЦВЗ, основанного на структурах, отличных от LSB. В данной исследовании предложено в качестве нового подхода использовать взвешенную модель стеганографического контейнера (далее -WCE) [3]. В работе рассмотрен подход по формированию структуры взвешенного контейнера, формируемую отдельно для группы документов схожих между собой по визуальным и структурным признакам, позволяющую уменьшить вероятность успешной атаки фальсификации при сохранении качества результирующего стегоконтейнера.

В работе качестве защищаемых типов цифровых копий документов (контейнеров) для встраивания были выбраны следующие документы: цифровая копия паспорта РФ, СНИЛС, ИНН. Выбор контейнеров обоснованы тем, что пользователи постоянно использую данные документы, например в банковской сфере и в сфере страхования. Для проверки эффективности предложенного подхода был проведен ряд практических экспериментов. Так, для каждого класса документов брались три разные незаполненные цифровые копии. В качестве сообщения, которое необходимо встроить, была сформирована псевдослучайная последовательность, имитирующая шифрование сообщения. Процедура встраивания заключается в следующем, берется цифровая копия документа в формате Bitmap, в эту цифровую копию выполняется встраивание данных таким образом, что постепенно формируются все варианты структур контейнера, использующих 5 битовых плоскостей. В результате для одного файла формируется 125 стегоконтейнеров. Для полученных наборов стегоконтейнеров информации определяется один или несколько контейнеров, наиболее подходящих для встраивания по критерию заметности вносимых искажений (PSNR и SSIM). При атаке фальсификации на выбранные структуры контейнеров для всех трех типов документов вероятность успешной фальсификации не превысила 0,008. Экспериментальное исследование показало высокую эффективность предложенного подхода для защиты цифровых копий документов от атаки фальсификации ЦВЗ.

Список литературы:

1. Моркель Т., Элофф Дж. Х. П., Оливье М. С. Обзор стеганографии изображений //ISSA. – 2005. – Т. 1. – № 2. – С. 1-11.

2. Аль-Шатанави О. М., Эль-Эмам Н. Н. Новый алгоритм стеганографии изображений на основе метода MLSB со случайным выбором пикселей //Международный журнал сетевой безопасности и ее приложений. – 2015. – Т. 7. – №. 2. – С. 37.

3. Волошина Н., Жданов К., Беззатеев С. Оптимальное взвешенное нанесение водяных знаков на неподвижные изображения //XIV Международный симпозиум 2014 года по проблемам избыточности в информационно-управляющих системах. – IEEE, 2014. – С. 98-102.

Солдатова А.Ю., Ярмак А.В., Павленко Е.Ю.
Санкт-Петербургский политехнический университет Петра Великого

ОБНАРУЖЕНИЕ МОШЕННИЧЕСТВА С МОБИЛЬНОЙ РЕКЛАМОЙ НА ОСНОВЕ АНАЛИЗА РАБОТЫ ANDROID-ПРИЛОЖЕНИЙ

Высокий рост затрат на мобильную рекламу, выраженный в увеличении объема мирового рынка соответствующей рекламной отрасли с 240 млрд долларов в 2020 году до 290 млрд долларов в 2021 году [1], провоцирует возрастающую активность злоумышленников в данной сфере. Жертвами мошенничества с мобильной рекламой становятся как рекламодатели, теряющие прибыль, так и пользователи устройств, скачавшие рекламное ПО или перешедшие на недоверенный сайт вследствие манипулирования с рекламой в приложении.

Мошенничество с мобильной рекламой – вид киберпреступлений, представляющий собой манипулирование рекламным контентом с целью получения дохода, распространения нежелательного и/или вредоносного ПО. Задачами реализации мошенничества с рекламой являются имитация реального поведения потребителя для обмана рекламодателей и видимость использования легального приложения для рядового пользователя.

В исследовании проведен анализ существующих способов мошенничества с мобильной рекламой, на основе которого введена классификация по уровню реализации механизмов мошенничества, представленная в таблице 1. Уровень трафика соответствует модели «фиктивный пользователь – фиктивное действие», уровень приложения – «реальный пользователь – реальное действие», уровень устройства – «фиктивный пользователь – реальное действие».

Предметом исследования является мошенничество с мобильной рекламой на уровне приложения. Актуальность выбора объекта исследования обусловлена лидирующей позицией устройств под управлением операционной системы Android по охвату рекламного контента [2].

Таблица 1. Классификация механизмов мошенничества по уровням реализации

	Уровень реализации механизмов мошенничества		
	Трафик	Приложение	Устройство
Механизм мошенничества	Подмена SDK	Скрытая реклама	Флуд нажатий
	Подмена идентификатора приложения	Наложение рекламы	Инъекция нажатий

	Эмуляторы	Навязчивая реклама	Сброс идентификатора устройства
	Флуд нажатий через VPN	Перехват нажатий	Кликовые фермы
	Флуд нажатий через proxy-сервера	Флуд нажатий	Боты
		Инъекция нажатий	
		Атака с исправлением UI	
		Подмена SDK	

К типовым методам обнаружения мошенничества с мобильной рекламой можно отнести анализ сетевого трафика. В частности, в работах [3-5] предложены методы обнаружения приложений ОС Android, реализующих преимущественно мошенничество с нажатиями, путем анализа сетевого трафика и поведения приложения без взаимодействия с пользователем. MAdFraud [3] идентифицирует приложения на основе наличия запросов рекламы при работе программы в фоновом режиме и нажатии на рекламный баннер без участия пользователя. AdSherlock [4] на основе предвычисленных шаблонов анализирует трафик для выявления рекламных запросов. FraudDetective [6] определяет классифицирует приложения на основе полной трассировки стека от наблюдаемой активности мошенничества с рекламой до события взаимодействия с пользователем, т. е. трассировки в фоновом режиме. Однако данные работы обнаруживают не все механизмы мошенничества с нажатиями (например, атаку с исправлением пользовательского интерфейса) и не рассматривают мошенничество с показами. С другой стороны, DECAF [6] реализует обнаружение мошенничества с показами для приложений Windows Phone на основе анализа состояний пользовательского интерфейса, однако не учитывает мошенничества с кликами. Таким образом, имеющиеся исследования не позволяют распознавать известные механизмы мошенничества с рекламой в полном объеме.

Для решения задачи обнаружения мобильного мошенничества с рекламой предлагается использовать подход, основанный на исследовании признаков, полученных в результате статического и динамического анализа работы Android-приложения. В рамках доклада подробно рассмотрены:

- требования платформ-магазинов приложений и рекламных сетей к рекламе в распространяемых приложениях,
- примеры реализаций механизмов мошенничества, выделенных при анализе Android-приложений,
- эвристические признаки мошенничества с мобильной рекламой,
- значимые признаки из статического и динамического анализа,
- возможность автоматизации процесса принятия решения о наличии/отсутствии манипулирования с рекламой с помощью методов машинного обучения: влияние типа модели, состава признаков, несбалансированности классов на точность обнаружения,
- преимущества и недостатки представленной реализации.

Таким образом, результаты исследования могут быть использованы для улучшения методологической базы обнаружения мошенничества с мобильной рекламой в Android-приложениях.

Список литературы

1. The New Normal in 2021: Five Things You Need to Know in Mobile // data.ai URL: https://www.data.ai/en/insights/market-data/2021-five-things-you-need-to-know-in-mobile/ (дата обращения: 09.04.2022).
2. iOS и Android: что перспективнее для рекламодателей? // byyd URL: https://www.byyd.me/ru/blog/2020/10/ios-vs-android/ (дата обращения: 01.05.2022).
3. Jonathan Crussell, Ryan Stevens, Hao Chen MAdFraud: Investigating Ad Fraud in Android Applications // MobiSys. - 2014. - №12
4. Mahesh Bathula, Rama Chaithanya Tanguturi, Srinivasa Rao Madala Click Fraud Detection Approaches to analyze the Ad Clicks Performed by Malicious Code // Journal of Physics Conference Series. - 2021. - №2089
5. Joongyum Kim, Jung-hwan Park, Sooel Son The Abuser Inside Apps: Finding the Culprit Committing Mobile Ad Fraud // Network and Distributed Systems Security (NDSS) Symposium. - 2021
6. Bin Liu, Suman Nath, Ramesh Govindan, and Jie Liu. 2014. DECAF: Detecting and Characterizing Ad Fraud in Mobile Apps.. In NSDI. 57–70.

Григорьева Н.М., Платонов В.В.
Санкт-Петербургский Политехнический университет Петра Великого

ЗАЩИТА ОТ СОСТЯЗАТЕЛЬНЫХ АТАК НА СИСТЕМЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ АВТОЕНКОДЕРА

В настоящее время системы распознавания изображений широко применяются во многих областях человеческой деятельности от систем распознавания лиц в метрополитене до работы в беспилотных автомобилях и летательных аппаратах. Многие такие системы справляются с распознаванием образов даже лучше человека, этим и обусловлено их повсеместное внедрение. Данные обстоятельства требуют обеспечения высокой безопасности таких систем, особенно учитывая их высокую уязвимость перед состязательными атаками.

