ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

Ջամղարյան Թիմուր Վյաչեսլավի

ՏԵՂԵԿԱՏՎԱԿԱՆ ՑԱՆՑՈՒՄ ՏՎՅԱԼՆԵՐԻ ԱՄԲՈՂՋԱԿԱՆՈՒԹՅԱՆ ԱՊԱՀՈՎՄԱՆ ՀԱՄԱԿԱՐԳԻ ՄՇԱԿՈՒՄԸ

Ե.13.04 – «Հաշվողական մեքենաների, համալիրների, համակարգերի և ցանցերի մաթեմատիկական և ծրագրային ապահովում» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի հայցման ատենախոսության

Ս Ե Ղ Մ Ա Գ Ի Ր

Երևան – 2023

ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ НАН РА

Джамгарян Тимур Вячеславович

РАЗРАБОТКА СИСТЕМЫ ОБЕСПЕЧЕНИЯ ЦЕЛОСТНОСТИ ДАННЫХ В ИНФОРМАЦИОННОЙ СЕТИ

А В Т О Р Е Ф Е Р А Т

диссертации на соискание учёной степени кандидата технических наук по специальности 05.13.04 «Математическое и программное обеспечение вычислительных машин, комплексов, систем и сетей»

Ереван – 2023

Ատենախոսության թեման հաստատվել է Հայաստանի ազգային պոլիտեխնիկական համալսարանում:

Գիտական ղեկավար՝	տեխ. գիտ. թեկնածու	Ռոբերտ Գրիգորի Հակոբյան
Պաշտոնական ընդդիմախոսներ՝	տեխ. գիտ. դոկտոր	Հրաչյա Վոլոդյայի Ասցատրյան
	ֆիզ. մաթ. գիտ. թեկնածու	Աշոտ Նշանի Հարությունյան
Առաջատար կազմակերպություն՝	Երևանի կապի միջոցների գիտահետազոտական ինստիտուտ	

Պաշտպանությունը կայանալու է 2024թ. հունվարի 30-ին, ժամը 15:00-ին ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում գործող 037 «Ինֆորմատիկա» մասնագիտական խորհրդի նիստում, հետևյալ հասցեով՝ Երևան, 0014, Պ. Սևակի 1:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ ԻԱՊԻ գրադարանում:

Սեղմագիրը առաքված է 2023թ. դեկտեմբերի 28-ին:

037 Մասնագիտական խորհրդի
գիտական քարտուղար, ֆ.մ.գ.դ. Մ. Ե. Հարությունյան

Тема диссертации утверждена в Национальном политехническом университете Армении.

Научный руководитель:	кандидат тех. наук	Акопян Роберт Григорьевич
Официальные оппоненты:	доктор тех. наук	Асцатрян Грачья Володиевич
	кандидат физ.-мат. наук	Арутюнян Ашот Ншанович
Ведущая организация:	Ереванский научно-исследовательский институт средств связи	

Защита состоится 30 января 2024г. на заседании специализированного совета 037 «Информатика» Института проблем информатики и автоматизации НАН РА по адресу: 0014, г. Ереван, ул. П. Севака 1.

С диссертацией можно ознакомиться в библиотеке ИПИА НАН РА.

Автореферат разослан 28 декабря 2023г.

Учёный секретарь,
Специализированного совета 037
доктор физ.-мат. наук М. Е. Арутюнян

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Активное развитие технологии машинного обучения (Machine learning, ML) и конвергенция физических сетей с виртуальными привели к масштабному увеличению атак на сетевую инфраструктуру (СИ) способных нарушить целостность циркулирующих в СИ данных. Необнаруженное вторжение в СИ способно нарушить целостность циркулирующих в ней данных. Архитекторы инфраструктурной системы защиты (ИСЗ) выстраивают безопасность системы на основе заданных компонент: правил и политик безопасности, количество которых конечно. Одной из важных задач, стоящей перед архитекторами ИСЗ, является обеспечение целостности циркулирующих в СИ данных. Различные механизмы и методы, обеспечивающие целостность данных в СИ, становятся неактуальными, при применении злоумышленниками методов, основанных на машинном обучении. При этом необходимо отметить, что все компоненты ИСЗ, которые разрабатывались десятилетиями, наряду с развитием вычислительной техники, на сегодня начинают отставать от атак в тех случаях, где генератором атаки является система с использованием технологий M2M&ML (Machine-to-Machine, M2M). Злоумышленники, применяя технологию машинного обучения способны модифицировать циркулирующие в СИ потоки данных, нарушив их целостность или полностью подменив их. Многочисленные уязвимости программного и аппаратного обеспечения, различные программные среды, сам сетевой трафик - все подлежит анализу с целью выявления точки внедрения в атакуемую СИ, эскалации и дальнейшей модификации (нарушения целостности) сетевого трафика. Системы, построенные на основе машинного обучения, способны анализировать исходный код используемого ПО и с высокой скоростью атаковать его на основе базы известных уязвимостей. В наихудшем варианте, данные системы могут сами генерировать векторы угроз. В связи с этим возникает парадоксальная ситуация, при которой иногда лучше использовать проприетарные программные среды, нежели среды с открытым исходным кодом, так как системы с машинным обучением будут не в состоянии проанализировать на уязвимости «закрытый» код проприетарных продуктов. То есть нарушение целостности проприетарных систем будет обнаружена быстрее, чем систем с открытым исходным кодом. Извлечение известных и/или конструирование новых признаков из исследуемого ПО способны создать дополнительный вектор атаки.

Системы обнаружения вторжений (СОВ), в связке с системой предотвращения вторжений и различными файрволлами, являются элементами защиты при построении многоэшелонированной ИСЗ. Обнаружение ими различных атак

производится в рамках описанных наборов правил. Также общим недостатком является сложность динамической переконфигурации системы защиты в режиме реального времени при изменении и/или модификации границ зон безопасности СИ. СОВ должны обнаруживать не только вторжения в СИ, но и определять попытку изменения/подмены циркулирующего сетевого трафика. Борьба с атаками, при которых злоумышленники используют машинное обучение, показывает не очень высокую эффективность детерминированных средств защиты. В результате чего исследователи сетевой и системной безопасности начали активно исследовать применение машинного обучения в СОВ. Но большинство разработанных и применяемых СОВ с машинным обучением работают на уровне доступа и/или уровне распределения сетевой инфраструктуры. Уровень ядра сети, как правило, не бывает защищён, так как операторы связи, которые формируют уровень ядра сети, не всегда устанавливают СОВ, отвечая только за доставку больших объёмов трафика. Атаки на целостность циркулирующих в СИ данных, как правило начинаются с анализа состояния как СИ, так и самого сетевого трафика. Возникает необходимость нейтрализации угроз целостности данных в СИ и обнаружения атаки, на ранней стадии, до ее активной эскалации. Учитывая этот факт, актуальной становиться разработка СОВ с машинным обучением, способной обнаружить нарушение целостности передаваемых на уровне ядра сети данных.

Цель работы заключается в разработке системы обнаружения вторжений с машинным обучением, работающей на уровне ядра сети и способной обнаруживать нарушение целостности передаваемых данных.

Достижение данной цели предполагает решение следующих задач:

1. Решение исследовательской задачи обнаружения полиморфного вредоносного ПО с использованием неполных наборов данных для нейтрализации атак на целостность.
2. Решение задачи обнаружения обфусцированного, скомпилированного вредоносного ПО, при применении злоумышленником инструментов на основе машинного обучения.
3. Решение задачи выявления нарушений целостности циркулирующих в СИ данных с использованием нейронных сетей, функционирующих в режиме нехватки наборов данных.
4. Разработка многокритериальной, стратифицированной модели СИ, позволяющей определить количество необходимых критериев, изменяя которые СИ будет выведена из состояния устойчивости и нарушится целостность циркулирующих данных при применении злоумышленником в качестве инструмента атаки методов машинного обучения.

Объектом исследования являются вероятные атаки на целостность циркулирующих в СИ данных, механизмы их обнаружения и системы обнаружения вторжений с применением нейронных сетей.

Предметом исследования являются модели, алгоритмы генерации наборов данных вредоносного ПО, методики обнаружения нарушения целостности циркулирующих в СИ данных, различные нейронные сети.

Методы исследования. В диссертационной работе для обеспечения целостности циркулирующих в СИ данных, использованы методы математического моделирования на основе метода оценки устойчивости Ляпунова, статистических критериев Пирсона, Манна-Уитни, Фишера, метода вычисления редакционного расстояния, метода Фадеева-Леверье, метода k ближайших соседей, способ минимизации функций методом неопределённых коэффициентов.

Научная новизна диссертационной работы заключается в следующем:

- Предложен метод обнаружения вредоносного полиморфного ПО, улучшающий по заданным параметрам, результаты существующих методов, что позволяет более эффективно выявлять нарушения целостности циркулирующих в СИ данных.
- Предложен метод обнаружения обфусцированного, скомпилированного вредоносного ПО, превосходящий результаты существующих при идентичных входных данных, что создаёт возможность обнаруживать данное ПО.
- Предложен метод обнаружения вторжений при атаке на целостность (при прочих равных условиях) как против циркулирующих в СИ данных, так и самой СОВ, обеспечивающий обнаружение вторжения быстрее, чем существующие методы.
- Усовершенствована «модель угроз» в рамках повышения устойчивости СИ и целостности, циркулирующих в СИ данных при ее многокритериальной, стратифицированной оптимизации в режиме динамической переконфигурации, при деструктивном воздействии на неё.

Практическая значимость работы.

- Разработана система обнаружения вторжений на основе машинного обучения, которая, более быстро обнаруживает нарушения целостности передаваемого сетевого трафика с меньшим количеством входных данных.

Практическое сравнение по заданным параметрам проводилось с программными Snort, Suricata, OSSEC, Zeek, и программно-аппаратными NGFW (Next-Generation Firewall) и UTM (Unified Thread Management) СОВ на основе файлов сетевого трафика, с внедрённым вредоносным ПО. Разработанная система работает в комплексе с системой SIEM (security information and event management), а также с системой локального и удалённого журналирования событий.

Степень достоверности результатов. Достоверность изложенных в диссертационной работе результатов обеспечивается анализом исследований в области обеспечения безопасности СИ от атак и угроз с применением злоумышленниками методов машинного обучения, подтверждением научных расчётов практической реализацией, сравнением при равных условиях с существующими аналогичными разработками, внедрением разработанного ПО в государственных структурах, а также публикацией в рецензируемых изданиях в том числе международного уровня.

Положения, выносимые на защиту:

- Метод обнаружения вредоносного полиморфного ПО при атаке на целостность циркулирующих в СИ данных.
- Метод обнаружения обфусцированного, скомпилированного вредоносного ПО способного нарушить целостность циркулирующих в СИ данных.
- Метод обнаружения вторжений с меньшим временем срабатывания СОВ против атаки на целостность как циркулирующих в СИ данных, так и самой СОВ.
- Усовершенствованная «модель угроз» с допущением, о применении злоумышленником инструментов на основе машинного обучения.

Внедрение. Разработанная система обнаружения вторжений с машинным обучением внедрена в систему связи Вооружённых сил Республики Армения. Внедрённая система обнаружения вторжений обеспечивает выявление нарушения целостности циркулирующих в сетевой инфраструктуре данных, нейтрализуя большинство атак на неё.