Актуальность исследования заключается в том, что на данный момент для защиты от состязательных атак на фиксированном наборе данных в основном используется обучение системы распознавания (целевой модели) на предварительно созданных состязательных примерах в дополнение к обычному набору изображений. Однако, если набор данных сложно ограничить, то приходится переобучать модель каждый раз, когда становится известно о новом типе атак, что достаточно неудобно и ресурсозатратно. Предлагаемый в работе подход подавать на вход целевой модели изображения, которые уже прошли этап предобработки с помощью автоенкодера, позволит не изменять целевую модель и сделать общую систему модульной и легко масштабируемой.

В ходе работы исследовались различные виды архитектур автоенкодеров в качестве средства защиты от состязательных атак (конкретно, от атак уклонения) на системы распознавания изображений. Были разработаны архитектуры автоенкодеров, которые продемонстрировали их эффективность от рассмотренных атак, и сделан вывод, что автоенкодер является достаточно эффективным средством защиты и повышения устойчивости систем распознавания изображений от состязательных атак. Полученные результаты могут быть использованы в качестве основы для проектирования автоенкодеров более сложных архитектур или целых систем автоенкодеров для дальнейших исследований.

Список литературы:

1. Yassine Bakhti, Sid Ahmed Fezza, Wassim Hamidouche, Olivier Déforges. A Defense Against Adversarial Attacks Using Deep Denoising Sparse Autoencoder // IEEE Access, IEEE. – 2019. – №. 7. – С.160397-160407.

2. A Neuro-Inspired Autoencoding Defense Against Adversarial Perturbations. [Электронный ресурс]. URL: https://arxiv.org/pdf/2011.10867.pdf – (дата обращения: 10.05.2022).

3. Detecting Adversarial Samples from Artifacts. [Электронный ресурс]. URL: https://arxiv.org/pdf/1703.00410.pdf – (дата обращения: 12.05.2022).

Аль-Барри М.Х., Саенко И.Б.
Военная академия связи, Санкт-Петербург

ФОРМИРОВАНИЕ И ИСПОЛЬЗОВАНИЕ ПРИЗНАКОВОГО ПРОСТРАНСТВА ДЛЯ ОБНАРУЖЕНИЯ АНОМАЛЬНЫХ SQL-ЗАПРОСОВ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

В настоящее время методы машинного обучения широко применяются для обнаружения аномалий в различных областях, связанных с безопасностью функционирования сложных систем, включая компьютерные системы и сети. Например, эти методы позволяют достаточно хорошо обнаруживать аномалии сетевого компьютерного трафика, вызванные компьютерными атаками или действиями инсайдеров. Работа пользователей с базами данных также подлежит защите от компьютерных атак и несанкционированного доступа. В этой связи методы машинного обучения, как предполагается, также могут обеспечить эффективное обнаружение аномального поведения пользователей баз данных, которое выражается в присутствии в общем потоке SQL-запросов, поступающих на обработку в СУБД, аномальных запросов. Под аномальными SQL-запросами при этом понимаются такие запросы к базам данных, которые могут быть вызваны компьютерными атаками. Такие атаки обладают определенной спецификой, выделяющих этот тип атак из других. Во-первых, они могут проявляться в форме выполнения специальным образом модифицированных SQL-запросов – SQL-инъекций. В SQL-инъекциях к телу нормального запроса добавляется специальным образом сконструированный код на языке SQL, наносящий вред базе данных или позволяющий извлечь из нее запрещенную информацию. Во-вторых, эти атаки, как правило, являются целевыми. В-третьих, к их числу могут относится атаки «подмены удаленного пользователя», при которой нарушитель может обращаться к базе данных с обычными, не модифицированными запросами. В этом случае реализуется попытка несанкционированного доступа к таблицам базы данных. Этими причинами обусловливается актуальность данного исследования, которое преследует две цели. Первой целью является исследование возможности использования методов машинного обучения для обнаружения аномальных SQL-запросов. Второй целью является исследование возможности оптимизации признакового пространства, используемого в методах машинного обучения, в целях повышения эффективности обнаружения аномальных SQL-запросов.

Наборы данных, которые в настоящей работе использовались в методах машинного обучения, были взяты из регистрационных журналов СУБД. Каждую строку такого журнала можно представить совокупностью трех полей: время запроса; пользователь, являющийся автором запроса; полный текст SQL-запроса. Первоначальное признаковое пространство было сформировано из трех групп признаков. Первая группа признаков была образована из количеств вхождения в SQL-запрос ключевых слов языка SQL. Например, такими словами являлись: SELECT, DELETE, UPDATE, FROM, WHERE и т.д. С помощью значений этих признаков, как предполагалось, можно определить уровень сложности SQL-запроса и его тип. Вторая группа

признаков включала в себя количества вхождений в SQL-запрос специальных конструкций, свойственных SQL-инъекциям. Примерами таких конструкций являются: “Execute”, “1 = 1” и другие. Третья группа признаков была образована количествами вхождения различных имен таблиц данных в SQL-запросы. С помощью этой группы признаков, как предполагалось, можно обнаруживать аномальные SQL-запросы, в которых пользователи предпринимают попытки несанкционированного доступа.

Реализация моделей машинного обучения была сделана в системе Orange 3.32. Эксперименты на полном признаковом пространстве проводились для следующих моделей машинного обучения: SVM, DT, LR, KNN, RF, BN, ANN. Все классификаторы показали точность, превышающую 0,92.

Оптимизация признакового пространства проводилась на основе оценки информативности признаков. Использовались следующие метрики: Info.Gain, Gain ratio, ANOVA, а также нормированное среднее значение этих метрик. Была проведена серия экспериментов, проверяющих возможность повышения точности обнаружения аномальных запросов за счет сокращения количества признаков. Критерий сокращения был следующим – остаются те признаки, метрика информативности которых превышает среднее значение этой метрики по всем признакам. Эксперименты на сокращенном признаковом пространстве подтвердили эффективность предложенного подхода. Точность обнаружения аномальных SQL-запросов повысилась до 0,98.

Таким образом, полученные результаты подтверждают возможность успешного использования методов машинного обучения для обнаружения SQL-запросов и правомерность предложенного подхода к сокращению используемого этими методами признакового пространства. Программная реализация этого подхода и внедрение ее в систему защиты базы данных является дальнейшим направлением исследований.

6. Криптографические методы обеспечения безопасности распределенных систем

Маршалко Г.Б [(1)]., Дали Ф.А. [(1)], Савиных А.Н.[(2)]
[(1)]Академия криптографии Российской Федерации,
[(2)]Московский технический университет связи и информатики (МТУСИ)

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ОБЕЗЛИЧИВАНИЯ ПЕРСОНАЛЬНЫХ ДАННЫХ

Российское законодательство [1] определяет персональные данные (ПнД) как любую информацию, относящуюся к прямо или косвенно определенному или определяемому физическому лицу (субъекту персональных данных). Концепция обезличивания ПнД предполагает наличие методов, преобразующих такие данные к виду, не позволяющему отнести их к определенному субъекту.

Будем рассматривать данные, представленные в виде таблицы $T(A_1, \ldots, A_n)$размеров $m \times n$, где каждая строка $s_j, j = 1, \ldots, m$соответствует субъекту персональных данных, а столбцы – различным атрибутам A_i, которые могут быть уникальными (номер паспорта, СНИЛС), чувствительными (размер зарплаты, диагноз заболевания), неконфиденциальными (не влияющие на безопасность ПнД, что зависит от конкретной ситуации). Для таких данных принято [2] рассматривать три базовых типа угроз:

– выделение субъекта ПнД, когда из набора данных можно выделить строку $s_{j'}$, которая соответствует конкретному субъекту ПнД;

– связывание данных, когда для двух таблиц $T^{(1)}, T^{(2)}$возможно определить, по крайне мере пару строк $s_{j'}^{(1)}, s_{j''}^{(2)}$, соответствующих одному субъекту ПнД.

– определение истинных значений, когда предполагая, что в некоторой ячейке на позиции с номером (i, j) находится значение $t_{ij} = f(x')$, где f– функция, а x' – истинное значение атрибута персональных данных субъекта, по t_{ij}и какой-либо дополнительной информации определить x'.

Собственно задача обезличивания предполагает наличие оператора ПнД, владеющего персональными данными, и аналитика, которому такие данные требуются для проведения каких-либо исследований. Считается, что аналитик не является нарушителем (т. е. не нарушает протокол взаимодействия с оператором), но может проводить анализ данных, в том числе с использованием дополнительной информации, адаптивных запросов к оператору, пытаясь получить новую информацию о записях. Задачей оператора является передача аналитику таким образом обезличенных данных, чтобы последний не мог реализовать указанные выше угрозы.

К настоящему моменту сложилось два подхода к обезличиванию:

– классический, подразумевающий, что оператор преобразует исходную таблицу T, к некоторому обезличенному виду T', и передает ее аналитику;

– сервисный, когда оператор не передает таблицу целиком, а последовательно обезличивает результаты обработки запросов аналитика к исходной таблице T.

Наиболее известными методами, относящимися к первому классу, является метод k-анонимности [3] и его производные (l- разнообразие [4], t- близость [5]).

Суть метода k- анонимности заключается в приведении исходной таблицы Tк такому виду T', что любой набор значений чувствительных атрибутов $A_{i_1}, \ldots, A_{i_k}$, встречается в T', по крайне мере kраз. Для такой таблицы произвольная строка может быт соотнесена с группой, состоящей из не менее чем, kсубъектов. Преобразование реализуется комбинацией методов подавления (удаления строк) и обобщения (например, преобразования номера паспорта 4403 123456 к виду 4403). Очевидно, что для такой таблицы все уникальные идентификаторы должны быть подавлены или обобщены. Вместе с тем, наиболее сложной проблемой является наличие минимальных уникальных комбинаций значений чувствительных атрибутов (МУКЗА), позволяющих однозначно идентифицировать человека. В соответствии с [3], например, такой комбинацией является тройка (почтовый индекс, пол, дата рождения). Задача поиска МУКЗА является вычислительной сложной и решается с помощью переборных алгоритмов [6]. Очевидно, что для сохранения статистических свойств данных таблица T'должна быть получена минимальным числом преобразований исходной таблицы. Показано [7], что эта задача является NP-сложной. В целом, данный класс методов способен работать для баз, где $m \gg n$, т.е. имеющих много записей с относительно малым количеством атрибутов.

Однако наибольшей проблемой, для k- анонимности и ее обобщений является возможность использования аналитиком дополнительной информации для реализации обозначенных угроз. Это может быть сделано за счет:

– обогащения данных: дополнения таблицы T'атрибутами из других таблиц (объединения таблиц);

– использование информации, о наличии связей атрибутов, продиктованной «жизненным опытом» (например, о том, что оклад топ-менеджера существенно превосходит оклад рядового менеджера), некоторых априорных сведений о наблюдаемых объектах.

В целом данный класс методов при отсутствии ограничений на распространение и обогащение таблицы T'не способен обеспечить гарантированную защиту от указанных угроз. Можно утверждать, что они могут обеспечивать защиту таблицы T'при условии ее обработки аналитиком в защищенном контуре без возможности дообогащения.

Сервисная модель, к которой относится статистическое обезличивание [8] (differential privacy), подразумевает, что оператор ПнД реализует интерфейс с набором обезличивающих статистических функций анализа таблицы T. При очередном запросе аналитика на вычисление статистики (математического ожидания, собственных значений и т. п.) оператором вызывается подходящая обезличивающая функция f, которая вычисляет соответствующее значение от элементов таблицы T.

Смысл метода статистического обезличивания заключается в том, чтобы подобрать вид обезличивающих функций fтаким образом, чтобы аналитик не мог статистически (при заданном пороге ε) отличить результаты их работы на исходной таблице Tи на таблице T', отличающейся от Tодной записью$T^{(1)} = (T, s)$, т. е. фактически не смог соотнести результат с конкретным субъектом. Технически это достигается наложением аддитивного шума в процессе вычисления функции, параметры которого зависят от вида функции f, размера таблицы T, максимально допустимого количества запросов аналитика на вычисление f и требуемой точности (разницы значений функции на исходных и зашумленных данных)

Минусами данного подхода является то, что:

– он работает с числовыми данными;

– необходима разработка обезличивающего варианта каждой используемой аналитиком функции.

– необходим контроль числа попыток обращения к обезличивающему аналитика к оператору.

С другой стороны, такой подход:

– существенно менее чувствителен к наличию дополнительной информации, по сравнению с классическими методами;

– позволяет получить численные оценки уровня обезличенности данных (границу на максимальное число запросов к таблице T, после которого будет возможна реализация рассматриваемых угроз);

– позволяет получить численную оценку безопасности композиции статистически обезличивающих функции и, следовательно, допускает масштабирование;

– позволяет отказаться от необходимости передачи баз ПнД третьим лицам.

В целом, именно этот подход представляется более перспективным для реализации в системах, обрабатывающих большие объемы данных.

Список литературы:

1. Федеральный закон «О персональных данных» от 27.07.2006 № 152-ФЗ
2. Opinion 05/2014 on Anonymisation Techniques, Article 29 data protection working party, 0829/14/EN, WP216.
3. Sweeney L. k-anonymity: A model for protecting privacy //International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. — 2002. — Т. 10. — №. 05. — С. 557-570.
4. A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. In Proc. 22nd Intnl. Conf. Data Engg. (ICDE), page 24, 2006.
5. Li, Ninghui et al. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. 2007 IEEE 23rd International Conference on Data Engineering (2007): 106-115.
6. Elliot M. J., Manning A. M., Ford R. W. A computational algorithm for handling the special uniques problem //International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. — 2002. — Т. 10. — №. 05. — С. 493-509.
7. P. Bonizzoni, G.D. Vedova, R. Dondi, the k-anonymity problem is hard, Proc. 17th International conference on fundamentals of computation theory, 2007.
8. Dwork C. et al. The algorithmic foundations of differential privacy //Found. Trends Theor. Comput. Sci. – 2014. – Т. 9. – №. 3-4. – С. 211-407.

Кустов Е.Ф.[(1)], Леевик А.Г. [(1)], Голованов А.А. [(1)], Беззатеев С.В. [(1,2)]
[(1)]Университет ИТМО, Санкт-Петербург
[(2)]Санкт-Петербургский Государственный Университет Аэрокосмического Приборостроения, Санкт-Петербург

К ВОПРОСУ ОБ ИСПОЛЬЗОВАНИИ ПОСТКВАНТОВЫХ СХЕМ ЭЛЕКТРОННОЙ ПОДПИСИ CRYSTALS DILITHIUM И FALCON

Электронные подписи повсеместно используются в информационных системах для обеспечения целостности, подтверждения авторства и неотказуемости информации, например, в системах электронного документооборота, системах распределённого реестра, банковских системах и т. д.

На сегодняшний день безопасность большинства схем электронной подписи основана на задачах дискретного логарифма и факторизации чисел. Однако, данные алгоритмы станут неприменимы, когда появятся достаточно мощные квантовые компьютеры, способные решить данные задачи за полиномиальное время, используя алгоритм Шора [1]. На данный момент компанией IBM уже разработан 20-кубитный квантовый компьютер [2], и масштабирование квантовой системы до размеров, которые представляли бы угрозу используемым схемам электронной подписи, это вопрос ближайших десятилетий.

В целях противодействия данной угрозе активно исследуются алгоритмы так называемой постквантовой криптографии. Её основными направлениями являются:

- криптография на кодах, исправляющих ошибки;
- криптография на теории решёток;
- криптография на многомерных уравнениях;
- криптография на хеш-функциях;
- криптография на изогениях эллиптических кривых.

Национальный институт стандартов и технологий США в целях стандартизации постквантовых алгоритмов проводит конкурс и к третьему раунду осталось всего три схемы электронной подписи. Среди них присутствуют две схемы, построенные на теории решеток (Falcon [3] и CRYSTALS Dilithium [4]), и одна схема, построенная на многомерных уравнениях (Rainbow [5]).

В недавней работе был проведён успешный криптоанализ схемы Rainbow, который поставил под вопрос степень её безопасности [6]. В свете вышеизложенного оправданным является использование электронной подписи, построенной на решётках.

Схема электронной подписи Falcon является модификацией схемы NTRUSign, использующей специальные NTRU-решётки. Также в данной схеме используется быстрые преобразования Фурье (FFT), которые сокращают время работы алгоритмов подписи и верификации. Безопасность схемы основывается на проблеме short integer solution (SIS), которая подразумевает сложность нахождения вектора кратчайшей длины, являющегося решением системы линейных уравнений.

В схеме CRYSTALS Dilithium представлен другой подход к построению электронной подписи на решётках. Безопасность алгоритмов в CRYSTALS Dilithium основана как на проблеме SIS, так и на проблеме обучения с ошибками (Learning With Errors, LWE). Данная схема построена на решётке в кольце многочленов по некоторому модулю и для ускорения вычислений авторы используют дискретное преобразование Фурье (NTT).

В данной работе мы провели анализ используемого математического аппарата представленных выше схем и сравнили их по следующим показателям:

- размеры ключей;
- размеры подписей;
- время работы алгоритмов генерации ключевой пары, подписания и верификации подписи.

По результатам анализа обе схемы обладают своими преимуществами и недостатками и нельзя определить среди них фаворита. Выбор той или иной схемы зависит от требований к конкретной задаче.

Список использованных источников:

1. Shor PW. Algorithms for quantum computation: discrete logarithms and factoring. InProceedings 35th annual symposium on foundations of computer science 1994 Nov 20 (pp. 124-134).
2. Vu C. IBM announces advances to IBM quantum systems & ecosystem. – 2017.
3. Fouque P. A. et al. Falcon: Fast-Fourier lattice-based compact signatures over NTRU //Submission to the NIST's post-quantum cryptography standardization process. – 2018. – Т. 36. – №. 5.
4. Lyubashevsky V. et al. Crystals-dilithium //Submission to the NIST Post-Quantum Cryptography Standardization [NIS]. – 2017.
5. Ding J., Schmidt D. Rainbow, a new multivariable polynomial signature scheme //International conference on applied cryptography and network security. – Springer, Berlin, Heidelberg, 2005. – С. 164-175.
6. Beullens W. Breaking rainbow takes a weekend on a laptop //Cryptology ePrint Archive. – 2022.