Применение разработанной системы повышает безопасность и устойчивость сетевой инфраструктуры Вооружённых сил Республики Армения.

Апробация результатов работы. Основные результаты диссертационного исследования были представлены на ряде конференций, в том числе:

- II международная научно-теоретическая конференция «CYBER-QORGAY 2021», «Состояние, перспективы и тенденции развития сферы информационной (кибер) безопасности», 31.03.2021, Республика Казахстан.
- XIII международная конференция по компьютерным наукам и информационным технологиям (CSIT-2021), 27.09.2021 - 01.10.2021, Ереван, Армения.
- Международная научно-теоретическая конференция, «Силовые структуры как инструмент обеспечения национальной безопасности. Опыт проведения совместных операций», Алматы, Казахстан, 21.12.2022.
- XIV международная конференция по компьютерным наукам и информационным технологиям (CSIT-2023), 25.09.2023 -30.09.2023, Ереван, Армения.
- На научных семинарах Института проблем информатики и автоматизации НАН РА.
- На научных семинарах кафедры.

Публикации. Основные результаты диссертационной работы представлены в 14 печатных изданиях, 12-из которых опубликованы в журналах, рекомендованных ВАК, один в зарубежном издании, индексированным в Scopus, 2 в прочих изданиях. Список публикаций представлен в конце автореферата.

Структура и объём диссертационной работы. Диссертация состоит из введения, четырёх глав, заключения и 6 приложений. Основной материал изложен на 117 страницах, полный объём диссертации составляет 138 страниц с 74 рисунками. Список литературы содержит 98 наименований.

СОДЕРЖАНИЕ РАБОТЫ

Введение. Во введении обоснована актуальность диссертационного исследования, указана цель работы, задачи, подлежащие решению, для достижения поставленной цели. Указаны объект и предмет исследования, научная и практическая значимость работы, основные результаты диссертационного исследования.

В главе 1 проведён анализ существующих архитектур построения сетей и систем обнаружения вторжений. Рассмотрена трёхуровневая иерархическая модель построения СИ и вероятные атаки на СИ. Определены основные недостатки «детерминированных» СОВ при атаке с применением машинного обучения. Предметно рассмотрены существующие СОВ с открытым исходным кодом, а также исследования, посвящённые СОВ с машинным обучением. Определены и сформулированы основные задачи, стоящие перед исследователями СОВ с машинным обучением.

Рассмотрено требование устойчивого функционирования СИ и целостности циркулирующих данных. Обосновано, что для обеспечения целостности циркулирующих в СИ данных при использовании злоумышленником методов машинного обучения, также необходимо применение инструментов и методов, основанных на машинном обучении. Сформулирована задача расчёта устойчивости СИ и целостности циркулирующих данных, при наличии ее многокритериальной, стратифицированной модели.

Предметно проведён обзор некоторых СОВ с открытым исходным кодом. Анализ исходного кода проприетарных СОВ на основе NGFW и UTM не производился ввиду его закрытости. Изучены исследования посвящённые существующим СОВ с машинным обучением. Определены типы нейронных сетей и статистические критерии, применяемые для обнаружения атаки на целостность циркулирующих данных. Выбор нейронных сетей осуществлялся на основе принципа взаимного дополнения и решения формулы функционирования нейросетевой классификационной модели (1).

где,
$$Y(z)=\varphi\left(\sum_{i=1}^{N_2}\omega_{i1}^{(3)}.\varphi\left(\sum_{j=1}^{N_1}\omega_{ij}^{(2)}.\varphi\left(\sum_{k=1}^{n}\omega_{jk}^{(1)}.Z_k+\Theta_j^{(1)}\right)+\Theta_i^{(2)}\right)+\Theta_1^{(3)}\right) \quad (1)$$

φ - функция активации нейронной сети,

$\omega_{ij}^{(k)}$ - «веса» на входе нейрона i-го слоя,

$\Theta_i^{(n)}$ -параметр смещения выходного слоя,

Z_k -сигналы на входе/выходе слоёв нейронной сети.

Аналитическое выражение (1) также использовалось для конечной оценки суммарной эффективности всей СОВ с машинным обучением при борьбе с атаками на целостность циркулирующих данных, включающей в свой состав несколько нейронных сетей. На основе аналитического подхода и экспериментов для использования в СОВ с машинным обучением с целью обеспечения целостности циркулирующих в СИ данных из всего множества нейронных сетей, отобраны:

- Капсульная нейронная сеть (CapsNet).
- Рекуррентная нейронная сеть (Recurrent Neural Network, RNN).
- Генеративно-состязательная сеть (Generative Adversarial Network, GAN).
- Свёрточная нейронная сеть (Convolutional neural network, CNN).

Сформулирована постановка задачи диссертационного исследования. Указаны задачи, подлежащие решению, для обеспечения целостности передаваемых данных в информационной сети.

В главе 2 рассмотрена разработанная математическая модель и критерии оценки устойчивости СИ и целостности, циркулирующих в ней данных, при ее динамической переконфигурации. Разработана модель атаки на целостность циркулирующих в СИ данных при применении злоумышленником нейронных сетей. Модель атаки на целостность циркулирующих в СИ данных разработана на предположении, что злоумышленник обладает возможностями, описанными в модели угроз Долева-Яо. Разработан алгоритм и ПО генерации и подготовки наборов данных вредоносного ПО для обучения используемых нейронных сетей обнаружению атак на целостность циркулирующих данных. Рассмотрена математическая модель многокритериальной СИ. В качестве метода определения устойчивости СИ и целостности циркулирующих данных использовался первый метод Ляпунова, для оценки оптимальности динамической переконфигурации – принцип Парето.