Трифонов С.Е.[(1)], Лекарь Л.А.[(2)]

[(1)]АО «ПНИЭИ», г. Пенза, [(2)]ФКУ НПО «СТиС» МВД России, г. Москва

ЭФФЕКТИВНОЕ РЕШЕНИЕ ВОПРОСА КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ НА ОСНОВЕ ОТЕЧЕСТВЕННЫХ СОВРЕМЕННЫХ ВЫСОКОТЕХНОЛОГИЧНЫХ ИЗДЕЛИЙ

Суть решения заключается в раздельной реализации мобильных устройств телекоммуникаций (средства радиосвязи, средства удаленного широкополосного доступа на базе планшета и смартфона) и съемных легко отчуждаемых средств криптографической защиты информации (СКЗИ)

СКЗИ реализуется в форм-факторе microCD-карты и сертифицируется в системе ФСБ России по классам КС1-КС3 (в мобильных устройствах под управлением ОС «Аврора»)

Существующая практика программной реализацией СКЗИ , а также аппаратной реализацией СКЗИ на компонентной элементной базе приводит к необходимости сертифицировать полностью мобильные устройства телекоммуникаций со всеми вытекающими проблемами.

Рассматриваемые решения позволяют сертифицировать отчуждаемые СКЗИ в микроэлектронном исполнении как доверенную вычислительную среду, а для мобильного

устройства предъявляются требования оценки влияния и к ним не предъявляются специальные требования в полном объеме.

СКЗИ предназначены для построения современных конфиденциальных сетей передачи информации с минимальными организационно-техническими, эксплуатационными и финансовыми затратами. При этом обеспечивается:

— доверенная среда хранения идентифицирующих и аутентифицирующих параметров, параметров, паролей, ключей, файлов и данных пользователя с защитой от НДС;

— доверенная среда хранения, выгрузки и запуска СПО в среде мобильных платформ;

— легко отчуждаемое (съемное) шифровальное средство; промежуточный носитель. защищенный от НСД; загрузочный диск с доверенной ОС; криптографический токен с интерфейсом PKCS#l 1

В архитектуре microCD-карты используется микроконтроллер «Курган», обеспечивающий производительность до 2 Мбит/сек в режимах криптографической обработки, а также имитозащиту информации согласно ГОСТ Р 34.12-2015 «Магма» и

ГОСТ Р 34.13-2015.

Для реализации конфиденциальных сетей передачи информации обеспечена встречная работа СКЗИ microCD-карты с криптомаршрутизатором серверной инфракструктуры и станцией генерации ключей, имеющими действующие сертификаты ФСБ России.

В отличии от распространенных решений ViPNet компании Инфотекс рассмотренная реализация мобильной компоненты защищенной системы передачи информации имеет преимущества:

— в СКЗИ на аппаратном уровне реализованы современные отечественные криптоалгоритмы в соответствии с требованиями регулятора;

— СКЗИ является самостоятельным аппаратным изделием, подключаемым к мобильному устройству, следовательно весь комплекс соответствующих «тяжелых» специальных требований по учету и использованию СКЗИ будет применяться только к специализированной microCD-карте, а не к мобильному устройству в целом.-в данном случае смартфон и планшет будут рассматриваться регулятором только в качестве среды функционирования СКЗИ, требований к которой значительно меньше. При этом будет обеспечиваться возможность использования данных мобильных устройств без подключенных СКЗИ в штатном режиме. Таким образом, существенно улучшаются эксплуатационные характеристики мобильной компоненты;

— в качестве мобильных устройств используются смартфоны и планшеты отечественных производителей, функционирующие под управлением отечественной ОС Аврора. Такой подход обеспечит возможность сертификации СКЗИ по наивысшему классу защищенности КСЗ.

Таким образом, современные высокотехнологичные средства криптографической защиты информации на основе специализированной элементной базы отечественной разработки позволит разработать на их основе новое поколение средств защищенной передачи информации, качественно отличающиеся практически по всем тактика-техническим характеристикам от современных аналогов

Петренко А.С., Петренко С.А.
СПбГЭТУ «ЛЭТИ», г. Санкт-Петербург

МЕТОДИКА ОЦЕНКИ КВАНТОВОЙ УСТОЙЧИВОСТИ СОВРЕМЕННЫХ БЛОКЧЕЙН-ПЛАТФОРМ

Актуальность методики объясняется тем, что применение известных механизмов консенсуса (более распространен *Proof-of-Work)* и криптографических примитивов (*SHA256, Ethash, Kessak, blake256, RSA, DSA, ECDSA, ГОСТ Р 34.10* и др.) блокчейн-платформ Цифровой экономики Российской Федерации уже недостаточно для нейтрализации *квантовой угрозы* и обеспечения требуемой *квантовой устойчивости*. Здесь под квантовой устойчивостью понимается некоторое системное свойство блокчейн-платформы (*Waves Enterprise (Waves, Vostok), Hyperledger Fabric (Linux, IBM), Мастерчейн (Сбербанк), Microsoft Azure Blockchain, Enterprise Ethereum Alliance и др.*), интуитивно определяемое как *способность сохранять постоянство (неизменность) упомянутой платформы и ее ключевых системных свойств* в условиях криптографических атак с использованием квантовых компьютеров.

Основные этапы методики. Предлагаемая методика содержит следующие этапы.

Этап 1. Решение задачи факторизации. На этом этапе учтена вероятностная природа известного квантового алгоритма факторизации Шора [1-4]. Здесь первый источник случайности встроен в классическое вероятностное сведение разложения на множители к нахождению периода некоторой функции. Второй источник появляется из необходимости наблюдения квантовой памяти, которое также выдаёт случайные результаты. Поэтому, были определены следующие шаги решения задачи факторизации:

1) Выбор случайного остатка a по модулю N;
2) Проверка $НОД(a, N) = 1$;
3) Нахождение порядка r остатка a по модулю N;
4) Если r четен, вычислено $НОД(a^{r/2} - 1, N)$.

Задача разложения числа N на множители, свелась к нахождению периода r для случайно подобранного числа a.

Этап 2. Определение периода функции. На этом этапе с помощью преобразования *Фурье* не потребовалось вычислять все значения $f(x)$. Задача свелась к решению, похожему на решение задачи *Дойча*, в которой учитываются не все значения функции, а только некоторые её свойства [2-4]. В результате, полученные представления *Фурье-преобразования* в форме произведения (рис. 1) позволили промоделировать требуемые квантовые цепи (рис. 2) для работы с квантовым компьютером IBM *Q (20, 100 и 170 кубитов)*, который был выбран для апробации предлагаемой методики.

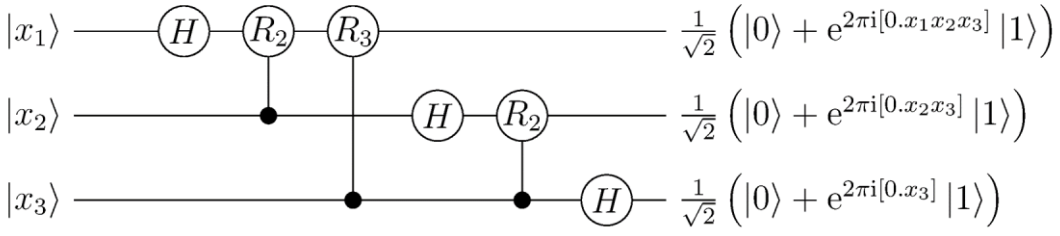

Рисунок 1 - Гейтовое представление квантового преобразования Фурье

Здесь *гейт* R_k обозначает унитарное преобразование вида:

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix}$$

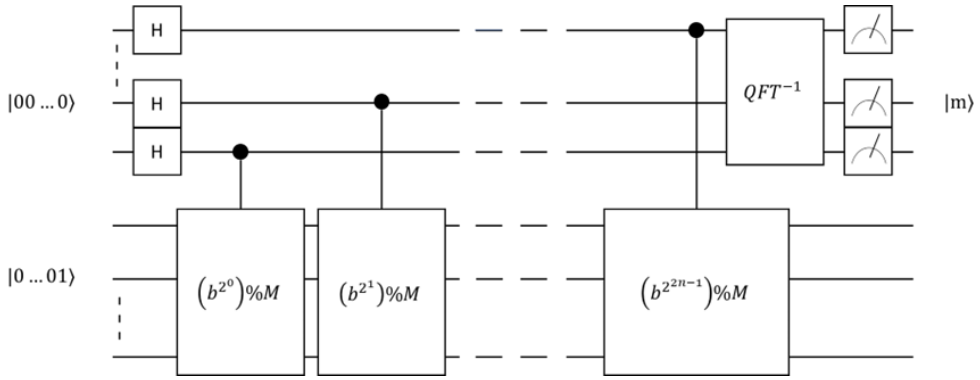

Рисунок 2 - Квантовое представление модифицированного алгоритма факторизации Шора

Этап 3. Собственно оценка квантовой устойчивости блокчейн-платформ. Для этого была опробована схема прямого подключения к квантовой (16, 20 и 100-кубитной системе *IBM Q)* с помощью платформы *IBM Cloud.* Получен доступ к *IBM Quantum Experience и осуществлен з*апуск соответствующее приложения на квантовой схеме для работы с отдельными кубитами. Также были апробированы возможные гибридные схемы из квантовых вычислителей *IBM Q* и симуляторов на суперЭВМ пятого поколения (*СуперЭВМ Ломоносов-2, Торнадо, СКИФ, вычислители на ПЛИС*). При этом были задействованы специально разработанные известные и авторские библиотеки для моделирования квантовых алгоритмов на квантовых схемах.