Проведена генерация наборов данных вредоносного ПО для обучения нейронных сетей обнаружению нарушению целостности передаваемых данных, на основе разработанного алгоритма. Для создания «синтетических» наборов данных применялся метод аугментации. В качестве инструмента для увеличения наборов данных использовался метод бустинга.

Обработка, преобразование и расширение базы входных наборов данных реализована на основе механизма *«рекуррентной сети с вниманием»* (рис. 1).

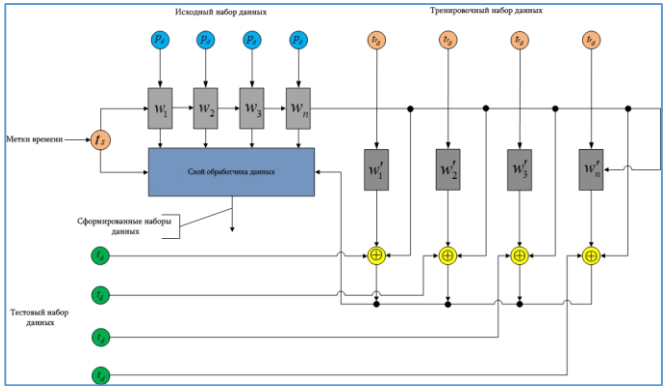

Рис. 1. Механизм расширения базы наборов данных на основе «рекуррентной сети с вниманием»

В рамках данного исследования с применением вредоносного ПО *athena, abc, cheeba, dyre, december_3, engrat, surtr, stasi, otario, dm, v-sign, tequila, flip*, *grum, mimikatz,* суммарно сформировано 6х55.000 Мбайт коллекций наборов данных различной размерности (20, 40, 128, 256, 512, 1024 байт). Представлен разработанный алгоритм и ПО подготовки наборов данных («фильтрации от шумов») для обучения нейронных сетей обнаружению нарушения целостности передаваемых данных в СИ. В качестве инструмента определения статистических различий между выборками применялся статистический критерий Манна-Уитни. Визуализированные результаты испытаний представлены на рис. 2, 3.

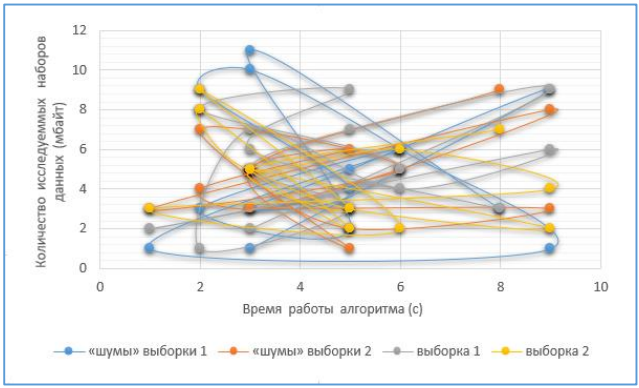

Рис. 2. Выходной набор данных без обработки разработанным ПО

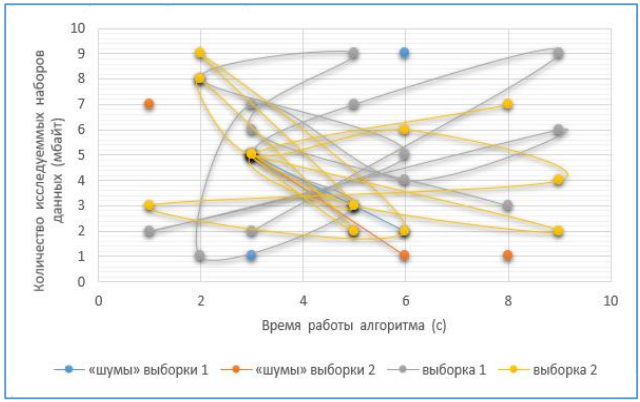

Рис. 3. Выходной набор данных после обработки разработанным ПО при трёх итерациях

Разработанный алгоритм и ПО позволяет при наличии сегментов исходного кода (не менее (5÷5,5) %) различного вредоносного ПО создавать обучающие наборы данных для нейронных сетей, входящих в СОВ с целью нейтрализации атак на целостность передаваемых в СИ данных.

В главе 3 проведена модификация многомерной логистической функции (*softmax)* с целью повышения точности обнаружения нарушения целостности передаваемых данных. Разработана и исследована программная модель обнаружения нарушения целостности передаваемых данных и выявления вредоносного ПО с помощью капсульной нейронной сети с применением трансферного обучения и кусочно-контекстного нечёткого хеширования. Использованы «матрицы состояний» для определения вредоносного ПО. Проверка предлагаемого решения осуществлялась на наборах данных, сформированных на основе полиморфного вредоносного ПО *abc, cheeba, december_3, stasi, otario, dm, v-sign, tequila, flip.* Наборы данных формировались на основе вредоносного ПО *athena*, *dyre, engrat, grum, mimikatz, surtr.* На рис. 4 представлен отчёт сервиса *virustotal* при исследовании одного из обфусцированных образцов вредоносного ПО *mimikatz* детектированного, нейронными сетями из состава разработанной СОВ.