Оценка результативности методики. К достоинствам предлагаемой методики относится повышение на *20%-30% оперативности,* а также повышение на *10%-15% результативности* исследования криптостойкости схем асимметричного шифрования (*RSA, Эль-Гамаля*), цифровой подписи (*DSA, ECDSA или RSA-PSS*) и хеш-функций **(***SHA256, Ethash, Kessak, blake256 и др.)* расмотренных блокчейн-платформ.

К дальнейшим возможным направлениям развития разработанной методики относится:

– снижение квантовой декогеренции путём использования специальных исправляющих алгоритмов, в том числе, на основе методов коррекции ошибок;

– добавление в библиотеку платформы других квантовых и пост-квантовых алгоритмов (*Гровера, Саймона, Экера, Берштейна-Вазирани, Харроу, Хассилима, Ллойда, Денисенко, Ключкарева, Гребнева, Федотова, Чижова, Бельского и* др.);

– совершенствование методов определения пригодности криптосистем для криптоатаки;

– добавление в объект криптоанализа новых криптосистем и примитивов других блокчейн-платформ.

Список литературы:

1. Корольков А.В. О некоторых прикладных аспектах квантовой криптографии в контексте развития квантовых вычислений и появления квантовых компьютеров. / А.В.

Корольков // Вопросы кибербезопасности № 1(9) – 2015. – М.: Журнал «Вопросы кибербезопасности», 2015. – с. 6-13.

2. Петренко А.С., Романченко А.М. Перспективный метод криптоанализа на основе алгоритма Шора// Защита информации. Inside №2 2020. – СПб.: Изд. Афина, 2020. – с. 17–23.

3. Shor P. Algorithms for quantum computation: discrete logarithms and factoring [Text] /Shor P.// Foundations of Computer Science.—1994.—№10. —134p.

4. Shor P. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Computing 26, 1484 – 1509 (1997).

Карантаев В.Г., Латышов К.В.
Национальный исследовательский университет «МЭИ», г. Москва

ЧАСТНЫЕ ВОПРОСЫ РЕАЛИЗАЦИИ ВСТРОЕННЫХ КРИПТОГРАФИЧЕСКИХ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ ПРИ МЕЖСЕТЕВОМ ВЗАИМОДЕЙСТВИИ УСТРОЙСТВ ЦИФРОВОЙ ПОДСТАНЦИИ С ИСПОЛЬЗОВАНИЕМ СТЕКА ПРОТОКОЛОВ МЭК 61850-8-1 (MMS)

Вопросам анализа влияния компьютерных атак на подсистемы цифровых подстанций (ЦПС), устойчивости их функционирования, реализации встроенных средств защиты информации, в частности, криптографической защиты информации был посвящен ряд публикаций и докладов в период с 2017-2021 годы [1-4]. Вопросы создания современных микропроцессорных интеллектуальных электронных устройств (ИЭУ) подсистемы релейной защиты и автоматики ЦПС неразрывно связаны с уровнем развития нормативно-технических документов (НТД) и оценкой текущего состояния международных и национальных стандартов. Вопросам анализа актуальных НТД посвящены публикация и доклад на расширенном заседании исследовательского комитета В5 РНК СИГРЭ, состоявшегося 21.12.2021.

В последние несколько лет можно отметить следующие тенденции в развитии ИЭУ РЗА:

- Разработка ИЭУ с новыми видами аппаратной и программной архитектуры: модульные, многопроцессорные;
- Реализация РЗА с новыми видами программной архитектуры: мультиагентные распределенные системы;
- Переход к использованию встраиваемых операционных систем;
- Разработка и использование в ИЭУ механизмов безопасности, образующих встроенные средства защиты информации (СЗИ);

Существующие тенденции в развитии ИЭУ и систем РЗА приводят к необходимости анализа актуальных техник и тактик нарушителя и обосновывают необходимость исследований в области реализации криптографических методов защиты в автоматической технологической системе, встраиваемых средств криптографической защиты информации (ВСКЗИ).

Например, применение криптографических протоколов TLS 1.2, TLS 1.3 позволит предотвратить ряд техник и тактик MITRE ICS AT&T, использованных в ходе атаки Industroyer [5]. К ним относятся: тактика Discovery и соответствующие ей техники Control

Device Identification, Network Connection Enumeration, а также тактика Collection и техники Automated Collecion, Role Identification.

Целью настоящей работы было исследование нефункциональных свойств работы протокола MMS при защищенном межсетевом взаимодействии.

На специально разработанном стенде исследовались свойства прототипа криптографически защищенного протокола, с использованием набора криптографических протоколов TLS 1.2.

В Таблице 1 приведена информация о шифр наборах и основных критических операциях.

Таблица 1 – Затраты времени на различные алгоритмы шифрования

Шифр набор	**Установка соединения, мс**	**Проверка сертификата, мс**	**Шифрование, мс**	**Дешифрование, мс**
TLS_RSA_WITH_AES_128_CBC_SHA	101,821	0,001201	0,015668	0,035625
TLS_RSA_WITH_AES_256_CBC_SHA	107,962	0,001412	0,015681	0,041221
TLS_RSA_WITH_AES_128_CBC_SHA256	130,884	0,002176	0,015984	0,050579
TLS_RSA_WITH_AES_256_CBC_SHA256	139,963	0,004838	0,019468	0,052833

Выводы:

- Исследования известных угроз объектов критической информационной инфраструктуры ЦПС и технико-экономическое сравнение СЗИ позволяет сделать вывод о необходимости преимущественного использования ВСЗИ и ВСКЗИ в ИЭУ РЗА ЦПС средних классов напряжения.
- Необходимо продолжить научно-прикладные исследования и организовать НИОКР по реализации прототипа защищенного протокола MMS с использованием TLS 1.3, определённого в отечественных рекомендациях по стандартизации.
- Требуется продолжить исследование нефункциональных свойств защищенного протокола MMS. Организовать штормые испытания на полунатурном стенде ЦПС.
- Необходимо разработать отраслевые требования: методики проектирования подсистем ЦПС в защищенном виде с использованием ВСЗИ и ВСКЗИ, методики оценки нефункциональных свойств ИЭУ РЗА при использовании ВСЗИ и ВСКЗИ, методики приемо-сдаточных испытаний.

Список литературы:

1. Частные вопросы реализации средств криптографической защиты информации для защиты подсистемы релейной защиты и автоматики. Карантаев В.Г., к.т.н. ОАО «ИнфоТеКС», Россия. Сборник аннотаций докладов международной Конференции «Цифровая подстанция. Стандарт IEC 61850» 3–5 октября 2017 г., Москва. http://iec61850ru.ru/upload/docs/Sbornik_annotacii_dokladov_conferencii.pdf

2. Карантаев, В. Г. Вопросы кибербезопасности в меняющейся электроэнергетической отрасли / В. Г. Карантаев // Релейщик. – 2019. – № 1(33). – С. 48-51. – EDN HPJXIA.

3. Карантаев, В. Г. Возможные методы анализа последствий влияния кибератак на системы РЗ и па цифровых и высокоавтоматизированных подстанций / В. Г. Карантаев, В. И. Карпенко // Электроэнергетика глазами молодежи: Материалы XI Международной научно-технической конференции. В 2-х томах, Ставрополь, 15–17 сентября 2020 года. – Ставрополь: Северо-Кавказский федеральный университет, 2020. – С. 62-65.

4. Карантаев В.Г. Вопросы реализации доверенных Интеллектуальных Электронных Устройств // Релейная защита и автоматика энергосистем. — 2021. — Том 1 — С. 336–343.

5. ESET, Win32/Industroyer: новая угроза для промышленных систем управления, 2017 г.

Поликарпов А.А.
ООО «Специальный Технологический Центр», Санкт-Петербург

ПРОТОТИПИРОВАНИЕ ПРОТОКОЛОВ CRISP И PROTOQA

Общие сведения

В феврале — марте 2022 года в СТЦ была проведена работа по прототипированию протоколов CRISP и ProtoQa на основе МР 26.4.001-2019, Р 1323565.1.029-2019, МР 26.4.004-2021, ГОСТ 34.12-2015, ГОСТ 34.13-2015, ГОСТ 34.13-2018, Р 1323565.1.017-2018.

Целью работы была проверка практической реализуемости протоколов CRISP и ProtoQa на языке С для дальнейшего встраивания в разрабатываемую СТЦ аппаратуру.

В результате работы были написаны программные модули, реализующие сборку/разборку пакетов протокола CRISP, сборку/разборку пакетов протокола ProtoQa, создание и вскрытие ключевых контейнеров и, используя эти модули, написаны две утилиты:

— имитатор СВРК — однопоточный UDP сервер, обрабатывающий запросы стека протоколов в соответствии с жёстко заданными в исходном коде данными (ключами, данными для передачи, идентификационными данными);

— имитатор СКЗИ — UDP клиент, подключающийся к имитатору СВРК, создающий сессию в соответствии с протоколом ProtoQa, запрашивающий ключ и случайные данные в этой сессии и закрывающий сессию после обмена.

Описание стека протоколов

Полезная нагрузка (ключ/ключи или случайные данные) передаются в ключевом контейнере — зашифрованной и снабжённой имитозащитой структуре данных.

Ключевой контейнер передаётся в пакете протокола ProtoQa. Протокол сессионный, параметры сессии устанавливаются в начале обмена и задают метод идентификации ключей, алгоритм шифрования, используемый ключевым контейнером, декларируют возможности участников обмена. Пакеты протокола несут адресную информацию и нумеруются, что даёт возможность обрабатывать различные запросы независимо друг от друга и отбрасывать повторные пакеты.