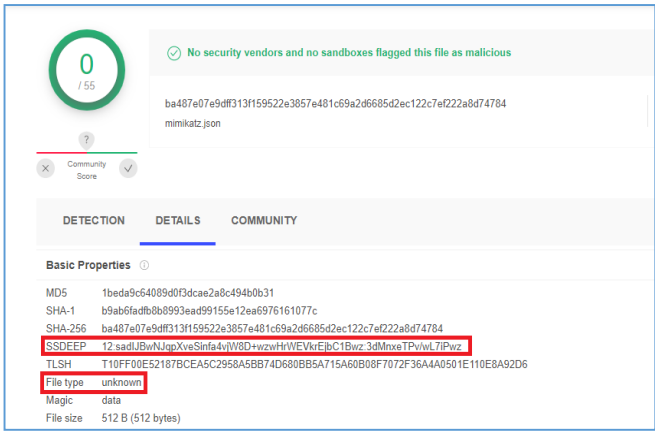

Рис. 4. Отчет сервиса *virustotal*

Как видно из рис. 4 сервис *virustotal, не* обнаружил ни тип файла, ни принадлежность его к тому или иному виду вредоносного ПО в отличии от разработанной СОВ. При наличии не менее чем (3,13÷3,27) % фрагмента исходного кода вредоносного ПО (для ПО *ssdeep* данный параметр равен (5,4÷6,2) %) капсульная нейронная сеть показывает результаты лучше, чем свёрточная нейронная сеть на (5,6÷6,4) % и (21,3÷22,4) % чем ПО *ssdeep* при обнаружении обфусцированного вредоносного ПО. В качестве математического аппарата для генерации «весовых» коэффициентов при инициализации нейронной сети на основе «матрицы состояний» использовался метод Фадеева-Леверье. Исследование проводилось с матрицами размерностью 16x16, 32x32, 64x64, 128x128, 256x256, 512x512, 1024x1024, 2048x2048, 4096x4096 элементов. Результаты исследования по заданным параметрам сравнивались с различными программными и программно-аппаратными СОВ на основе файлов сетевого трафика с внедрённым вредоносным ПО.

Сравнение осуществлялось по следующим заданным параметрам:

- обнаружение обфусцированного вредоносного ПО в режиме активной модификации циркулирующих данных,
- обнаружение обфусцированного скомпилированного вредоносного ПО,
- количество минимально необходимых наборов данных для обнаружения полиморфного вредоносного ПО,
- точность обнаружения вредоносного ПО на 500Мб сетевого трафика при скорости передачи данных до 1Гбит/с,

- скорость реакции системы обнаружения вторжений с машинным обучением на атаку целостности данных при скорости передачи 1Гбит/с.

Применение вышеуказанных методов позволяет улучшить на 19,3% результаты, основанные на существующих методах, при обнаружении обфусцированного и/или скомпилированного вредоносного ПО, нарушающего целостность циркулирующих в СИ данных.

В главе 4 разработан алгоритм и ПО повышения достоверности СОВ с машинным обучением. Проведена оценка эффективности нейронных сетей из состава СОВ с машинным обучением. Предложено решение задачи повышения достоверности СОВ с машинным обучением на основе рекуррентной нейронной сети. В качестве математического аппарата выбран метод *k ближайших соседей* (k Nearest Neighbors, kNN). Проведена интеграция рекуррентной нейронной сети с СОВ Snort. Проведена оценка эффективности нейронных сетей входящих в состав СОВ с машинным обучением. В разработанной СОВ в качестве одного из механизмов управления выбрана генеративная модель, так как в отличии от детерминированной модели она представляет не просто фиксированные вычисления, выдающие при неизменных входных данных соответствующий им неизменно тот же набор выходных данных, а содержит стохастический элемент, позволяющий генерировать новые правила на основе известных сценариев атак. В качестве инструмента анализа работы генеративной модели использовался критерий Фишера. На рис. 5, 6 приведены результаты работы, разработанной СОВ с машинным обучением и «детерминированной» СОВ.

Рис. 5. Количество инцидентов, зафиксированных «детерминированной» системой обнаружения вторжений в реальной сети.

Рис. 6. Количество инцидентов, зафиксированных системой обнаружения вторжений с машинным обучением в реальной сети.

Как видно из графиков на рис. 8 и 9 количество инцидентов в интервалах времени 01^{00}–02^{00}, 07^{00}–08^{00}, 11^{00}–13^{00}, 19^{00}–20^{00} «детерминированная» СОВ аномалий не обнаруживала, а разработанная СОВ с применением машинного обучения фиксировала (в том числе и атак на целостность передаваемых данных). Сетевой трафик был проверен сторонними средствами объективного контроля и анализа которые подтвердили факт атаки (в том числе и атак на целостность циркулирующих данных) и её обнаружение разработанной СОВ с машинным обучением.

Разработан алгоритм и ПО, осуществляющее расчёт производительности. В качестве математического аппарата уменьшения размерности вычисляемых метрик производительности выбран *способ* минимизации функций *методом неопределённых коэффициентов*. Показано, что возможно проведение теста производительности меньшим объёмом выборки, без уменьшения ее репрезентативности. В качестве измеряемых метрик, из всего пространства производительности выбраны:

- метрика 1- количество запросов в секунду (Request Per Second, RPS),
- метрика 2- временной интервал между началом и окончанием срабатывания СОВ на вредоносное ПО (Latency),
- метрика 3 - пропускная способность капсульной, сверточной и генеративно-состязательных сетей (Throughput),
- метрика 4 - утилизация процессора и оперативного запоминающего устройства.