Протокол ProtoQa использует в качестве транспорта пакеты протокола CRISP. Протокол CRISP использует симметричное шифрование на заранее распределённых ключах. Каждый пакет протокола CRISP содержит идентификатор алгоритма шифрования и идентификатор ключа, с помощью которых он зашифрован. Идентификатор ключа содержит идентификаторы отправителя, получателя и номер производного ключа. Так же пакет CRISP протокола содержит порядковый номер пакета, что позволяет отбрасывать повторные и устаревшие пакеты.

Пакеты протокола CRISP передаются через любой протокол транспортного уровня, например UDP. Фактически, CRISP является надстройкой над протоколом транспортного уровня, дополняющий транспортный уровень функциями шифрования, имитозащиты и защиты от повторов.

Порядок использования стека протоколов

Порядок использования стека протоколов со стороны СКЗИ:

1. открыть сессию и согласовать параметры сессии;

2. запросить ключи или случайные данные;

3. закрыть сессию.

Порядок использования стека протоколов со стороны СВРК:

1. ожидать открытия сессии со стороны СКЗИ;

2. согласовать параметры сессии с СКЗИ;

3. по запросам СКЗИ в рамках открытой сессии отдавать ему случайные данные или ключи, при необходимости вырабатывая новые ключи;

4. в случае выработки новых ключей оповестить о готовности ключей второй СКЗИ, парный тому, что запросил новые ключи;

5. принять закрытие сессии с СКЗИ.

Предварительные условия для использования стека протоколов

Использование стека протоколов CRISP/ProtoQa подразумевает предварительное установление соединения по какому-либо протоколу транспортного уровня.

Для работы стека протоколов требуется предварительное распределение ключей для каждой пары СКЗИ-СВРК.

Для работы по протоколу CRISP требуется, чтобы на передающее и принимающее устройство был распределён базовый ключ, из которого при работе будут выработаны ключи шифрования и имитозащиты. Для передачи в обратном направлении должен быть распределён ещё один базовый ключ.

Для использования криптоконтейнера требуется, чтобы на передающее и принимающее устройство была распределена ключевая информация, необходимая и достаточная для формирования ключей, использующихся в алгоритмах экспорта и импорта.

Выводы и вопросы

Для реализации программных модулей, осуществляющих обработку протоколов CRISP и ProtoQa, работу с криптоконтейнерами, информации в стандартах достаточно.

Для реализации утилит имитаторов СКЗИ и СВРК, которые затем были бы пригодны для отладки разрабатываемой аппаратуры или отладки сопряжения с аппаратурой, разрабатываемой другими организациями, необходимо определить:

1. правила формирования меток целевых ключей;

2. правила формирования идентификаторов целевых ключей;

3. правила формирования ключей, использующихся в алгоритмах экспорта и импорта при обработке ключевых контейнеров.

Если такие правила не определены, невозможно гарантировать совместимость СКЗИ и СВРК, разрабатываемых разными разработчиками.

В ходе работы возникли вопросы, ответы на которые отсутствуют в стандартах:

1. Сессия протокола ProtoQa поддерживается постоянно всё время работы СКЗИ, или открывается и закрывается по мере надобности?

2. Как СВРК уведомляет СКЗИ о выработке для него новых ключей по запросу другого СКЗИ, если в данный момент нет открытой сессии?

Воробьев Е.Г., Альшанская Т.В.
СПбГЭТУ «ЛЭТИ», Самарский университет

ПЕРСПЕКТИВЫ РАЗВИТИЯ КВАНТОВЫХ ТЕХНОЛОГИЙ В ПРОМЫШЛЕННОСТИ РОССИИ

Геополитическая ситуация в мировом сообществе, вторая квантовая революция способствуют мобилизации ресурсов ведущих исследовательских центров и коммерческих компаний в нашей стране. Технологии на основе принципов квантовой механики применяют для разработки инновационных решений на новом квантовом уровне. В России комплексное развитие квантовых технологий представлено «дорожной картой», разработанной Госкорпорацией «Росатом» (2019г. – 2024 г.), в составе национальной программы «Цифровая экономика» и включает три направления: вычисления, коммуникации и сенсоры. Кроме того, в 2015 г. принята «дорожная карта» Национальной технологической инициативы (НТИ) РФ, как долгосрочная программа частно-государственного партнерства по содействию развитию новых перспективных рынков на базе высокотехнологичных решений. Создан Центр квантовых коммуникаций. В рамках этих инициатив запущен проект Сэйфнет как «кокон безопасности» для IoT, беспилотного транспорта, «умных» городов, систем энергоснабжения и здравоохранения также основан на применении к 2035 года квантовых технологий для реализации вопросов информационной безопасности, в частности критических информационных инфраструктур.

В Реализацию этих направлений, подписав соответствующие соглашения о намерениях с правительством РФ, разделили между собой крупнейшие компании с государственным участием. Квантовыми вычислениями, то есть разработкой отечественного квантового компьютера, занимается «Росатом», квантовыми коммуникациями – О АО РЖД, сенсорами «Ростех». Считается, что вторая волна практического применения квантовых технологий включает несколько направлений, перспективных для бизнеса (по материалам западных и российских источников):

1. Квантовые вычисления (компьютеры), возможные области применения:

- для управления движением транспортных средств (воздушных, морских, наземных),
- прогнозирования погоды,
- предупреждения чрезвычайных ситуаций,
- в оборонной, горнодобывающей и автомобильной промышленности, медицине и других отраслях.

В настоящее время разработаны 7 языков квантового программирования, делаются попытки внедрения их в учебный процесс ВУЗов РФ. Также производятся квантовые симуляторы для

исследовательских целей и доказательства превосходства квантовых компьютеров над обычными при решении вычислительно сложных задач.

В России разработана первая интегральная схема на базе пяти сверхпроводниковых кубитов в держателе. созданная специалистами Московского физико-технического института (МФТИ) в Лаборатории искусственных квантовых систем (ЛИКС). По прогнозу Market Research Future, среднегодовой темп роста рынка квантовых вычислений в 2022-2023 годах составит 34% и достигнет 2,82 млрд долларов.

2. Защищенные квантовые коммуникации:

Квантовые коммуникационные сети подразумевают передачу квантовой информации между двумя удаленными в пространстве квантовыми системами способом, защищенным от злоумышленников. Интерес к квантовым коммуникациям отчасти связан с развитием big data – все эти большие данные нужно обрабатывать и передавать с высокой скоростью и защищенным способом. По прогнозам IDC, к 2023 году почти 60% данных будут обрабатываться с использованием облачных сервисов. Это существенно увеличит потребность в квантовых коммуникациях и шифровании, поскольку значительное количество передаваемых в облако данных должно быть защищено. Квантовая криптография может применяться для защиты данных в коммуникационных сетях различного назначения, включая спутниковые каналы передачи данных. Большие перспективы откроет появление устройств квантовой криптографии на рынке для владельцев ЦОД, банков, телекоммуникационных компаний, интернет провайдеров.

В декабре 2017 года Сбербанк и РКЦ запустили квантовую сеть протяженностью 25 км. В сентябре 2019 года Казанский квантовый центр, «Ростелеком» и «Таттелеком» успешно провели эксперимент по распределению квантовых ключей на волоконно-оптической линии связи (ВОЛС) протяженностью 143 километра. В этом эксперименте использовался криогенный однофотонный детектор российского производителя СКОНТЕЛ и система квантового распределения ключей университета ИТМО. В настоящее время разработана и протестирована квантовая линия: Санкт-Петербург-Москва, закладываются направления Москва-Казань и Казань-Самара. Всего сейчас имеются шесть готовых к эксплуатации протоколов для систем квантового распределения ключей: 3 США, 1 Швейцария, 1 Китай, 1 Россия.

По прогнозу Markets and Markets, ожидается, что объем мирового рынка квантовой криптографии вырастет со 101 мил. долларов в 2018 году до 506 миллионов долларов к 2023 году при среднегодовом темпе роста 37,9%. Как считают аналитики CIR, квантовый интернет быстро становится реальностью, и к 2023 г. принесет 3,7 млрд. долларов в виде оборудования и услуг.

3. Квантовые сенсоры. Основное применение ожидается в медицине и биологии: анализ генома, диагностика заболеваний, в том числе онкологических, исследование процессов, происходящих в теле человека, внутренних органов, тканей, клеток и молекул. Кроме того, высокочувствительные датчики нового поколения будут применять и в других областях: навигация (космическая отрасль, беспилотный транспорт), оборона и безопасность, геологоразведочные работы, нефтедобыча и строительство, технологии интернета вещей. Другим примером коммерциализации этих приборов служит счетчик фотонов, разработанный в Московском педагогическом государственном университете (МПГУ). Его использует РКЦ для своих разработок в сфере квантовой криптографии. Квантовые сенсоры могут использоваться в системах квантовой криптографии для обеспечения случайности квантовых ключей. Например, квантовый генератор случайных чисел, который создали физики МГУ в 2017 году. По оценке BCC Research, объем рынка квантовых сенсоров вырастет с 161 млн. долларов в 2019 году до 299.9 млн. долларов в 2024 году.