Применение метода позволило снизить количество измеряемых/обрабатываемых наборов данных в 2,13 раз, не снижая их репрезентативность. Полученные результаты позволяют с вероятностью 0,78–0,81 заранее рассчитать требуемый аппаратный ресурс для развёртывания СОВ с машинным обучением для заданных метрик и селективно конфигурировать разработанную СОВ под требуемые задачи.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИОННОЙ РАБОТЫ

- Предложен метод обнаружения вредоносного полиморфного ПО, улучшающий по заданным параметрам, результаты существующих методов, что позволяет более эффективно выявлять нарушения целостности циркулирующих в СИ данных [8,10,13,14].
- Предложен метод обнаружения обфусцированного, скомпилированного вредоносного ПО, превосходящий результаты существующих при идентичных входных данных, что создаёт возможность обнаруживать данное ПО [6,7,10,14].
- Предложен метод обнаружения вторжений при атаке на целостность (при прочих равных условиях) циркулирующих в СИ данных и самой СОВ, обеспечивающий обнаружение вторжения быстрее, чем существующие методы [1,5,7,10,11].
- Усовершенствована «модель угроз» в рамках повышения устойчивости СИ и целостности, циркулирующих в СИ данных при ее многокритериальной, стратифицированной оптимизации в режиме динамической переконфигурации, при деструктивном воздействии на неё [1,2,3,4,7,8].
- Разработана система обнаружения вторжений на основе машинного обучения, которая, более быстро обнаруживает нарушения целостности передаваемого сетевого трафика с меньшим количеством входных данных [1-4,8,9,12,13,14].

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

[1] T. Jamgharyan, V.Ispiryan, «Model of Generative Network Attack», CSIT Conference 2021, Yerevan, Armenia, September 27 - October 1, pp. 90-94. https://csit.am/2021/proceedings/IS/IS_3.pdf

[2] Т. Джамгарян, Т. Шахназарян, «Обеспечение целостности сети передачи данных, при атаке на доступность серверов аутентификации», сборник трудов международной конференции «Состояние, перспективы, и тенденции развития сферы информационной (кибер) безопасности» «CYBER-QORGAY-2021», Казахстан, Алматы 31.03.2021, стр. 76-82.

[3] T. Jamgharyan, «Application of a generative-adversarial network for managing a precise stochastically changeable signal sources», Bulletin of High Technology, 2(16)/2021, pp. 49-58, Stepanakert. http://bulletin.am/wp-content/uploads/2021/11/7.pdf

[4] Թ.Ջամղարյան, «Ներխուժումների հայտնաբերման համակարգի արդիականացում գեներատիվ մոդելի կիրառմամբ», Հայկական բանակ, 2(108). 2021, էջ.69-79. https://razmavaraget.files.wordpress.com/2022/01/hb2-final.pdf

[5] T. Jamgharyan, «Research of the Data Preparation Algorithm for Training Generative-Adversarial Network», Bulletin of High Technology, 1(19)/2022, pp. 40-50, Stepanakert. http://bulletin.am/wp-content/uploads/2022/05/5..pdf

[6] T. Jamgharyan, «Research of Obfuscated Malware with a Capsule Neural Network», Mathematical Problems of Computer Science 58, pp. 67–83, 2022. https://doi.org/10.51408/1963-0094

[7] Т. Джамгарян, А. Саргсян, «Применение машинного обучения для обеспечения безопасности сетевой инфраструктуры», сборник трудов международной научно-теоретической конференции «Силовые структуры как инструмент национальной безопасности. Опыт проведения совместных операций» Алматы, Казахстан, 21.12.2022, стр. 5-8.

[8] T. Jamgharyan, V. Ispiryan. «Network infrastructures assessment stability». *AIP Conference Proceedings*, 22 May 2023, 2757 (1): 090001. https://doi.org/10.1063/5.0136237 (indexed in Scopus).

[9] R. Hakobyan, T. Jamgharyan, «Research of Algorithm for Expanding the Database of Training Datasets of a Generative-Adversarial Network», Bulletin of High Technology 1(25)/2023, pp. 59-66, Stepanakert. https://doi.org/10.56243/18294898-2023.1-59

[10] T. Jamgharyan, «Research the Multidimensional Logistic Function in the Intrusion Detection System with Machine Learning», Bulletin of High Technology 2(26)/2023, pp. 64-70, Stepanakert. https://doi.org/10.56243/18294898-2023.2-64
[11] T. Jamgharyan, «Research of Model Increasing Reliability Intrusion Detection Systems», Mathematical Problems of Computer Science», (2023). 59, pp. 69–81. https://doi.org/10.51408/1963-0103
[12] Թ.Ջամղարյան, «Կիբեռոլորտում ներխուժումների հայտնաբերման համակարգերում մեքենայական ուսուցման կիրառման որոշ խնդիրների վերլուծություն», Հայկական բանակ, 1(115). 2023, էջ.55-67. https://razmavaraget.files.wordpress.com/2023/08/hb-2023-n1.pdf
[13] R. Hakobyan, T. Jamgharyan, «Polymorphic Malware Analysis Model», CSIT Conference 2023, Yerevan, Armenia, September 25-30, pp. 85-88. https://doi.org/10.51408/csit2023_17
[14] T. Jamgharyan, «Network Intrusion Detection System Performance Measurement Model», CSIT Conference 2023, Yerevan, Armenia, September 25-30, pp. 251-254. https://doi.org/10.51408/csit2023_59

Timur V. Jamgharyan

DEVELOPMENT OF A SYSTEM FOR ENSURING DATA INTEGRITY IN THE INFORMATION NETWORK

ABSTRACT

The relevance of the work. The active development of machine learning (ML) technology and the convergence of physical networks with virtual ones have led to a large-scale increase in attacks on the network infrastructure (NI) capable of violating the integrity of data transmitted to the NI. Architects of an infrastructure security system (ISS) build system security based on specified components: rules and security policies, the number of which is finite. One of the important tasks facing satellite architects is to ensure the integrity of transmitted data in the information network. Various mechanisms and methods that ensure the integrity of data in an information network become irrelevant when attackers ML-based methods. It should be noted that all components of artificial intelligence systems, which have been developed over decades, along with the development of computer technology, today begin to lag behind attacks in cases where the attack generator is a system using M2M&ML (Machine-to-Machine, M2M) technologies. Attackers using ML technology are able to modify data streams transmitted to the NI, violating their integrity or completely replacing them. Attacks on the integrity of transmitted data in the NI, as a rule, begin with an analysis of the state of both the NI and the network traffic itself. In order to

neutralize threats to data integrity in information systems, there is a need to detect an attack at an early stage, before its active escalation. Considering this fact, the development of intrusion detection system (IDS) with ML, capable of detecting violations of the integrity of transmitted data, at the network core level, becomes relevant.