Область квантовых технологий особенно привлекательна не только перспективой бурного роста в ближайшие годы, но и отсутствием стандартов. Квантовый консорциум, принимает

участие в разработке и внедрении, что способствует развитию рынка квантовых систем связи на территории РФ, кроме того, есть ресурсы позволяющие обойти его нынешних лидеров.

Фомин Д.Б.
Национальный исследовательский университет «Высшая школа экономики»

ОДНА МОДЕЛЬ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ КОНФИДЕНЦИАЛЬНОЙ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

В настоящее время все активнее развивается прикладная область анализа информации, связанная с обработкой большого количества данных в том числе методами машинного и глубинного обучения. Одним из способов гарантирования конфиденциальности обрабатываемой такими методами информации является использование алгоритмов гомоморфного шифрования, идея которых впервые была предложена в 0.

Введем необходимые обозначения. По 0 рассмотрим вероятностную модель алгоритма шифрования $\Sigma = (X, K, Y, E, D, P_X, P_K)$, где X – множество открытых текстов, Y – множество шифртектов, $E: K \times X \to Y$ – функция зашифрования, $D: K \times Y \to X$ – функция расшифрования, P_X, P_K – соответственно вероятностные распределения на множествах открытых текстов и множестве ключей. При этом, для любого $x \in X$ выполняется равенство: $D(K, E(K, x)) = x$.

Гомоморфные алгоритмы шифрования принято разделять на частично гомоморфные, ограниченно гомоморфные и полностью гомоморфные алгоритмы шифрования. Пусть на множестве X задана бинарная операция $*: X \times X \to X$. Алгоритм шифрования называется *частично гомоморфным*, если существует эффективный алгоритм, который для любых $x_1, x_2 \in X$ и для любого $k \in K$, получив на вход только $E(k, x_1)$ и $E(k, x_2)$, выдает значение $E(k, x_1 * x_2)$. Под *ограниченно гомоморфными* алгоритмами шифрования будем подразумевать такие Σ, что для любой функции $f: X^m \to X$ от m переменных из некоторого класса функций F, для любых $x_1, \ldots, x_m \in X$ и любого ключа $k \in K$ существует эффективный алгоритм, который, получив на вход только $\{E(k, x_i), i = 1, \ldots, m\}$, выдает значение $E(k, f(x_1, \ldots, x_m))$. *Полностью гомоморфным* алгоритмом шифрования называется ограниченно гомоморфный алгоритм в случае, когда F есть множество всех возможных функций $X^m \to X$ для всех $m \in \square$.

Вопросам обоснования стойкости гомоморфных алгоритмов шифрования посвящено большое количество работ [3-7], однако зачастую в них рассматриваются вопросы безопасности конкретных решений, обоснования свойств которых сводят к решению некоторых (не всегда хорошо изученных) задач. Рассмотрим вопрос построения универсальной модели нарушителя, позволяющий в том числе предложить универсальные методы анализа гомоморфных алгоритмов шифрования.

Рассмотрим для начала модель практической стойкости не гомоморфных алгоритмов шифрования, аналогично 0. Пусть здесь и далее $щ, \kappa$ – независимые случайные величины, принимающие соответственно значения на множестве X и K и имеющие распределения $\mathrm{P}(щ = x) = P_X(x), \forall x \in X$, $\mathrm{P}(\kappa = k) = P_K(k), \forall k \in K$. Через $\hat{\xi}$ обозначим реализацию случайной величины ξ. Тогда $\hat{\kappa}$ – реализация случайной величины κ, называемая ключом и неизвестная

нарушителю, $\text{ш}\hat{\text{ш}}_1,\ldots,\text{ш}\hat{\text{ш}}_n$ – реализация n независимых, одинаково распределенных случайных величин, имеющих распределение, равное распределению случайной величины ш. Для каждого $i \in \overline{1,n}$ определим случайную величину γ_i: $\gamma_i = E\left(\kappa, \text{ш}_i\right)$, откуда $\hat{\gamma}_i = E\left(\hat{\kappa}, \text{ш}\hat{\text{ш}}_i\right)$. Тогда для каждого $i \in \overline{1,n}$ значение $\left(\text{ш}\hat{\text{ш}}_i, \hat{\gamma}_i\right)$ суть пара открытый-шифрованный текст.

Пусть у нарушителя имеется множество шифртекстов $\hat{Y} = \left\{\gamma_i, i = 1,\ldots,m\right\}$, $m < n$, а также множество пар открытый-шифрованный текст $XY = \left\{\left(\text{ш}\hat{\text{ш}}_i, \hat{\gamma}_i\right), i = m+1,\ldots,n\right\}$, полученных при неизвестном ключе $\hat{\kappa}$. Алгоритм шифрования Σ можно считать практически ε-стойким с уровнем стойкости T если произвольный нарушитель, обладающий вычислительными возможностями не превосходящими T, для некоторого фиксированного $\varepsilon > 0$, может предъявить $i \in \overline{1,m}$, такое, что $\left|\text{P}\left(\text{ш} = \text{ш}\hat{\text{ш}}_i \middle| \hat{Y}, XY\right) - P_X\left(\text{ш}\hat{\text{ш}}_i\right)\right| \geq \varepsilon$. Заметим, что в случае, когда у противника неограниченные вычислительные возможности, он может определить ключ $\hat{\kappa}$ и расшифровать все сообщения. Тогда для каждого $i \in \overline{1,m}$, указанный модуль разности вероятностей будет равен $\left|1 - P_X\left(\text{ш}\hat{\text{ш}}_i\right)\right|$.

Заметим, что такое обоснование стойкости корректно применимо к алгоритмам шифрования, к которым не предъявляются требования гомоморфности. Для гомоморфных же алгоритмов такое определение стойкости не содержательно ввиду предъявляемым дополнительных требований. Для обоснования стойкости гомоморфных алгоритмов шифрования можно предложить следующую модель. Пусть как и ранее у нарушителя имеется множество шифртекстов $\hat{Y} = \left\{\gamma_i, i = 1,\ldots,m\right\}$, $m < n$, а также пары открытый-шифрованный текст $XY = \left\{\left(\text{ш}\hat{\text{ш}}_i, \hat{\gamma}_i\right), i = m+1,\ldots,n\right\}$, полученные на неизвестном ключе $\hat{\kappa}$. В этом случае, помимо получения информации об открытых текстах, соответствующих множеству шифртекстов $\hat{Y}$, нарушителю интересны также значения функций, которые могут быть вычислены для рассматриваемой схемы гомоморфного шифрования. Для ограниченно гомоморфных алгоритмов шифрования указанные значения сильно зависят от класса функций, относительно которых алгоритм шифрования является гомоморфным. Рассмотрим два наиболее простых случая.

Для простоты изложения предположим, что во множестве X с заданной на нем бинарной операцией $*$ существует нейтральный относительно $*$ элемент θ. При анализе частично гомоморфной схемы интерес представляет информация о произвольном значении вида: $\text{ш}\hat{\text{ш}}_1^{a_1} * \ldots * \text{ш}\hat{\text{ш}}_n^{a_n}$, где $a_i \in \{0,1\}$, $x_i^1 = x_i$, $x_i^0 = \theta$ и существует $i \in \overline{1,m}$ такой, что $a_i \neq 0$. Число таких наборов $\left(a_1,\ldots,a_n\right)$ равно $2^n - 2^{n-m+1}$. Тогда, частично гомоморфный алгоритм шифрования Σ будем считать практически ε-стойким с уровнем стойкости T, если произвольный нарушитель, обладающий вычислительными возможностями не превосходящими T, для некоторого фиксированного $\varepsilon > 0$, может предъявить $a = \left(a_1,\ldots,a_n\right) : \exists i \in \overline{1,m}, a_i = 1$, такое, что

$$\left|\text{P}\left(\text{ш} = \text{ш}\hat{\text{ш}}_1^{a_1} * \ldots * \text{ш}\hat{\text{ш}}_n^{a_n} \middle| \hat{Y}, XY\right) - P_X\left(\text{ш}\hat{\text{ш}}_1^{a_1} * \ldots * \text{ш}\hat{\text{ш}}_n^{a_n}\right)\right| \geq \varepsilon.$$

Если на множестве X заданы две операции $+$, $\cdot$, позволяющие определить на X структуру поля, и алгоритм Σ гомоморфен по двум этим операциям, то он очевидно полностью гомоморфен.

В этом случае, интерес представляет не $2^n - 2^{n-m+1}$ значений, а $|X|^{|X|^n} - |X|^{|X|^{n-m+1}}$, но практическая стойкость может быть определена аналогично ранее изложенному.

На примере частично гомоморфной схемы покажем, что существуют универсальные методы анализа гомоморфных алгоритмов шифрования в предложенной модели. Действительно, пусть $2^n - 2^{n-m+1} = \alpha \cdot \sqrt{|Y|}$, $2^{n-m+1} = \beta \cdot \sqrt{|Y|}$, $\alpha, \beta \in (0, \sqrt{|Y|})$, тогда, по парадоксу дней рождения (см. напр. 0), среди множества значений $\left\{ш\hat{u}_1^{a_1} * \ldots * ш\hat{u}_m^{a_m} : a_i \in \{0,1\}, i \in \overline{1,m}, \exists i \in \overline{1,m} : a_i = 1\right\}$ и $\left\{ш\hat{u}_{n+1}^{a_{n+1}} * \ldots * ш\hat{u}_m^{a_m} : a_i \in \{0,1\}, i \in \overline{n+1,m}\right\}$ есть пересечение с вероятностью, которая может быть оценена величиной $1 - \exp\{-\alpha \cdot \beta\}$. Иными словами, при достаточном объеме материала, нарушитель может получить информацию о значении некоторой функции от открытых текстов. Заметим, что указанная вероятность существенно зависит от вероятностного распределения, заданного на X.