The purpose of the work is developing an intrusion detection system with machine learning, operating at the network core level and capable of detecting violations of the integrity of transmitted data.

The object of the research is possible and probable attacks, attacks on the integrity of data transmitted in the network infrastructure, mechanisms for their detection and intrusion detection systems using artificial neural networks.

The subject of the research are models, algorithms for generating malware datasets, methods for detecting violations of the integrity of data transmitted in the network infrastructure, artificial neural networks.

Scientific novelty

- A method for detecting polymorphic malware is proposed, which improves, according to given parameters, the results of existing methods, which makes it possible to more effectively detect violations of the integrity of data circulating in the NI.
- A method for detecting obfuscated, compiled malware is proposed, which is superior to existing results with identical input data, which makes it possible to detect this software.
- An intrusion detection method is proposed for an attack on the integrity (all other things being equal) of data circulating in the NI and the IDS itself, which provides intrusion detection faster than existing methods.
- The «threat model» has been improved as part of increasing the stability of the NI and the integrity of the data circulating in the NI during its multicriteria, stratified optimization in the dynamic reconfiguration mode, with a destructive impact on it.

Structure and volume of dissertation work. The dissertation consists of an introduction, four chapters, a conclusion and 6 applications.

Introduction. The introduction substantiates the relevance of the dissertation research, indicates the purpose of the work, the tasks to be solved in order to achieve the goal.

Chapter 1 analyzes existing architectures for building networks and IDS. A three-level hierarchical model of NI construction and possible attacks on NI levels are considered. The main shortcomings of «deterministic» IDS in an attack using ML are identified. It is substantiated that in order to ensure the integrity of the data

transmitted to the NI when an attacker uses ML methods, it is also necessary to use tools and methods based on ML.

Chapter 2 considers the developed mathematical model and criteria for assessing the stability of a multicriteria network infrastructure during its dynamic reconfiguration. A model has been developed for attacking the integrity of data transmitted to the NI when an attacker uses artificial neural networks. The model of an attack on the integrity of data transmitted to the NI is developed on the assumption that the attacker has the capabilities described in the Dolev-Yao threat model.

Chapter 3 a modification of the multidimensional logistic function (*softmax*) was carried out in order to increase the accuracy of detecting violations of the integrity of transmitted data. A software model for detecting violations of the integrity of transmitted data and identifying malware using a capsule neural network using transfer learning and piecewise context hashing has been developed and studied. «state matrices» were used to identify malware.

Chapter 4 An algorithm and software for increasing the reliability of IDS with ML has been developed. It has been determined that attackers will primarily try to violate the integrity of the ML IDS itself, forcing it to produce incorrect results.

The effectiveness of neural networks from IDS with ML was assessed.

Conclusion presents short description of performed work.

Թիմուր Վյաչեսլավի Ջամղարյան

ՏԵՂԵԿԱՏՎԱԿԱՆ ՅԱՆՑՈՒՄ ՏՎՅԱԼՆԵՐԻ ԱՄԲՈՂՋԱԿԱՆՈՒԹՅԱՆ ԱՊԱՀՈՎՄԱՆ ՀԱՄԱԿԱՐԳԻ ՄՇԱԿՈՒՄԸ

ԱՄՓՈՓՈՒՄ

Մեքենայական ուսուցման (Machine Learning, ML) տեխնոլոգիայի ակտիվ զարգացումը և ֆիզիկական ցանցերի կոնվերգենցիան վիրտուալ ցանցերի հետ հանգեցրել են ցանցային ենթակառուցվածքի (ՑԵ) դեմ գրոհների լայնածավալ աճի, որոնք կարող են խախտել ՑԵ փոխանցվող տվյալների ամբողջականությունը: Ենթակառուցվածքային անվտանգության համակարգերի ճարտարապետները ՑԵ անվտանգությունը կառուցելիս, հիմնվում են սահմանված կանոնների և անվտանգության քաղաքականության վրա, որոնց թիվը սահմանափակ է: Ենթակառուցվածքային անվտանգության համակարգերի ճարտարապետների առջև ծառացած կարևոր խնդիրներից մեկը ՑԵ փոխանցվող տվյալների ամբողջականության ապահովումն է: Տարբեր մեխանիզմներ և մեթոդներ, որոնք ապահովում են տվյալների ամբողջականությունը ՑԵ, կորցնում են իրենց փոխազդման հնարավորությունը, երբ հարձակվողները կիրառում են ML հիմնված մեթոդներ: Հարկ է նշել, որ ML համակարգերի բոլոր բաղադրիչները, որոնք