Отдельно отметим, что для ряда алгоритмов шифрования функция зашифрования E является случайной функцией (см. напр. [3-7, 10]). Иными словами, можно считать, что фиксировав значение ключа и открытого текста, значение функции шифрования зависит от случайного параметра $r \in R$, имеющего распределение P_R. При наличии неравновероятности на множестве R, вероятность пересечения описанных выше множеств также будет больше, чем в равновероятном случае.

Список литературы:

1. R. L Rivest, L. Adleman, M. L Dertouzos, et al. On data banks and privacy homomorphisms. Foundations of secure computation, 4(11):169–180, 1978.

2. Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии. Учебное пособие. М.: Гелиос АРВ, 2001. 480 с.

3. C. Gentry. Fully homomorphic encryption using ideal lattices. Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing, pages 169-178, 2009.

4. K. Gjøsteen and M. Strand. Fully homomorphic encryption must be fat or ugly? https://eprint/iacr/org/2016/105, 2016.

5. M. Van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan. Fully homomorphic encryption over the integers. Annual International Conference of the Theory and Applications of Cryptographic Techniques, pages 24-43, 2010.

6. V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. EUROCRYPT 2010, pages 1-23, 2010.

7. O. Regev. On lattices, learning with errors, random linear codes, and cryptography. Journal of the ACM (JACM), 56, 2009.

8. N. Koblitz and A. Menezes, Another Look at “Provable Security”. II, Progress in Cryptology – INDOCRYPT 2006, pages 148-175, 2006.

9. B. Prenel. Cryptographic hash functions. European Transactions on Telecommunications, vol. 5, issue 4, pages 431-448, 1994.

10. T. Okamoto, S. Uchiyama. A new public-key cryptosystem as secure as factoring. In International conference on the theory and applications of cryptographic techniques, стр. 308–318. Springer, 1998.

Шкоркина Е.Н.
Санкт-Петербургский политехнический университет Петра Великого

АУТЕНТИФИКАЦИЯ В ГРАНИЧНОЙ ВЫЧИСЛИТЕЛЬНОЙ АРХИТЕКТУРЕ С ВЫРАБОТКОЙ КЛЮЧЕЙ ЗАЩИТЫ ДАННЫХ

** Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-90110.*

Важным условием функционирования таких систем, как мобильные, автомобильные или летающие самоорганизующиеся сети (MANET, VANET, FANET), Интернета вещей (ИВ) и киберфизических систем в целом, в последнее время стала обработка и анализ данных в режиме реального времени. Для обеспечения этого аспекта были изобретены и развиваются граничные вычисления, архитектура которых характеризуется расположением низкоресурсных электронных потребительских устройств ИВ в физической и логической близости от граничного сервера. Однако это влечет за собой необходимость адаптации протоколов взаимодействия и защитных механизмов, основополагающим из которых является аутентификация.

В соответствии с иерархической архитектурой граничных вычислений, в которой граничный сервер дополняет облачный, можно выделить пять моделей взаимодействия:

1. Граничный сервер с облачным сервером. Позволяет передавать предварительно обработанные данные для хранения и обработки совместно с данными от других серверов.
2. Граничный сервер с граничным сервером. Синхронизация данных, в том числе аутентификационных, об устройствах ИВ.
3. Устройство ИВ с граничным сервером. Передача данных, накопленных, например, датчиками и сенсорами, для предварительной обработки.
4. Устройство ИВ с облачным сервером через граничный. Прямая передача накопленных данных в случае, если решение о необходимости отправлять данные на облачный сервер принимает само устройство ИВ.
5. Два устройства ИВ друг с другом через граничный сервер. Передача команд управления, запросов состояния или прочих данных.

Уже известные решения аутентификации для данных моделей отличаются сетевым уровнем встраивания, принципами принятия решения о предоставлении доступа, составом криптографических алгоритмов и объемом необходимых ресурсов. Так, модели 1 и 3 не отличаются от классических парных взаимодействий. Однако известны решения [1], учитывающие низкие ресурсы одной из сторон и повышенную мобильность сетей VANET.

В модели 2 с целью повышения степени подлинности передаваемых данных по цепочке серверов могут использоваться перспективные методы аутентификации на основе технологии блокчейн [2] и групповых подписей [3, 4].

Модели 4 и 5 характерны именно для рассматриваемой архитектуры и для эффективного использования всех преимуществ граничных вычислений требуют разработки новых протоколов. Протокол аутентификации устройств в модели 4 представлен в работе [5]. Протокол аутентификации управляющего устройства (УУ) на исполняемом (ИУ) через граничный сервер в модели 5 опубликован в работе [6] и позволяет противостоять атакам несанкционированного доступа к ИУ, подмены граничного сервера и истощения энергоресурсов ИУ. Предлагаемый протокол является его модернизированной версией, позволяющей дополнительно:

1. Выработать в ходе аутентификации ключи для последующей передачи команд управления и ответов на них путем формирования граничным сервером значений вида $Sign \oplus r \| r$ и $Sign \oplus q \| q$ ($Sign$ – подпись за уникальные идентификационные данные граничного сервера, r и q – случайно сгенерированные числа длины подписи). Первое из них передается УУ в виде шифрограммы, а второе – ИУ в виде шифрограммы с имитовставкой. Значения $K_1 = Sign \oplus r$ и $K_2 = Sign \oplus q$ сохраняются устройствами и сервером после успешной аутентификации в качестве ключей передачи данных с помощью алгоритма аутентифицированного шифрования (K_1 – ключ защищенной передачи данных между УУ и сервером, а K_2 – между сервером и ИУ). При выборе алгоритма, длина ключа которого равна n бит и меньше длины $Sign$, могут использоваться, например, n старших бит.

2. Выполнять упрощённую аутентификацию при отправке команд управления ИУ через другой граничный сервер. Пусть протокол был выполнен через граничный сервер ES_1, а УУ находится в зоне действия ES_2. Тогда в результате первой фазы протокола будет получен ключ $K_{ES_2} = Sign_{ES_2} \oplus r' \| r'$, с помощью которого управляющая команда может быть отправлена в защищенном виде серверу ES_2. Тот в свою очередь должен отправить ее ES_1 с последующей отправкой ИУ на ранее выработанном ключе K_2.

3. Нивелировать атаки по известному открытому тексту благодаря передаче каждый раз разного открытого текста (выполняется маскирование значения $Sign$ случайным числом), первая половина которого всегда является ключом аутентифицированного шифрования, и, как следствие, разного шифртекста.

Гибкость протокола по-прежнему обеспечивается выбором любого подходящего набора криптографических алгоритмов для конкретных вычислительных возможностей устройств, взаимодействующих в рамках протокола, и пропускной способности граничного сервера.

Список использованных источников:

1. Yang A. et al. Delegating authentication to edge: A decentralized authentication architecture for vehicular networks //IEEE Transactions on Intelligent Transportation Systems. – 2020.

2. Zhaofeng M. et al. Blockchain-based decentralized authentication modeling scheme in edge and IoT environment //IEEE Internet of Things Journal. – 2020. – V. 8. – №. 4. – P. 2116–2123.

3. Aleksandrova E. B. Methods of group authentication for low-resource vehicle and flying self-organizing networks //Automatic Control and Computer Sciences. – 2017. – V. 51. – №. 8. – P. 947–958.

4. Aleksandrova E. B., Shtyrkina A. A., Yarmak A. V. Post-quantum group-oriented authentication in IoT //Nonlinear Phenomena in Complex Systems. – 2020. – V. 23. – №. 4. – P. 405–413.

5. Shahidinejad A. et al. Light-edge: A lightweight authentication protocol for IoT de-vices in an edge-cloud environment //IEEE Consumer Electronics Magazine. – 2021.

6. Aleksandrova E. B., Oblogina A. Y., Shkorkina E. N. Authentication of Control Devices in the Internet of Things with the Architecture of Edge Computing //Automatic Control and Computer Sciences. – 2021. – V. 55. – №. 8. – P. 1087–1091.

Содержание

МЕТОДЫ И ТЕХНИЧЕСКИЕ СРЕДСТВА ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ИНФОРМАЦИИ
СБОРНИК МАТЕРИАЛОВ
31-ой научно-технической конференции
Памяти Петра Дмитриевича Зегжды

27 – 30 июня 2022 года Санкт-Петербург

Лицензия ЛР№020593 от 07.08.97 г.
Налоговая льгота – Общероссийский классификатор продукции
ОК 005-93, т.2; 95 3004 – научная и производственная литература.

Подписано в печать 20.06.22 г. Формат бумаги 60х84/8. Усл.печ. л. 14,5.
Тираж 110 экз. Заказ №

Отпечатано с готового оригинал-макета, предоставленного
оргкомитетом конференции,
в Цифровом типографском центре
Политехнического университета Петра Великого.
195251, Санкт-Петербург, Политехническая, 29