մշակվել են տասնամյակների ընթացքում, համակարգչային տեխնոլոգիաների զարգացմանը զուգընթացները այսօր սկսում են հետ մնալ հարձակումներից այն դեպքերում, երբ չարագործը կիրառում է որպես գրոհի գեներատոր M2M&ML (Machine-to-Machine, M2M) տեխնոլոգիան: M2M&ML տեխնոլոգիայի հիման վրա կառուցված գրոհի կանոնները փոխելու և սովորելու արագությունն այնքան մեծ է, որ անվտանգության համակարգերը կառավարող ոչ մի ադմինիստրատոր չի կարողանա արձագանքել փոփոխվող իրավիճակին: ՑԵ-ում փոխանցվող տվյալների ամբողջականության դեմ գրոհները, որպես կանոն, սկսվում են ինչպես ՑԵ-ի ճարտարապետության, այնպես էլ ցանցային թրաֆիկի վիճակի վերլուծությամբ: Տեղեկատվական համակարգերում տվյալների ամբողջականությանը սպառնացող վտանգները չեզոքացնելու համար անհրաժեշտ է հարձակումը հայտնաբերել վաղ փուլում՝ մինչև դրա ակտիվ էսկալացիան: Հաշվի առնելով այս հանգամանքը՝ արդիական է դառնում ML-ով ներխուժման հայտնաբերման համակարգի մշակումը, որն ի վիճակի է հայտնաբերել փոխանցված տվյալների ամբողջականության խախտումները ցանցի միջուկի մակարդակում:

Աշխատանքի նպատակն է մշակել մեքենայական ուսուցումով ներխուժումների հայտնաբերման համակարգ, որը գործելու է ցանցի միջուկի մակարդակում և հայտնաբերելու է փոխանցվող տվյալների ամբողջականության խախտումները:

Հետազոտության առարկան հավանական գրոհները ցանցային ենթակառուցվածքում փոխանցվող տվյալների ամբողջականության դեմ, դրանց հայտնաբերման մեխանիզմները և արհեստական նեյրոնային ցանցերի օգտագործմամբ ներխուժման հայտնաբերման համակարգերը:

Ատենախոսության գիտական նորույթը:

- Առաջարկվել է վնասաբեր պոլիմորֆ ծրագրային ապահովման հայտնաբերման մեթոդ, որը գերազանցում է գոյություն ունեցող մեթոդների վրա հիմնված արդյունքները, ինչը հնարավորություն է տալիս ավելի արդյունավետ հայտնաբերել փոխանցվող տվյալների ամբողջականության խախտումը:
- Առաջարկվել է փոխանցվող տվյալների ամբողջականությունը խախտող վնասաբեր մթագրված և յուրացված ծրագրային ապահովման հայտնաբերման մեթոդ, որը գերազանցում է գոյություն ունեցող մեթոդների արդյունքները, ինչը հնարավորություն է տալիս ավելի արդյունավետ հայտնաբերել փոխանցվող տվյալների ամբողջականության խախտումը:
- Առաջարկվել է ՑԵ շրջանառվող տվյալների և ներխուժումների հայտնաբերման համակարգի ամբողջականության (այլ հավասար պայմանների դեպքում) խախտման դեպքում հայտնաբերման մեթոդ, որն

ապահովում է ներխուժման հայտնաբերումն ավելի արագ, քան գոյություն ունեցող մեթոդները:

- Ձեռք բերված գիտական արդյունքների հիման վրա, բարելավվել է «սպառնալիքների մոդելը» ՅԵ շրջանառվող տվյալների ամբողջականության բարձրացման շրջանակներում, իր բազմաչափորոշիչային, ստրատիֆիկացված բարելավման ժամանակ, դինամիկ վերակազմավորման ռեժիմում:

Աշխատանքի կառուցվածքը հետևյալն է՝ ներածություն, չորս գլուխ, եզրակացություն, 6 հավելված:

Ներածությունը նկարագրում է դիտարկվող ոլորտը, թեմայի արդիականությունը և աշխատանքում ստացված գիտական և գործնական արդյունքները:

Առաջին գլուխը ներկայացնում է հետազոտվող ոլորտի հիմնական խնդիրները, կոշտ տրամաբանությամբ և մեքենայական ուսուցումով ներխուժումների հայտնաբերման համակարգերի հիմնական չլուծված խնդիրները:

Երկրորդ գլուխը ներկայացնում է բազմաչափորոշիչային ցանցային ենթակառուցվածքի կայունության մոդելը, հավանական գրոհը ցանցային ենթակառուցվածքում փոխանցվող տվյալների ամբողջականության դեմ: Մշակված են հաշվեկարգեր և ծրագրային ապահովում մեքենայական ուսուցումով ներխուժումների հայտնաբերման համակարգի համար տվյալների հավաքածուների գեներացման և նախապատրաստման համար, որը հայտնաբերելու է փոխանցվող տվյալների ամբողջականության խախտումները:

Երրորդ գլուխը ներկայացնում է ռեկուրրենտ, փաթույթային, գեներատիվ-մրցակցային և պարկուճային նեյրոնային ցանցերի կիրառմամբ վնասաբեր ծրագրային ապահովման հայտնաբերման մեթոդները, որի կիրառմամբ չարագործը կարող է խախտել փոխանցվող տվյալների ամբողջականությունը: Մասնավորապես ուսումնասիրվել է «իրավիճակների մատրիցների» գաղափարի հիման վրա, վնասաբեր պոլիմորֆ ծրագրային ապահովման միջոցով, փոխանցվող տվյալների ամբողջականության խախտման հայտնաբերումը:

Չորրորդ գլուխը ներկայացնում է մշակված մեքենայական ուսուցումով ներխուժումների հայտնաբերման համակարգի ամբողջականության դեմ գրոհի դեպքում, ելքային տվյալների հավաստիության բարձրացման և արդյունավետության գնահատման հաշվեկարգերը և մշակված ծրագրային ապահովումը:

Եզրակացությունը ներկայացնում է կատարված աշխատանքի արդյունքները:

Ծավալը 24 էջ: Տպաքանակը 100:
ՀՀ ԳԱԱ ԻԱՊԻ կոմփյուտերային պոլիգրաֆիայի լաբորատորիա:
Երևան, Պ.Սևակի 1